N° d'ordre

Année 2012

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE E2M2

DIPLOME DE DOCTORAT
(arrêté du 7 août 2006)

soutenue publiquement le 11 décembre 2012

par

M. PRUNIER Jérôme

Echantillonnage individu-centré en génétique du paysage : Etude de l'impact de la fragmentation d'origine anthropique sur la dispersion du triton alpestre *Ichthyosaura alpestris*.

Directeur de thèse:	M. JOLY Pierre
Co-directeur:	M. POMPANON François
Co-encadrants:	M. LENA Jean-Paul
	M. KAUFMANN Bernard

JURY:	Mme STEVENS Virginie (rapporteur)
	M. BROQUET Thomas (rapporteur)
	M. MIAUD Claude
	M. VAN DYCK Hans
	M. JOLY Pierre
	M. POMPANON François
	M. LENA Jean-Paul
	M. MICHELOT Jean-Louis

À la mémoire de mon père

RESUME

Echantillonnage individu-centré en génétique du paysage : Etude de l'impact de la fragmentation d'origine anthropique sur la dispersion du triton alpestre *Ichthyosaura alpestris.*

Les activités d'origine anthropique entraînent des modifications profondes du paysage : dans ce contexte, le succès de dispersion revêt une grande importance pour la persistance à long terme des populations animales et végétales. La génétique du paysage est une discipline récente qui permet de détecter et de modéliser les flux de dispersion de manière indirecte par le biais de l'identification des flux de gènes dans l'espace. Bien qu'il ait été suggéré de longue date qu'un échantillonnage individu-centré pouvait permettre de s'affranchir de certains biais liés à une délimitation *a priori* des populations, les études portant sur des organismes présentant une distribution spatiale discontinue restent malgré tout le plus souvent basées sur un échantillonnage réalisé à l'échelle populationnelle. C'est sur cette thématique qu'ont porté mes travaux de thèse. Des résultats obtenus par simulations montrent qu'augmenter le nombre de points échantillonnés dans l'espace au détriment du nombre d'individus échantillonnés par agrégat peut permettre une meilleure détection de l'influence de la matrice sur les flux de gènes, quelque soit le régime de dispersion et le contexte paysager considérés. Appliqué à l'étude de la dispersion du triton alpestre *Ichthyosaura alpestris*, d'une part en paysage fragmenté par des infrastructures linéaires de transport (Isolation par barrières aux flux de gènes IBB) et d'autre part en paysage agricole (Isolation par résistance de la matrice paysagère IBR), l'échantillonnage individu-centré apparaît comme une alternative flexible et efficace à l'échantillonnage populationnel classique.

MOTS-CLEFS

Amphibiens ; Conservation ; Fragmentation ; Flux de gènes ; Génétique du paysage ; Unité d'échantillonnage.

SUMMARY

Individual-based sampling scheme in landscape genetics: assessing the impacts of anthropogenic fragmentation on dispersal patterns in the alpine newt *Ichthyosaura alpestris.*

Anthropogenic activities lead to profound landscape alterations: dispersal success thus holds a paramount importance for long-term persistence of populations. Landscape genetics is a recent discipline aiming to detect and describe dispersal patterns through indirect estimations of gene flow. Although the use of an individual-based sampling scheme has been proposed for a long time to get round the issue of an *a priori* delimitation of population boundaries, most landscape genetic studies are still based on a population-based sampling scheme. Results from simulations showed that decreasing the sampling coverage of individuals within populations in favour of a better sampling coverage of aggregates through space allowed a better detection of the impacts of landscape on gene flow, whatever the individual dispersal behaviour or the landscape configuration. When considering the dispersal patterns in the alpine newt *Ichthyosaura alpestris*, both in the vicinity of large transport infrastructures (Isolation-by-barrier IBB) and in an agricultural landscape (Isolation-by-landscape-resistance IBR), the individual-based sampling scheme proved to be a flexible and efficient methodological alternative to the more conventional population-based sampling scheme.

KEYWORDS

Amphibians ; Conservation ; Fragmentation ; Gene flow ; Landscape genetics ; Sampling units.

Organismes de rattachements et financements

Ces travaux ont été réalisés dans le cadre d'un contrat CIFRE (Convention Industrielle de Formation par la Recherche) au sein :

- Du *Laboratoire d'Ecologie des Hydrosystèmes Naturels et Anthropisés* (*LEHNA*), Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.
- Du *bureau d'études Ecosphère, Agence Sud-Est*
  16 rue Garon 69560 Sainte-Colombe, France.

Ils n'auraient pu être menés à bien sans le soutien :

- De l'*Association Nationale de la Recherche et de la Technologie* (ANRT).
- De *Réseau Ferré de France* (RFF).
- D'*Autoroutes Paris-Rhin-Rhône* (APRR).

REMERCIEMENTS

Après quelques mois passés à radio-pister des crapauds sur le site de la réserve naturelle de l'Etang du Grand-Lemps dans le cadre de mon stage de master 2, je m'apprêtais à quitter définitivement l'université... mais à la sortie de mon oral de fin d'étude, Pierre Joly est venu me proposer de le rencontrer « pour discuter de la suite »... une suite ?.... C'était en septembre 2006.

Six ans après, je me retrouve à rédiger ces quelques lignes. Que de choses se sont passées depuis ! Pour faire très (très) bref : deux ans et demi pour monter le projet et trouver les partenaires nécessaires à sa réalisation, puis trois ans et demi de doctorat. Cette grande épopée n'aurait jamais pu être menée à bien sans l'accompagnement d'une foule de personne, à qui je destine tous mes remerciements.

En tout premier lieu, je tiens à remercier l'ensemble des membres de mon jury de thèse pour l'intérêt qu'ils auront porté à mes travaux, et tout particulièrement Virginie Stevens et Thomas Broquet, tous deux rapporteurs, qui auront pris le temps de relire, digérer et commenter ce manuscrit.

Je remercie infiniment Pierre Joly, mon directeur de thèse, qui m'a justement proposé cette « suite » qui fut à l'origine de tout, François Pompanon, mon co-directeur de thèse, ainsi que Jean-Louis Michelot et Jean-Christophe Kovacs, qui m'ont accueilli dans l'univers écosphérien : tous m'ont accordé leur confiance, j'espère avoir été à la hauteur.

Je tiens également à remercier chaleureusement Anne Guerrero (Réseau Ferré de France), Etienne Cuenot (Autoroutes Paris-Rhin-Rhône) et Jean Carsignol (CETE de l'Est), qui ont largement contribué à modeler ce projet. Sans eux, la « suite » serait tout simplement restée dans les cartons ! Si les résultats de mes travaux ne correspondent pas tout à fait à ce qui avait été initialement envisagé, j'espère que ce projet leur aura tout de même apporté satisfaction.

Un immense merci à Jean-Paul Léna et à Bernard Kaufmann, qui se sont subrepticement retrouvés embrigadés dans mes « tritonades » : leur aide et leur soutien me furent précieux, notamment dans mes instants de doutes. L'investissement (sans limite) dont Jean-Paul est capable de faire preuve a été décisif pour l'accomplissement de ce travail : une fructueuse collaboration que j'espère avoir l'occasion de renouveler à l'avenir ! Merci également aux nombreux autres chercheurs dont j'ai croisé la route, qui se sont intéressés à mon sujet à un moment ou à un autre et m'ont souvent aiguillé de manière pertinente : je pense notamment à Pascal Allemand qui m'a ouvert les portes de l'univers Matlab (et on s'y sent bien), Serge Fenet et sa force de frappe à l'IN2P3, Damien Picard, Anne-Béatrice Dufour, Thibaut Jombart, Erin Landguth, Jean-Christophe Foltête, Patrick Giraudoux, sans oublier Claude Miaud et Stéphanie Manel, membres de mon comité de pilotage.

Un coup de chapeau à tous ceux qui ont volontairement donné de leur personne, tant sur le terrain que derrière une paillasse ou un écran, pour que ce projet puisse avancer : Maxime Clasquin (de chez « thésard & stagiaire »), Florent Revollon, Elodie Calonnier, Amandine Plaisance,

Céline De Fazio, Grégoire Perez, Florent Ribette, Jennifer Mangin, Julien Gauthier, Benjamin Virone, Nicolas Farnier, Lara Konecny et Odile Grolet. Merci également à Daniel Sirugue et Nicolas Varenguin, de la Société d'Histoire Naturelle d'Autun, et à Samuel Gomez, du Conservatoire des Sites Naturels Bourguignons : les données cartographiques et naturalistes qu'ils m'ont fournies se sont révélées indispensables pour la mise en place de mon plan d'échantillonnage.

Je remercie tous mes collègues de Vienne, qui, malgré ma présence intermittente, m'ont laissé me faire une petite place douillette à Ecosphère : notamment Laurent Simon, Elodie Calonnier (« trop mauvaise » mais finalement remerciée deux fois !), Karine Ragouvin, François Caron, Cécile Barbier et Yvain Dubois, avec qui j'aurai passé des moments très agréables à papoter de nature, de musique ou de Georges, bref de tout et de rien, tant au boulot qu'à l'extérieur. Je remercie de la même manière les collègues du laboratoire LEHNA, permanents et étudiants, pour nos échanges parfois savants, souvent navrants (j'en assume une part de responsabilité) mais toujours plaisants : dédicace spéciale à messieurs Brepson (« Garobrepson ! »), Wardziak et Troïanowski, à Cécile Capderrey, David Eme, Jérôme Gippet, Karine Salin, Loïc Teulier, Emilien Luquet (mon p'tit José), Sandrine Plenet, Nadjette Houriez (une urgence ? une galère administrative de dernière minute ? Nadjette !), Yann Voituron, Nathalie Mondy et Thierry Lengagne.

Je me dois également de saluer toutes les personnes rencontrées au cours des quatre mois passés à arpenter la Bourgogne à la recherche de tritons : beaucoup ont fait preuve de curiosité et de gentillesse, certains se sont montré plus méfiants, voire franchement hostiles... mais il ne me reste au final que de bons souvenirs (ainsi que quelques tics de langage, je dois bien l'avouer !). Je ne suis pas prêt d'oublier Marcel Morin (« Il est très bien »), Bernur et son chien Blrrp (en fin de journée), ainsi que les bougons des Soupirons (« M'en fous, moi j'les mange pas »).

Je remercie chaleureusement toutes les personnes que j'ai pu rencontrer avant de me lancer dans cette aventure, et qui m'ont apporté beaucoup, peut-être sans qu'elles le sachent vraiment : je pense à Agnès Janin et Grégory Maillet, qui, chacun à sa façon, m'a initié à l'écologie du paysage, à l'herpétologie, aux joies de la recherche et du terrain ; à Tim Beroud, qui m'a entraîné à son insu dans la spirale infernale de l'ornithologie, une passion qui m'a accompagné pendant toutes ces années et qui n'est pas prête de s'éteindre; enfin et surtout à Jean-Pierre Chalabert et sa famille, qui m'ont accueilli à bras ouverts en leur demeure, même dans les moments difficiles.

Merci également à toutes les personnes qui ont patiemment écouté mes histoires de tritons et d'autoroutes, merci à Yohann Piga pour ses illustrations qui m'auront accompagné jusqu'au bout, et enfin merci à toutes les personnes que j'aurais pu oublier de remercier (!) mais qui se reconnaîtront.

Je remercie enfin Claire-Lise et ma famille, de tout mon cœur, pour leur soutien sans faille au cours de toutes ces années, malgré les épreuves que nous avons dues affronter ensemble.

M



Mes remerciements ne seraient pas complets sans une pensée attendrie à l'égard des 1080 tritons alpestres qui se sont docilement prêtés à l'écouvillonnage.

Sommaire

INTRODUCTION 11

L'ETUDE DE LA VARIABILITE GENETIQUE AU SEIN DES PEUPLEMENTS STRUCTURES 13
LA NOTION DE CONNECTIVITE DU PAYSAGE 17
UNE DISCIPLINE RECENTE: LA GENETIQUE DU PAYSAGE 19
QUELLE UNITE D'ECHANTILLONNAGE POUR L'ESTIMATION INDIRECTE DES FLUX DE GENES ? 21
ECHANTILLONNAGE INDIVIDU-CENTRE : SIMULATIONS 24
... ET ETUDES DE CAS 25

CHAPITRE 1 27

CHAPITRE 2 63

CHAPITRE 3 85

CHAPITRE 4 117

CONCLUSIONS ET PERSPECTIVES 145

APPORT DE L'ECHANTILLONNAGE INDIVIDU-CENTRE EN GENETIQUE DU PAYSAGE 146
Plus de flexibilité en génétique du paysage 146
Maximisation du rapport signal / bruit (signal-to-noise ratio) 149
PRISE EN COMPTE DE LA NON-INDEPENDANCE DES DONNEES 152
Non-indépendance des données appariées 152
Colinéarité des variables paysagères 156
VERS UNE MEILLEURE INTEGRATION DES DIFFERENTES APPROCHES EN GENETIQUE DU PAYSAGE 157
Intégration de données démographiques 157
Concilier les avantages des approches individu-centrées et populationnelles 160
L'APPROCHE INDIVIDU-CENTREE ... AU-DELA DES QUESTIONS DE CONSERVATION ? 163

BIBLIOGRAPHIE 165

INTRODUCTION

La distribution spatiale de tous les organismes animaux et végétaux s'avère discontinue, à une échelle plus ou moins large d'observation (Hanski & Gilpin 1991). Cette structuration des peuplements (voir *Encadré 1*) est intimement liée à la distribution spatiale des ressources exploitées par les organismes (With 1997; Swihart *et al.* 2003). Or cette distribution est susceptible de changer au cours du temps : en plus de l'évolution naturelle des habitats (cycles de perturbations et de successions écologiques), l'évolution des pratiques agricoles et l'artificialisation des surfaces, conséquences de l'accroissement exponentiel des activités humaines, ont largement contribué à une modification profonde des paysages au cours des dernières décennies (Fahrig 2003; Fahrig *et al.* 2011). La diminution de la quantité de ressources disponibles pour les organismes (par perte d'habitat) entraînant une diminution de la taille des populations, de nombreuses espèces risquent de se trouver - ou se trouvent déjà - engagées dans un « vortex d'extinction » (Gilpin & Soulé 1986; Fagan & Holmes 2006; Palomares *et al.* 2012). Il s'agit d'un processus par lequel la diminution de la taille d'une population entraîne une

Encadré 1 : La notion de peuplement

La notion de peuplement renvoie, en écologie, à un groupe d'espèces partageant une même ressource (on parle également de communauté ou de guilde). Dans ce manuscrit, j'utiliserai ce terme dans sa définition française classique (« *manière dont un territoire est peuplé* », dictionnaire *Le petit Robert 2009*), faisant simplement référence à l'occupation de l'espace par une espèce. Les individus pouvant se répartir de manière plus ou moins continue dans l'espace, parler de « structuration des populations » serait trop réducteur dans le contexte de mes travaux. Inversement, un « peuplement structuré » ne renvoie pas uniquement à un groupe de populations bien différenciées (Figure M1).

détérioration des caractéristiques démographiques (p. ex. baisse du succès reproducteur ou de la survie des individus) et génétiques (p. ex. hausses de la consanguinité ou de la dérive génétique entraînant la perte de diversité génétique et l'accumulation de mutations délétères), détérioration qui à son tour entraîne une baisse d'effectifs, jusqu'à un seuil critique pour la viabilité des populations. Ce cercle vicieux se trouve souvent aggravé par l'éloignement croissant des ressources les unes par rapport aux autres (fragmentation), qui augmente les coûts de la recherche de ressources alimentaires (*foraging*) et réduit les échanges d'individus (et d'allèles) entre populations. N'épargnant aucun groupe taxonomique, perte d'habitat et fragmentation sont ainsi rapidement devenues un enjeu majeur en biologie de la conservation (Fahrig & Merriam 1994; Wiens 1995; Sala *et al.* 2000; Fischer & Lindenmayer 2007). De nombreuses approches peuvent être utilisées pour caractériser l'impact de ces deux processus sur la distribution spatiale des organismes et sur leurs capacités de déplacement : c'est ainsi le cas des études basées sur des suivis démographiques (p. ex. Schtickzelle *et al.* 2002; Perret *et al.* 2003; Janin *et al.* 2009), comportementaux (p. ex. Schtickzelle & Baguette 2003; Stevens *et al.* 2006a) ou encore physiologiques (p. ex. Janin *et al.* 2011; Janin *et al.* 2012). Un nouveau champ d'investigation s'est toutefois développé au cours des trente dernières années, lié à la démocratisation de l'accès aux approches moléculaires et aux outils statistiques associés (Manel *et al.* 2003; Holderegger & Wagner 2008) : les approches génétiques tiennent désormais une place à part entière en biologie de la conservation (Frankham *et al.* 2002). Des modèles les plus anciens en génétique des populations (Wright 1931) à l'avènement de la génétique du paysage (Haila 2002; Manel *et al.* 2003), l'étude de la dynamique et de la persistance des peuplements en paysage fragmenté n'a eu de cesse de progresser, tant sur le plan théorique qu'appliqué.

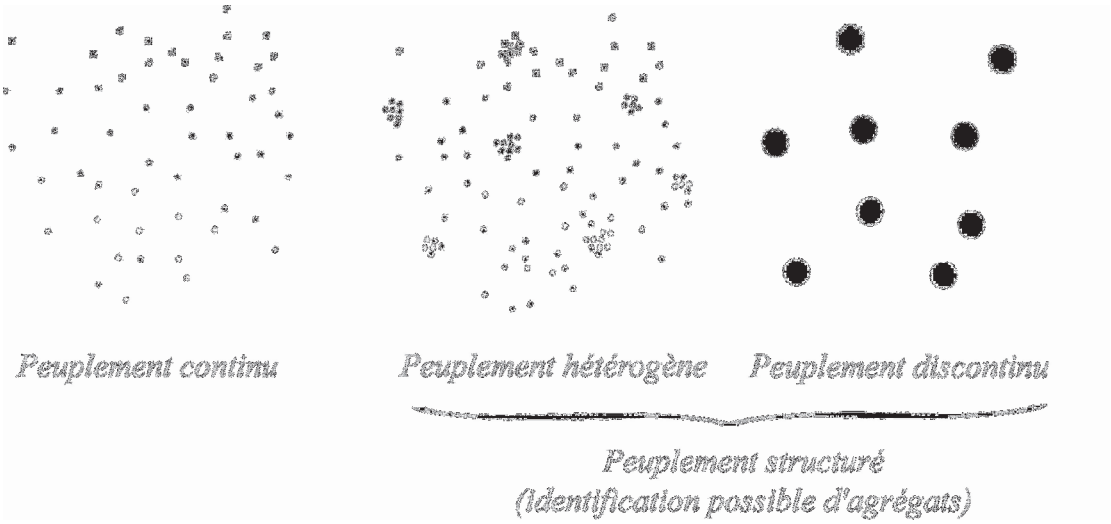


Figure M1 : La notion de peuplement renvoie, dans ce manuscrit, à l'occupation de l'espace par une espèce : un peuplement peut être continu ou structuré en agrégats d'individus plus ou moins bien délimités.

L'étude de la variabilité génétique au sein des peuplements structurés

Le maintien de la diversité génétique des peuplements est un objectif majeur en biologie de la conservation (Frankham 2005). La diversité génétique actuelle des populations conditionne en effet les capacités adaptatives des organismes aux conditions environnementales de demain: la fréquence des mutations naturelles étant souvent trop faible pour permettre aux populations de s'adapter aux variations environnementales liées aux activités humaines, c'est dans un pool génétique le plus diversifié possible qu'elles devront pouvoir puiser d'éventuelles solutions adaptatives. Or la réduction des effectifs de populations, souvent corrélée à la réduction de la disponibilité des ressources, entraîne, par dérive génétique ou dépression de consanguinité, une érosion de cette diversité génétique (Hedrick 2001; Frankham *et al.* 2002; Frankham 2005). La circulation des organismes entre populations pouvant permettre de compenser cette érosion en

assurant un meilleur brassage des gènes, l'étude de l'influence des flux de gènes sur le maintien de la diversité génétique est devenue un enjeu important de la génétique des populations.

Les premiers modèles théoriques de structuration génétique des populations, intégrant plusieurs populations connectées par échanges d'individus, sont les modèles en îles (*infinite and finite island models* ; Wright 1931; Latter 1973). Ils reposent sur un nombre fini ou infini de populations distinctes mais de taille identique, échangeant des individus à des taux identiques quelle que soit leur localisation dans l'espace. D'autres modèles théoriques de structuration génétique (ou de structuration démographique adaptés à la génétique) ont rapidement été proposés pour une modélisation plus réaliste des flux de gènes entre populations (Fahrig & Merriam 1994). Parmi ceux-ci : le modèle en pas japonais (*Stepping-stone model*, Kimura 1953), considérant que les flux de gènes ne peuvent avoir lieu qu'entre populations voisines ; le modèle île-continent (*island-mainland model* ; Boorman & Levitt 1973), dans lequel une population de taille infinie est la principale source de propagules en direction de populations de taille moindre ; ou encore le modèle sources-puits (*source-sink model* ; Pulliam 1988), considérant certaines populations comme sources de propagules à destination de populations-puits dont la persistance dépend principalement de cette immigration.

C'est toutefois le concept de métapopulation, proposé par Levins en 1969, qui semble s'être imposé comme paradigme en biologie de la conservation (Fahrig & Merriam 1994; Lewis *et al.* 1997; With 1997; Seppa & Laurila 1999; Haila 2002; Baguette & Schtickzelle 2003; Baguette & Stevens 2003), fournissant un cadre théorique fondamental pour l'étude tant démographique que génétique de nombreux organismes à peuplement structuré (p. ex. Olivieri *et al.* 1995; Lewis *et al.* 1997; Seppa & Laurila 1999; Baguette *et al.* 2000; Charbonnel *et al.* 2002b; Frank & Wissel 2002; Baguette & Schtickzelle 2003; Baguette & Stevens 2003). Au sens strict, une métapopulation peut être décrite comme une *population de populations* (Hanski & Gilpin 1991), c'est-à-dire un ensemble de sous-populations interagissant par des échanges d'individus dispersant entre fragments d'habitat favorable (ou ressources). Le maintien d'une métapopulation est un processus dynamique d'extinctions-colonisations, intimement lié aux capacités de déplacements des individus entre sous-populations: les sous-populations peuvent se maintenir si, au niveau d'un fragment d'habitat favorable, le taux d'extinction local reste

inférieur au taux de recolonisation (*rescue effect*), tandis que les fragments laissés vacants suite à une extinction locale peuvent être à la longue recolonisés par des individus dispersants (Hanski & Gilpin 1991; With 1997; Hanski 1998; Baguette & Schtickzelle 2003). La théorie des métapopulations repose sur la notion d'isolement par la distance (*isolation-by-distance* IBD; Wright 1943; Hanski & Gilpin 1991; Slatkin 1993): les taux de dispersion entre populations dépendent du taux de recrutement au sein de chaque sous-population mais également de leur éloignement relatif (Harrison 1991; Hanski 1998; Frank & Wissel 2002). Le concept de métapopulation a ainsi été adopté en génétique des populations comme un modèle de choix pour l'étude de la variabilité génétique intra- et interpopulationnelle (Slatkin 1977; Whitlock & Barton 1997; Whitlock & McCauley 1999; Pannell & Charlesworth 2000; Holsinger & Weir 2009; van Nouhuys 2009), les notions conjointes d'IBD et de dynamique d'extinction-colonisation venant enrichir les modèles classiques de structuration génétique des populations tels que les modèles en îles.

Tous ces modèles théoriques de structuration des peuplements peuvent être utilisés pour interpréter d'un point de vue biologique, sous forme par exemple de taux de dispersion entre populations ou de tailles efficaces, des mesures statistiques (ou *summary statistics* ; Lowe *et al.* 2004) obtenues à partir de données génétiques (Seppa & Laurila 1999). Ces mesures sont très souvent calculées à partir des fréquences alléliques observées en chaque population et permettent de caractériser un écart à l'équilibre d'Hardy-Weinberg (voir *Encadré 2*), écart pouvant notamment s'expliquer par une subdivision des peuplements en populations discrètes (effet Wahlund) : c'est ainsi le cas des *F-statistics* (Fst, Fis et Fit ; Wright 1951) et de leurs dérivés (p. ex. Weir & Cockerham 1984; Slatkin 1995). Le panel d'outils disponibles pour les généticiens s'est également formidablement enrichi ces dernières années, avec l'apparition de méthodes permettant de s'affranchir de certaines

Encadré 2 : L'équilibre d'Hardy-Weinberg

L'équilibre d'Hardy-Weinberg (1908) est une notion centrale en génétique des populations : elle permet de prédire l'évolution des fréquences génotypiques d'une génération à l'autre, les fréquences alléliques restant stable au cours du temps dans une population de taille infinie, panmictique (avec appariement aléatoire), en absence de tout processus de dérive, mutation ou de flux de gènes.

contraintes liées à l'utilisation des modèles théoriques de structuration génétique des peuplements (telles que l'hypothèse d'un équilibre dynamique entre dérive génétique et flux de gènes, à la base de l'utilisation des *F-statistics* ; Hutchison & Templeton 1999; Whitlock & McCauley 1999; Lowe *et al.* 2004; Faubet *et al.* 2007). C'est notamment le cas des approches bayesiennes (Wilson & Rannala 2003; Beaumont & Rannala 2004), permettant d'estimer la probabilité d'un paramètre d'intérêt (probabilité d'assignement des génotypes à une population, taux de migration, etc.) sur la base de données observées et d'hypothèses *a priori* concernant les processus évolutifs en jeu sur la zone d'étude. Les approches bayesiennes ont notamment permis la mise au point de méthodes de partitionnement des données génétiques (*Bayesian clustering methods* ; p. ex. Pritchard *et al.* 2000; Francois *et al.* 2006), largement employées en écologie moléculaire et venant compléter les méthodes plus classiques d'assignement des individus à leur population d'origine (Manel *et al.* 2005). Les approches de coalescence (Donnelly & Tavare 1995; Kingman 2000) portent quant à elles sur l'étude de la transmission (ou généalogie) des allèles au cours du temps et autorisent la prise en compte d'évènements historiques (p. ex. effets fondateurs, *bottlenecks*) pour l'estimation des paramètres démographiques à l'origine des structurations génétiques observées (Marko & Hart 2011).

Toutes ces approches, lorsqu'elles sont appliquées à l'étude de la structuration génétique d'un peuplement discontinu, ont un point commun : elles sont très généralement basées sur un échantillonnage le plus complet possible des populations étudiées (Paetkau *et al.* 2004; Kalinowski 2005; Broquet & Petit 2009; Schwartz & McKelvey 2009). Cet exhaustivité d'échantillonnage autorise par exemple l'estimation non biaisée des fréquences alléliques observées au niveau de chaque population, ou permet de disposer du matériel génétique nécessaire à la détermination de l'origine des individus dispersants dans le cadre des tests d'assignement ou des analyses de parenté (Holderegger & Wagner 2008).

La notion de connectivité du paysage

Le concept de métapopulation, ainsi que les autres modèles théoriques de structuration des peuplements, reposent sur une vision binaire de la matrice paysagère (Harrison 1991; Wiens 1995; With 1997): le paysage est constitué de fragments d'habitat favorable, "îlots" accueillant les sous-populations, au sein d'une matrice (ou "océan") de non-habitat (With 1997; Haila 2002). Cette vision de la matrice paysagère, qui a notamment dicté la mise en place de réserves intégrales de biodiversité à partir des années 70 (Wiens 1995; van Nouhuys 2009), a longtemps prédominé en raison de la corrélation négative, attendue et très souvent observée, entre taille ou proximité des fragments d'habitat et risque d'extinction locale, sans que ne soit vraiment questionnée l'existence d'une relation de cause à effet réelle entre ces phénomènes (Haila 2002). Cette vision a aujourd'hui montré ses limites et a cédé le pas à ce qui est devenu une discipline à part entière, visant à étudier les interactions existant entre processus écologiques et patrons d'organisation spatiale du paysage: l'écologie du paysage (Turner *et al.* 2001). Les fragments d'habitats ne sont plus vus comme des îlots de ressources mais comme faisant partie intégrante d'un paysage hétérogène et complexe (With 1997; Fahrig *et al.* 2011), qui peut être caractérisé d'une part par sa composition, c'est-à-dire la nature et la proportion relative des différents habitats qui le composent, et d'autre part sa configuration, c'est-à-dire le degré de fragmentation des habitats et l'agencement les uns par rapport aux autres de ces fragments d'habitats (Fahrig & Merriam 1994; Fahrig 2003). Dans ce contexte, la distance seule ne suffit pas à expliquer tous les phénomènes de recolonisation observés au sein d'une métapopulation et la persistance des populations est intimement liée à la facilité avec laquelle les organismes se déplacent au sein de la matrice (Joly *et al.* 2001; Ricketts 2001; Chardon *et al.* 2003; Revilla *et al.* 2004; Stevens *et al.* 2004; Bender & Fahrig 2005). La connectivité entre fragments d'habitats n'est donc plus uniquement décrite d'un point de vue structurel (liens physiques entre fragments), mais également fonctionnel (Taylor *et al.* 1993; Fahrig & Merriam 1994; Tischendorf & Fahrig 2000b), selon le comportement des organismes vis à vis des différents éléments paysagers susceptibles d'être rencontrés au sein du paysage lors de

leurs déplacements (Dunning *et al.* 1992; Wiens 1995; Tischendorf & Fahrig 2000a; Van Dyck & Baguette 2005; Nathan *et al.* 2008).

La notion de connectivité fonctionnelle est venue raffiner le concept de dynamique métapopulationnelle (Hanski & Gaggiotti 2004). Le maintien d'une sous-population dépend notamment des possibilités d'accès des organismes à des fragments d'habitats complémentaires nécessaires à leur cycle de vie (complémentation) ou à des fragments d'habitats de substitution en cas d'altération du paysage (supplémentation ; Dunning *et al.* 1992). A une échelle plus large, des épisodes de dispersion, influencés par la configuration du paysage, assurent le renforcement de populations en déclin et la colonisation (ou la recolonisation) de fragments d'habitats favorables laissés vacants (Hanski & Gilpin 1991; Van Dyck & Baguette 2005; Semlitsch 2008). De la simple estimation d'occurrence ou d'abondance (e.g. Cushman 2006; Janin *et al.* 2009) aux suivis par capture-marquage-recapture (e.g. Lewis *et al.* 1997; Mousson *et al.* 1999; Schtickzelle *et al.* 2002; Perret *et al.* 2003), en passant par la télémétrie (e.g. Jehle & Arntzen 2000; Driezen *et al.* 2007), les expérimentations comportementales en mésocosme (e.g. Rosenberg *et al.* 1998; Boudjemadi *et al.* 1999; Stevens *et al.* 2004; Stevens *et al.* 2006a) ou encore les mesures d'indicateurs physiologiques de stress (e.g. Janin *et al.* 2011; Janin *et al.* 2012), l'écologie du paysage a vu se développer une large gamme d'approches permettant d'étudier la réponse fonctionnelle des organismes à la fragmentation du paysage. L'une des principales avancées méthodologiques provient notamment de la théorie de la percolation (Stauffer & Aharony 1991; Kindlmann & Burel 2008), à l'origine des méthodes dites de distance de coût (Adriaensen *et al.* 2003), qui a permis de répondre à la nécessité d'évaluer et de prédire l'impact de la fragmentation sur la connectivité fonctionnelle des habitats (Ray *et al.* 2002; Cushman *et al.* 2006; Stevens *et al.* 2006b; Wang *et al.* 2008; Janin *et al.* 2009).

L'hétérogénéité des paysages ne peut être pleinement appréhendée que par le biais d'une couverture d'échantillonnage optimale de l'espace : en écologie du paysage, les données doivent donc être récoltées de manière à disposer d'une représentation satisfaisante des éléments (ou des gradients) paysagers susceptibles d'impacter les processus biologiques, et ce à différentes échelles spatiales (Hirzel & Guisan 2002; Bennett *et al.* 2006; Janin *et al.* 2009; Fahrig *et al.* 2011).

D'un point de vue appliqué, la notion de connectivité fonctionnelle a fait évoluer les stratégies de conservation de la biodiversité, longtemps restées cantonnées à la protection de sites isolés abritant des espèces ou des milieux vulnérables, rares ou menacés. Cette logique a laissé place à une stratégie plus globale qui s'appuie sur la conception de réseaux d'espaces protégés (Allag-Dhuisme *et al.* 2010) : enrayer la perte de biodiversité passe en effet par la préservation des noyaux de biodiversité, mais également par le maintien des possibilités de déplacement de la faune et de la flore entre ces espaces par la préservation ou la création de continuités écologiques (ou corridors ; Brooker *et al.* 1999; Mech & Hallett 2001; Kirchner *et al.* 2003; Pinto & Keitt 2009). En France, cette stratégie de conservation de la biodiversité passe par un outil d'aménagement du territoire appelé « Trame verte et bleue » : il s'agit d'une mesure prioritaire demandée par le groupe « Préserver la biodiversité et les ressources naturelles » du Grenelle de l'environnement (2007).

Une discipline récente: la génétique du paysage

Le suivi direct du déplacement effectif des propagules au sein de la matrice par télémétrie ou par capture-marquage-recapture (Mousson *et al.* 1999; Perret *et al.* 2003; Driezen *et al.* 2007) peut s'avérer particulièrement ardu (Manel *et al.* 2003; Broquet *et al.* 2006b; Storfer *et al.* 2007; Murphy *et al.* 2010b; Spear *et al.* 2010). Au contraire, les développement récents de nouvelles techniques de séquençage automatique et l'accès à des marqueurs génétiques neutres hautement polymorphes (Holderegger & Wagner 2008; Abdelkrim *et al.* 2009; Storfer *et al.* 2009) ont autorisé le développement d'une grande variété d'outils génétiques permettant une estimation indirecte des flux de dispersions entre populations (Vos *et al.* 2001; Vandewoestijne & Baguette 2004; Storfer *et al.* 2007; Broquet & Petit 2009), au sein d'une discipline en pleine expansion: la génétique du paysage (Manel *et al.* 2003). A l'interface entre écologie du paysage et génétique des populations, la génétique du paysage s'attache à étudier les flux de gènes et les patrons de structuration génétique des populations au regard de la composition et de la configuration du paysage (Manel *et al.* 2003; Storfer *et al.* 2007), constituant par là

un axe de recherche prometteur en biologie de la conservation (Segelbacher *et al.* 2010). Reposant sur la notion de connectivité fonctionnelle (Keyghobadi 2007; Holderegger & Wagner 2008), cette discipline s'avère un outil de choix pour quantifier l'impact des variables paysagères et leur configuration sur les patrons de variation génétique, identifier les barrières aux flux de gènes ou encore déterminer l'échelle optimale de conservation de la diversité génétique au sein des populations (Hedrick 2001; Manel *et al.* 2003; Storfer *et al.* 2007).

Deux grandes approches sont couramment employées en génétique du paysage. Les *overlay methods,* approches que l'on pourrait qualifier de « *ad hoc* » (Anderson *et al.* 2010), sont directement issue de la génétique des populations. Elles reposent sur l'emploi d'un large panel d'outils permettant de caractériser la structure génétique des populations (e.g. Pritchard *et al.* 2000; Chen *et al.* 2007; Guillot 2008; Jombart *et al.* 2008), d'identifier des discontinuités génétiques (p. ex. Crida & Manel 2007; Kelly *et al.* 2010; Safner *et al.* 2011b) ou des évènements récents de dispersion (Sork *et al.* 1999; Jones & Ardren 2003; Manel *et al.* 2005). Néanmoins, même si certaines intègrent une information spatiale d'isolement par la distance via l'emploi de graphes de voisinage entre sites d'échantillonnage génétique (p. ex. Chen *et al.* 2007; Jombart *et al.* 2008), ces approches mettent principalement l'accent sur l'exploitation des données génétiques et ne permettent qu'une comparaison *a posteriori* entre structuration génétique et caractéristiques paysagères (p. ex. Coulon *et al.* 2006; Angelone & Holderegger 2009).

Au contraire, les approches de type « corrélatif » (Seppa & Laurila 1999; Broquet *et al.* 2006b; Cushman *et al.* 2006; Manel *et al.* 2009; Anderson *et al.* 2010; Emaresi *et al.* 2011) permettent une utilisation simultanée des données génétiques et paysagères : elles reposent sur l'utilisation de différentes approches statistiques telles que corrélations matricielles (Cushman *et al.* 2006; Holderegger & Wagner 2008; Legendre & Fortin 2010) , régressions linéaires simples (Goldberg & Waits 2010; Emaresi *et al.* 2011) ou encore modèles mixtes (Selkoe *et al.* 2010) pour tester des hypothèses concernant l'influence directe de la composition et de la configuration du paysage sur les flux de gènes. Dans ce cas, les flux de gènes sont estimés de manière indirecte par des mesures de distances génétiques (Slatkin 1985) : différentiation génétique entre populations (p. ex. Fst ; Rousset 1997; Stevens *et al.* 2006b; Selkoe *et al.* 2010; Emaresi *et al.* 2011) ou dissimilarité entre génotypes individuels (également, degré d'apparentement (ou

relatedness) entre individus; Legendre & Legendre 1998; Rousset 2000; Broquet *et al.* 2006b; Cushman *et al.* 2006; Elmer *et al.* 2007).

Quelle unité d'échantillonnage pour l'estimation indirecte des flux de gènes ?

La structuration génétique d'un peuplement dépend à la fois de la répartition spatiale des individus dans l'espace et de leur capacité de déplacement au sein du paysage : il s'avère donc crucial d'adopter un plan d'échantillonnage permettant de capter au mieux cette variabilité génétique, mais également la variabilité paysagère, c'est-à-dire la diversité des structures paysagères rencontrées, aux échelles de temps et d'espace considérées (Bennett *et al.* 2006; Storfer *et al.* 2007; Anderson *et al.* 2010). Si l'importance de la distribution spatiale des points d'échantillonnage a été soulignée par plusieurs études (Schwartz & McKelvey 2009; Anderson *et al.* 2010; Jaquiéry *et al.* 2011), aucune ne s'est réellement intéressée à l'influence de l'unité d'échantillonnage dans les approches corrélatives. Dans le cas d'une distribution continue des organismes, les estimations indirectes de flux de gènes passent naturellement par un échantillonnage individu-centré et la mesure d'un degré d'apparentement entre individus (Coulon *et al.* 2004; Broquet *et al.* 2006b; Cushman *et al.* 2006). En revanche, dans le cas d'une distribution discontinue des organismes, les flux de gènes sont très généralement estimés par le biais de mesures de différentiation entre « populations » (Broquet & Petit 2009; Anderson *et al.* 2010). Cette tendance systématique, en présence d'agrégats, à réaliser un échantillonnage populationnel à partir de populations identifiées *a priori* peut s'expliquer par la facilité d'échantillonnage offerte par de fortes densités locales d'individus (Angelone & Holderegger 2009), mais également par un héritage direct de la génétique des populations, du paradigme des métapopulations et autres modèles en îles : tous les agrégats d'individus facilement identifiables dans l'espace sont traditionnellement vus comme des populations à part entière, génétiquement bien différenciées les unes des autres, et échantillonnés comme tels (Hanski & Gilpin 1991; Waples & Gaggiotti 2006). Cette approche soulève toutefois plusieurs problèmes.

Lorsque l'hétérogénéité de la matrice paysagère favorise localement la dispersion, les agrégats d'individus risquent de ne refléter que très imparfaitement la structuration réelle des peuplements, et la délimitation des « populations » s'avérer particulièrement problématique (Figure M2 ; Harrison 1991; With 1997; Baguette 2004; Bender & Fahrig 2005; Cushman *et al.* 2006; Waples & Gaggiotti 2006; Mayer *et al.* 2009).

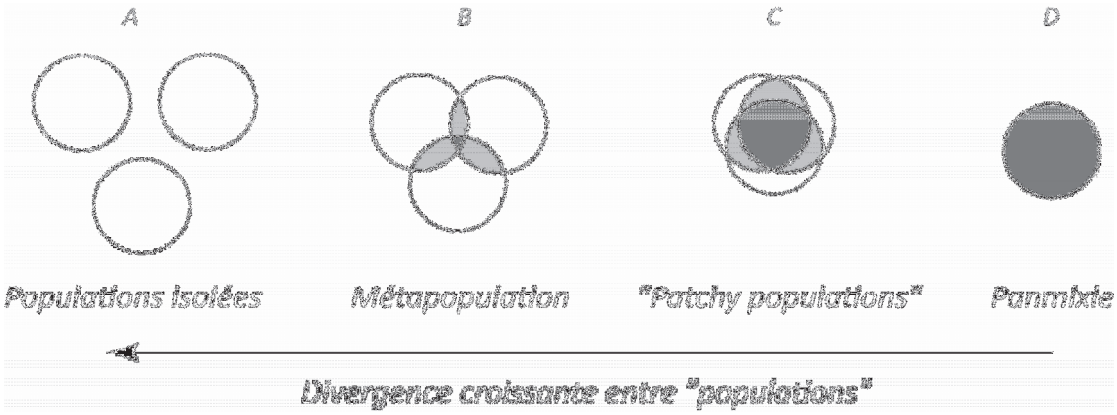


Figure M2 (*d'après Waples & Gaggiotti 2006*): Continuum de différentiation entre populations. Chaque cercle correspond à un agrégat d'individus. A: En l'absence totale de dispersion, les populations sont parfaitement identifiables car totalement isolées ; B: En situation de faible connectivité, les populations restent identifiables, la dispersion étant juste suffisante pour assurer la persistance des peuplements à l'échelle de la métapopulation ; C: En situation de connectivité élevée, les agrégats d'individus peuvent difficilement être considérés comme des populations à part entière ; D: En situation de connectivité totale, la délimitation des populations devient totalement arbitraire.

Par ailleurs, l'échantillonnage populationnel est très généralement associé à l'utilisation de mesures de différentiation génétique entre populations, reposant sur des hypothèses théoriques fortes (telles qu'un équilibre dynamique entre dérive génétique et flux de gènes lors de l'utilisation des Fst), rarement vérifiées dans la nature, et souvent ignorées (Waples 1998; Marko & Hart 2011). Enfin, cette approche implique une réduction de la couverture d'échantillonnage du paysage, c'est-à-dire une moindre représentativité de l'hétérogénéité du paysage dans les données, car, lors de la mise en œuvre de méthodes statistiques développées en génétique des populations, une grande part des prélèvements d'ADN doit nécessairement être réservée à l'exhaustivité d'échantillonnage des agrégats (Paetkau *et al.* 2004; Kalinowski 2005).

Changer d'unité d'échantillonnage en travaillant non plus à l'échelle de l'agrégat mais à l'échelle de l'individu permettrait de contourner ces problèmes : s'affranchissant de toute délimitation *a priori* des populations ainsi que des contraintes liées à l'utilisation de modèles théoriques de structuration génétique parfois trop peu réalistes, une approche individu-centrée autoriserait également une meilleure prise en compte de la variabilité paysagère, la diminution du nombre de prélèvements par agrégat entraînant une résolution d'échantillonnage plus fine du paysage (Figure M3).

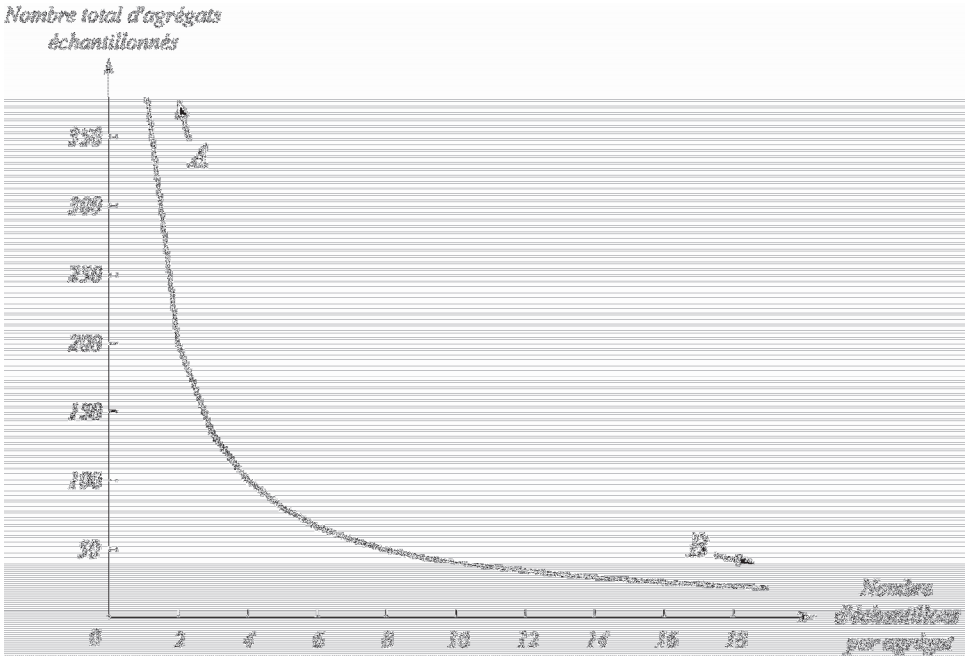


Figure M3 : Compromis entre couverture d'échantillonnage du paysage (A : accent mis sur la variabilité paysagère) et exhaustivité d'échantillonnage des agrégats (B : accent mis sur la variabilité génétique dans le cadre d'une approche populationnelle). L'exemple présente le cas théorique de la répartition spatiale de 400 échantillons : diminuer le nombre d'échantillons par agrégat autorise une couverture plus fine de la zone d'étude.

Si l'utilisation d'un échantillonnage génétique individu-centré en génétique du paysage a été suggéré de longue date (Manel *et al.* 2003; Anderson *et al.* 2010), il semblerait qu'une telle approche n'ait jamais été véritablement testée, ni mise en application dans le cadre de l'étude de peuplements structurés. Le cas des amphibiens est à ce titre tout à fait symptomatique : réputés ne disposer que de faibles capacités de dispersion, associées à un comportement souvent fortement philopatrique (Beebee 2005;

Cushman 2006), et présentant une distribution discontinue au niveau des sites aquatiques en période de reproduction, les amphibiens ont été très généralement considérés comme présentant une dynamique métapopulationnelle. Or malgré le fait que cette vision de la dynamique des peuplements d'amphibiens ait été plusieurs fois remise en question (Marsh & Trenham 2001; Jehle *et al.* 2005; Smith & Green 2005), l'écrasante majorité des études visant à étudier les flux de gènes chez ces animaux reposent encore sur un échantillonnage populationnel et l'utilisation de mesure de différentiation de type Fst (Spear *et al.* 2005; Wang 2009; Goldberg & Waits 2010; Murphy *et al.* 2010a; Emaresi *et al.* 2011 ; mais voir également Austin *et al.* 2011).

Lorsque l'on cherche à déterminer l'influence du paysage sur les flux de gènes chez un organisme présentant une distribution discontinue, ne serait-il pas plus avantageux de pratiquer un échantillonnage individu-centré ? Cette approche individu-centrée est-elle aussi efficace qu'une approche populationnelle classique ? N'entraîne-t-elle pas d'autres biais, qu'il serait dès lors nécessaire d'identifier et de prendre en compte ? Basés sur l'analyse de jeux de données simulées et empiriques, les différents chapitres de ce manuscrit visent à apporter à ces interrogations quelques éléments de réponses, repris et mis en perspective en fin de document.

Echantillonnage individu-centré : simulations...

En génétique du paysage, en permettant d'exercer un contrôle des processus responsables de la distribution de la variabilité génétique à l'échelle du paysage, les simulations rendent possible une comparaison fiable et sans ambiguïté de plusieurs approches méthodologiques sous différentes conditions environnementales (Balkenhol *et al*. 2009b; Epperson *et al.* 2010). Je me suis donc tout d'abord intéressé à une comparaison entre échantillonnage individu-centré et échantillonnage populationnel sur la base de jeux de données simulés sous différents régimes de dispersion et dans deux cas simples d'isolement par la distance (IBD) et d'isolement par barrière aux flux de gènes (*Isolation-by-barrier*, IBB ; Millions & Swanson 2007; Landguth *et al.* 2010) **[*Chapitre 1*]**.

... et études de cas.

Le triton alpestre *Ichthyosaura alpestris* est un amphibien bien répandu dans le quart nord-est de la France, qui coexiste avec des urodèles plus rares et à fort enjeu de conservation tels que le triton crêté (*Triturus cristatus* ; Joly *et al.* 2001; Emaresi *et al.* 2011), classé en Annexes 2 et 4 de la Directive Européenne Habitats (CEE 92/43). Le triton alpestre est par ailleurs classé « espèce déterminante Trame Verte et Bleue » pour la région Bourgogne (Sordello *et al.* 2011), ce qui souligne l'importance d'une meilleure compréhension de l'influence de la matrice paysagère sur les capacités de dispersion de cette espèce dans le cadre de la mise en place du Schéma Régional de Cohérence Ecologique bourguignon (Allag-Dhuisme *et al.* 2010). Ma première tâche a été de mettre au point une banque de 14 nouveaux marqueurs microsatellites polymorphes, afin de compléter la gamme de marqueurs existants pour le triton alpestre (Garner *et al.* 2003) . De plus, j'ai testé une procédure de prélèvement d'ADN (*skin swabbing*) limitant au maximum la manipulation des amphibiens, inspirée des méthodes de détection du champignon pathogène *Batrachochytrium dendrobatidis* (Berger *et al.* 1998; Hyatt *et al.* 2007; Soto-Azat *et al.* 2009) **[*Chapitre 2*]**.

Je me suis ensuite intéressé à l'étude de la dispersion du triton alpestre dans le cadre d'un échantillonnage individu-centré dans deux situations d'isolement par barrière aux flux de gènes (IBB) et d'isolement par la résistance que les milieux peuvent opposer aux mouvements des animaux (*Isolation-by-resistance*, IBR ; Cushman *et al.* 2006). La région Bourgogne est traversée par deux grandes infrastructures linéaires de transport (ILT) : l'autoroute A6 et la ligne à grande vitesse LGV-PSE. Cette problématique présentant un fort enjeu en termes de connectivité (Balkenhol & Waits 2009; Holderegger & Di Giulio 2010), j'ai tout d'abord cherché à déterminer l'impact potentiel de ces deux infrastructures sur la structuration génétique du triton alpestre par le biais de différentes approches *ad hoc* et corrélatives **[Etude de cas IBB : *Chapitre 3*]**.

L'évolution des pratiques agricoles au cours des dernières décennies a fortement impacté les milieux et leur mosaïque (déforestation, drainage, intensification, remembrement, etc.). Les nouveaux paysages agricoles se caractérisent par une érosion générale de la biodiversité, en particulier chez les amphibiens (Janin *et al.* 2009; Goldberg

& Waits 2010; Cosentino *et al.* 2011). Même dans un territoire relativement épargné telle que la Puisaye, une altération de la connectivité fonctionnelle du paysage n'est pas exclue : cette région, située à l'ouest d'Auxerre, a en effet connu comme tant d'autres une politique de remembrement à l'origine d'une modification plus ou moins marquée des pratiques agricoles. Caractériser l'influence de l'occupation des sols sur les possibilités de dispersion du triton alpestre est donc une étape indispensable à une meilleure prise en compte de l'hétérogénéité du paysage dans l'élaboration des réseaux écologiques de Bourgogne **[Etude de cas IBR : *Chapitre 4*]**.

CHAPITRE 1

Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling schemes.

Avant-propos.

En génétique du paysage, lorsqu'une espèce-cible est distribuée de manière discontinue dans le paysage, chaque agrégat est traditionnellement considéré comme une population à part entière (chaque individu est susceptible de s'accoupler avec tous les autres), et échantillonnée comme telle, par le prélèvement d'échantillons génétiques sur plusieurs dizaines d'individus (Broquet & Petit 2009). Des mesures de différentiation génétique (ou de distances génétiques) entre paires d'agrégats sont ensuite utilisées pour identifier les éléments paysagers associés à une augmentation ou à une diminution de la différentiation génétique entre populations. Cette approche, encore très largement utilisée, présente pourtant plusieurs inconvénients :

- Les limites réelles de chaque « population » sont inconnues (elles dépendent du domaine vital potentiel de chaque individu et de l'intensité de la dispersion), et ne correspondent pas nécessairement au simple agrégat échantillonné ;
- Les métriques généralement utilisées pour quantifier la différentiation génétique entre agrégats reposent sur des estimations de fréquences alléliques à l'échelle populationnelle, ce qui constitue une perte potentielle de l'information portée par chaque individu (« agencement » des allèles formant le génotype) ; certaines sont par ailleurs basées sur des hypothèses théoriques peu réalistes (Marko & Hart 2011);
- Augmenter l'effort d'échantillonnage au sein de chaque agrégat implique une réduction drastique du nombre d'agrégats qu'il sera possible d'échantillonner, entraînant un balayage moins important de l'hétérogénéité du paysage ; or c'est bien cette hétérogénéité du paysage, plutôt que l'information génétique en elle-même, qui est le facteur focal de la génétique du paysage.

Un échantillonnage individuel (entre 1 à 10 individus par agrégat seulement), couplé à un traitement statistique des données interindividuelles plutôt qu'interpopulationnelles, permettrait de contourner ces différents écueils, tout en augmentant de manière très importante le nombre d'agrégats échantillonnés dans l'espace.

Par le biais de simulations et d'analyses de jeux de données empiriques, cette première partie vise à mettre en évidence les avantages et les limites d'une telle approche.

Cet article est en cours de révision pour publication dans *Molecular Ecology* :

Prunier J., Kaufmann B., Fenet S., Picard D., Pompanon F., Joly P., Lena J. P. Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling schemes.

Chapitre 1

Original Article

Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling schemes.

J. PRUNIER[1, 2], B. KAUFMANN [2], S. FENET [3], D. PICARD [4], F. POMPANON [5], P. JOLY [2] and J.P. LENA[2]

[1] *Ecosphère, 3bis rue des Remises, 94100 Saint Maur des Fossés, France*

[2] *Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.*

[3] *Université de Lyon, UMR 5205 Laboratoire d'Informatique en Image et Systèmes d'information, Bat. Nautibus, Université Lyon 1; CNRS; Villeurbanne, F-69622, France.*

[4] *Université d'Angers, UMR CNRS 6554 Laboratoire d'Etudes Environnementales des Systèmes Anthropisés, 2 Boulevard Lavoisier, 49045 Angers Cedex 01, France.*

[5] *Université Joseph Fourier, Laboratoire d'Ecologie Alpine, CNRS, UMR 5553, 2233 Rue de la Piscine, 38041 Grenoble Cedex 9, France*

Keywords: CDPOP, restricted permutations, functional connectivity, *Ichthyosaura alpestris*, Mantel tests, patchy populations

Corresponding author:

Jérôme Prunier

Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.

Fax : (+33) 4 72 43 11 41

Email: jerome.prunier@gmail.com

Running title: Individual sampling in landscape genetics.

Abstract

Landscape genetics focuses on assessing the resistance of landscape matrices by relating observed spatial genetic structures to landscape characteristics. When studying organisms with a clustered spatial distribution, landscape genetic studies are often based on population-level analyses, i.e. on estimations of genetic distances between pairwise populations. However, small numbers of population pairs only provide a poor scan of landscape heterogeneity. We used matrix correlation analyses on empirical and simulated genetic datasets to investigate the efficiency of an individual-based sampling scheme at evaluating the impacts of landscape on isolation-by-distance and isolation-by-barrier genetic patterns. The individual-based approach proved to be a flexible and efficient methodological alternative to the more conventional population-based approach. It notably allows the design of powerful targeted sampling schemes to assess the impact of a priori defined landscape features on gene flow. Each strategy offering particular advantages, a combined use of both sampling schemes is discussed.

Introduction

In the context of accelerating landscape fragmentation worldwide, understanding the movements of individuals and genes across landscapes is of paramount importance, because they provide insight into the conservation status of populations (Cushman 2006; Safner *et al.* 2011a). Reaching this understanding relies on landscape genetics to provide theoretical and applied tools (Manel *et al.* 2003; Segelbacher *et al.* 2010). Two kinds of approaches may be distinguished in landscape genetic studies. Overlay methods, relying on the use of genetic information to characterize population genetic structure or to quantify gene flow among populations (e.g. bayesian clustering methods or assignment methods ; Pritchard *et al.* 2000; Manel *et al.* 2005), are widely used in genetic studies but do not allow to directly test for hypotheses regarding the impact of landscape on gene flow. In this paper, we thus focused on correlative methods, which rely on the direct comparison of genetic and landscape pairwise data, generally through the use of Mantel tests (Manel *et al.* 2003; Cushman *et al.* 2006). Genetic data are used to infer indirect measures of gene flow through the calculation of pairwise genetic distances, based either on measures of genetic differentiation among populations or on measures of genetic relatedness among individuals. How to conduct correlative analyses is not straightforward: the spatial distribution of individuals relative to populations varies tremendously across species and landscapes, and the temporal and spatial scales at which genes travel encompass several orders of magnitude (Schwartz & McKelvey 2009; Anderson *et al.* 2010). It is therefore crucial to propose sampling schemes combined with metrics used to estimate the impact of landscape on gene flow that are best fitted to the ranges of structures and scales found in nature.

Spatial and temporal distribution of individuals relative to populations

When individuals are continuously distributed, correlative analyses usually follow an individual-based sampling scheme (*ISS*). However, individuals are often distributed in aggregates, in which case each aggregate is generally sampled and treated as a discrete population (population-based sampling scheme, *PSS;* Broquet & Petit 2009; Anderson *et*

al. 2010). This widespread approach stems from the predominance of the metapopulation paradigm (Hanski 1999) and former theoretical models such as Wright's island model (1931), which assume restricted gene flow among local populations. Except for a few species, the relevancy of this paradigm to terrestrial landscapes is questionable (Baguette 2004), since the degree to which populations are isolated is mainly a function of the movement abilities of individuals relative to habitat heterogeneity (Tischendorf & Fahrig 2000a). Actually, observed aggregates may only be temporary (Anderson 2010) and individuals may be encountered at low densities between aggregates (Cushman *et al.* 2006). As a consequence, considering each spatial aggregate as a discrete population may be irrelevant in most cases, either because the concept simply does not apply or because of the challenging delineation of real population boundaries (Manel *et al.* 2003). The spatial organization of aggregated individuals may thus be better described using a patchy population model, which predicts weak genetic differentiation among subpopulations (Waples & Gaggiotti 2006; Mayer *et al.* 2009).

Indirect estimators of gene flow and sampling schemes

The efficiency of a *PSS* in landscape genetic studies may suffer from two kinds of limitations. The first flaw proceeds from the use of indirect estimators of gene flow based on population allelic frequencies, such as *Fst* (Wright 1951; Weir & Cockerham 1984), Nei's genetic distance (Nei *et al.* 1983) or Cavalli-Sforza and Edwards' chord distance (1967; see Ruzzante 1998 for further examples): these metrics may suffer from the population-level averaging of the genetic information over individuals (Kelly *et al.* 2010). Furthermore, some of these metrics (e.g. *Fst* or Nei's genetic distance) rely on assumptions, such as the equilibrium between genetic drift and gene flow, which may not be encountered in real situations (Marko & Hart 2011). The second flaw proceeds from the sampling scheme itself: *PSSs* often imply sampling a sufficient number of individuals (ideally, at least 50; Ruzzante 1998; Kalinowski 2005; Broquet & Petit 2009), across an optimized but often limited number of subpopulations, of which some may be discarded from the sampling scheme because densities are too low, or because of financial or operational constraints. This limited number of aggregates, by restricting the sample size of pairwise genetic distances, may impair the inferential power of correlative analyses

when tracking the effects of landscape features on dispersal (Legendre & Fortin 2010). Furthermore, *PSSs* also require an efficient coverage of neighbouring populations to ensure a sufficient power in estimating movement rates (Epperson 2003a; Broquet & Petit 2009), notably when using assignment methods (Paetkau *et al.* 2004; Holderegger & Wagner 2008): the existence of unsampled (or ghost) populations may actually affect the estimates of key parameters such as genetic differentiation among populations of interest (Slatkin 2005; Waples & Gaggiotti 2006). To cope with feasibility constraints, this need for an exhaustive sampling of aggregates area usually implies a drastic restriction in the extent of the study area, thus impairing the representativeness of environmental spatial heterogeneity in datasets (Anderson *et al.* 2010).

In landscape genetic studies, it may thus be preferable to assess genetic distances using measures of relatedness between individuals, without defining a priori populations on the basis of observed aggregates (Manel *et al.* 2003). Along with the obvious benefit of avoiding the challenging delineation of putative population boundaries, the *ISS* offers three potential advantages: (1) avoiding the use of averaged genetic information such as allelic frequencies; (2) enhancing the power of statistical analyses by increasing the sample size of pairwise genetic distances (Legendre & Fortin 2010); (3) ensuring a better coverage of the landscape heterogeneity: by ensuring an exhaustive sampling of aggregates and by increasing the number of sampled points over a broader spatial extent, over several replicate landscapes or in the direct vicinity of a targeted landscape feature (Anderson *et al.* 2010), the *ISS* may actually allow a better representation of landscape variables in datasets.

Despite all the arguments listed above, the use of an *ISS* in species with an apparently clustered distribution, though suggested for years (Manel *et al.* 2003), is still particularly uncommon in correlative analyses (but see Austin *et al.* 2011). Our hypothesis is that, when studying spatially aggregated individuals, the use of measures of relatedness between individuals should be preferred to the use of metrics based on allelic frequencies: along with a better inferential power, it allows applying an *ISS* which may provide a better spatial spreading out of pairwise distances across a given landscape, and therefore a better coverage of landscape heterogeneity.

Relative performance of ISSs and PSSs to detect spatial genetic patterns

To test this hypothesis, we used a threefold approach. We first conducted spatially explicit, individual-based simulations that allowed us to investigate the relative performance of random *ISSs* and *PSSs* to detect spatial genetic patterns according to varying dispersal abilities. We focused on two kinds of processes known to drive spatial genetic patterns: isolation-by-distance (*IBD*) and isolation-by-barrier (*IBB*). In most landscape genetic studies, *IBD* is considered as the null hypothesis, that is as the standard process driving genetic differentiation among individuals or populations (e.g. Broquet *et al.* 2006b; Emaresi *et al.* 2011). It was thus of prime importance to test for the efficiency of the *ISS* to detect such a pattern. Competing hypotheses, such as *IBB* models or more complex isolation-by-landscape resistance models (*IBR*), are then proposed to determine whether adding landscape variables may improve the prediction power of the standard *IBD* model (e.g. Cushman *et al.* 2006; Goldberg & Waits 2010). We thus focused on the efficiency of the *ISS* to explain spatial genetic structures in an *IBB* model, a simple but commonly tested hypothesis (Cushman *et al.* 2006; Landguth *et al.* 2010; Safner *et al.* 2011b). In the case of *IBB*, we examined how the efficiency of *ISSs* and *PSSs* could be enhanced by using a targeted sampling scheme, where individuals are sampled preferentially in the direct vicinity of the landscape feature supposed to act as a barrier (Anderson *et al.* 2010).

Second, we focused on several empirical datasets to test the relative performance of measures of genetic differentiation among individuals or populations to detect an *IBD* pattern, in the context of either an *ISS* or a *PSS*.

Material and methods

Simulated and empirical datasets

Simulated datasets.

Simulations were performed to test the ability of various sampling procedures to detect a significant genetic pattern under different gene flow regimes. For this purpose, we used CDPOP (Landguth & Cushman 2010) to simulate through 100 non-overlapping generations the genetic differentiation among 160 populations randomly placed in a 19 x 22 km homogeneous area (Fig. 1 and Table 1). Each population was initiated with 30 individuals and kept at a constant size over generations. Genetic polymorphism was set to 10 microsatellite loci and 10 alleles per locus (mean H_0 = 0.90), and mutation rate to zero (Cushman & Landguth 2010b). Only males were allowed multiple mating, and the litter size of paired animals was drawn according to a Poisson distribution with the mean set to three. Offspring sex was randomly assigned following a binomial distribution and an unbiased sex ratio. Dispersal was allowed only during the juvenile stage and the dispersal distance of juveniles was drawn from a probability distribution inversely proportional to a linear function. Maximal dispersal distance (*Dmax*) was fixed at 10, 20 and 30 % of the maximal extent of the study area to mimic a quasi-metapopulation dataset with low inter-aggregate dispersal (*Low*), a patchy population dataset with medium dispersal (*Medium*) and a quasi-continuous population dataset with high dispersal (*High*), respectively (see Appendix S1 in Supporting Information).

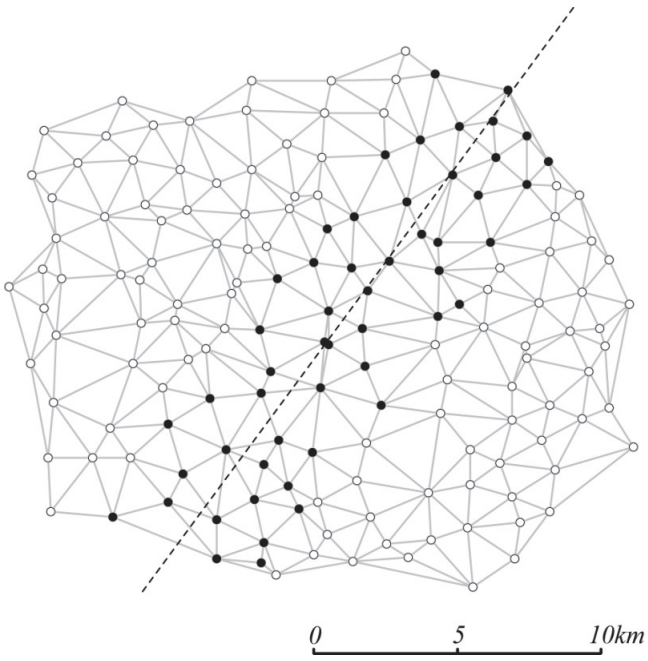


Figure 1 Random localization of the 160 aggregates of 30 individuals (referred to as populations), used to simulate genetic exchanges with CDPOP (Landguth & Cushman 2010). Grey lines stand for neighbourhood relations between populations as defined by a Delaunay triangulation. The barrier (dotted black line) segregates individuals in two sets of 80 populations. Black circles stand for aggregates located at less than 3000m from the barrier. The mean distance between neighbouring populations (as defined by the Delaunay triangulation) was 1685 ± 587m.

Table 1. Main characteristics of the datasets. SS: sampling scheme (P: population-based sampling scheme *PSS* / I: individual-based sampling scheme *ISS*). P: Number of sites with at least two sampled genotypes. N: total number of sampled genotypes. mN: mean number of sampled genotypes per site. Mrd: maximal recorded distance (in km). L: number of microsatellite loci. A: total number of alleles.

Dataset	SS	P	N	mN	Mrd (km)	L	A	Citations
Simulated *	P-I	160	4800	30.00	22.9	10	100	This study
Frog	P	10	200	20.00	9.3	7	136	Safner *et al.* 2011
Salamander	P	9	249	27.67	12.1	7	46	Unpublished data †
Toad	P	19	384	20.21	59.6	6	131	Unpublished data †
Snail	P	46	986	21.43	768.4	5	78	Charbonnel *et al.* 2002a, b
Newt	I	57	114	2.00	24.2	12	84	This study
	P	6	159	26.50	22.2			

* Three simulated datasets (*Low*, *Medium*, *High*) for *IBD* detection, and three simulated datasets (Low_{+barr}, $Medium_{+barr}$, $High_{+barr}$) for *IBB* detection.

† See Appendix S2 for details.

The evolution of genetic structuring over generations in these three simulated datasets was followed using Rousset's (1997) linearized *Fst* (*Fst*/(1-*Fst*)*;* hereafter denoted simply as *Fst*) as an indirect estimator of gene flow among all aggregates (p = 160) using all individuals in each aggregate (n = 30). Values of standard Mantel correlations between genetic and Euclidean distances (Fig. 2) decreased from *Low* to *High*, as expected by the *IBD* theory (Wright 1943): populations rarely exchanging migrants drifted apart genetically with time, while populations exchanging migrants more easily were genetically similar. Significant *IBD* could be detected in all datasets from the 2nd generation. We thus arbitrarily chose the 20th generation to evaluate the efficiency of the various sampling schemes in *IBD* detection.

To investigate the efficiency of the various sampling procedures to detect a recent barrier to dispersal, we simulated three new data sets (Low_{+barr}, $Medium_{+barr}$ and $High_{+barr}$) by placing an impermeable barrier to dispersal at the 20th generation, which segregated individuals into two sets of 80 populations. All the tests for *IBB* detection were realized at the 30th generation.

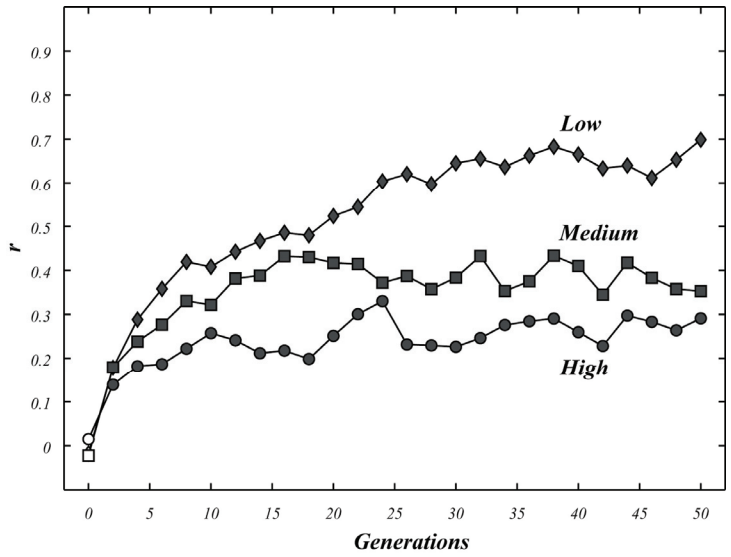


Figure 2 *IBD* detection through 50 generations for datasets *Low* (*Dmax* = 10% of the maximum study area extent, diamonds), *Medium* (*Dmax* = 20%, squares) and *High* (*Dmax* = 30%, circles), using a population-based approach over the whole dataset (see text for details). Symbols stand for the Mantel correlations r: a filled symbol indicates that the test was significant (α = 0.05).

Empirical datasets with a population-based sampling scheme (PSS).

To assess the efficiency of the use of measures of relatedness among individuals in realistic situations, we tested for *IBD* detection in four empirical datasets involving spatially clustered organisms sampled through a *PSS*, but differing in the extent of the study area, in sample size and in genetic polymorphism of microsatellite markers (Table 1). One dataset concerned a Malagasy freshwater snail *Biomphalaria pfeifferi* (*Snail*; Charbonnel *et al.* 2002a, b). Three datasets concerned pond-breeding amphibians sampled in France: the common frog *Rana temporaria* (*Frog*, Safner *et al.* 2011a), the fire salamander *Salamandra salamandra* (*Salamander*) and the common toad *Bufo bufo* (*Toad*). All datasets contained geographic coordinates and the genotype of each sampled individual at several microsatellite loci (see Table 1 and Appendix S2).

Empirical dataset with both an ISS and a PSS in Ichthyosaura alpestris.

To test whether an *ISS* may be more efficient than a *PSS* to detect a significant *IBD* pattern, we finally analysed data on the alpine newt *Ichthyosaura alpestris* (*Newt* in Table 1). The sampling was conducted in France during the 2010 breeding season, in an area of approximately 20 x 25 km (see Appendix S3). Six aggregates were sampled following a conventional *PSS* with at least n = 23 sampled individuals, while only one individual per sex was sampled in 51 aggregates (*ISS*, n = 2), resulting in a total of p = 57 sampled aggregates. The spatial extent of both the *PSS* and the *ISS* was quite similar (Table 1). Non-destructive genetic samples (Broquet *et al.* 2007) were taken from each captured individual, using non-sterile buccal swabs. DNA extraction, PCR amplifications and genotyping were performed using 12 microsatellite loci without null alleles, following Prunier *et al.* (2012).

Metrics used to compute pairwise genetic distance matrices

When comparing genetic distances between individuals (i.e. relatedness), we used the Bray-Curtis percentage dissimilarity measure (*Bc*; Legendre & Legendre 1998), as in Cushman et al. (2006). These authors showed that the Bray-Curtis percentage dissimilarity measure applied to individual genotypes was highly correlated to the a_r

metric (Rousset 2000) in their dataset. To test for the correlation of the a_r metric and the *Bc* coefficient in each of our simulated datasets, we randomly sampled 2 individuals in each population at generation 30 and calculated two genetic distance matrices, one based on the a_r coefficient (calculated using Genepop; Rousset 2008) and the other one based on the Bray-Curtis coefficient (using MATLAB). For all simulated datasets, matrices were significantly correlated (Mantel *r* coefficients ranging from 0.846 to 0.869, *P-values* < 0.001 with 9999 permutations). For programming convenience, we thus used the Bray-Curtis coefficient to calculate all pairwise genetic distances between individuals.

When computing genetic distances between populations, we used three measures of genetic differentiation based on allelic frequencies: (a) Rousset's (1997) linearized *Fst*, (b) Cavalli-Sforza and Edwards' Chord Distance (C_d, 1967) and (c) the Bray-Curtis dissimilarity coefficient applied to allelic frequencies (B_{cf}). *Fst* and C_d measures were chosen because they are commonly used in landscape genetic analyses, although estimates of gene flow using *F*-statistics may lead to erroneous conclusions, for instance if the assumption of equilibrium between genetic drift and gene flow is violated (Marko & Hart 2011). On the contrary, the C_d metric does not rely on any assumption and proved to be relatively insensitive to effective population size (Kalinowski 2002): It performed better than *Fst* in explaining the spatial distribution of genetic variation in a two-dimensional stepping-stone model in Dyer *et al.* (2010). The third genetic distance metric (B_{cf}) was chosen to test for the efficiency of the Bray-Curtis coefficient to detect genetic variations when based on the same information (allelic frequencies) as *Fst* and C_d metrics (see Appendix S4 for calculation).

Designing sampling schemes

For each simulated and empirical dataset (except *Newt*, see below), we designed random sampling procedures resulting in samples of size *np*, with *p* the number of randomly selected populations among the P available ones (Table 1) and *n* the number of randomly sub-sampled individuals per population, following either an *ISS* (*n* ranging from 1 to 10) or a *PSS* (*n* = 20, or less if necessary in empirical datasets). In *IBB* datasets, *p/2* aggregates were randomly sampled on either side of the barrier. For each *ISS*, a genetic distance matrix was created using the B_c metric between pairs of individuals. For each

PSS, four genetic distance matrices were created using the B_c, B_{cf}, C_d and *Fst* metrics between pairs of populations.

In addition to these random sampling procedures, we assessed the efficiency of a targeted *ISS* and a targeted *PSS* to detect the impact of a particular landscape feature on gene flow in *IBB* datasets (Low_{+barr}, $Medium_{+barr}$ and $High_{+barr}$). We selected the p = 52 populations located at less than 3000 m from the barrier (Fig. 1) and designed 6 sampling procedures by randomly sub-sampling n = 1 to 5 (*ISS*) or n = 20 (*PSS*) individuals per population, resulting in three sets of 6 genetic distance matrices (based on B_c for the *ISS* and *Fst* for the *PSS*).

To test if an *ISS* was more efficient than a *PSS* to detect a significant *IBD* pattern in *Newt*, we created genetic distance matrices either using *Fst* or C_d in the case of the *PSS* (p = 6, np = 159 genotypes) or using B_c in the case of the *ISS* (p = 57, np = 114 genotypes; see Table 1). In the case of the *ISS*, one individual per sex was randomly selected within each of the six sampled populations in order to compute *Bc*.

Matrix correlation analyses

We based our statistical approach on matrix correlation analyses using standard and partial Mantel tests (Smouse *et al.* 1986; Legendre & Fortin 2010). In *IBD* detection, each genetic distance matrix was compared to the corresponding pairwise Euclidean distance matrix using a standard Mantel test with 5000 permutations. In *IBB* detection, each genetic distance matrix was compared to the corresponding effective distance matrix, after controlling for the effect of the Euclidian distance matrix, using partial Mantel tests with 5000 permutations. In the effective distance matrix, distances were set to 0 when two sampled units were located on the same side of the barrier, and 1 when they were separated by the barrier (Epps *et al.* 2005). Partial Mantel tests were conducted by permuting residuals of genetic distances over Euclidean distances (null model *IBD*), as advised in Legendre & Fortin (2010). Each random or targeted sampling procedure was repeated 100 times, except in *Frog* and *Salamander* when n = 20 since the complete dataset was used, and in *Newt* because of a specific analysis (see below). A sampling procedure was considered efficient to detect a landscape effect when at least 95 % of the repetitions led to a significant Mantel test (P-value ≤ 0.05). The use of

measures of relatedness between individuals was compared to that of genetic differentiation between populations in terms of detection power, according to the minimum number *np* of genotypes required for a significant genetic structure to be detected. In dataset *Newt*, each pairwise genetic distance matrix was only used for *IBD* detection through a single standard Mantel test with 9999 permutations.

All continuous variables (genetic and Euclidean distances) were log-transformed following the *D=ln(d+1)* formula and standardized to meet linearity assumptions. When using measures of relatedness between individuals with $n \geq 2$, Mantel tests were conducted using restricted permutations (permuted objects were not the *np* individuals but the *p* blocks of *n* individuals) in order to take into consideration the non-independence of individuals belonging to the same population (Efron & Tibshirani 1993). All within-population distances (that is genetic distances between individuals from a same aggregate) were systematically removed from linearised semi-matrices, to calculate Mantel correlations without intra-site distance, as when using measures of differentiation between populations. All operations (sampling populations and individuals, calculating genetic and Euclidian distances and running Mantel tests with or without restricted permutations) were automated in the MATLAB software coding environment (Mathworks, Inc.). *Fst* measures were calculated as in Fstat (Goudet 2001), following Weir and Cockerham (1984), and then linearised following Rousset (1997). C_d measures were calculated as in Genetix 4.03 (Belkhir *et al.* 2004). We checked the validity of MATLAB scripts by comparing *Fst* and C_d measures in a few situations with Fstat and Genetix 4.03 outputs respectively, and Mantel correlation coefficients (*r*) and p-values with ZT (Bonnet & Van de Peer 2002) outputs.

Results

Relative performance of metrics and sampling schemes (ISS versus PSS) in simulated datasets

Relative performance of metrics in PSSs.

Whatever the spatial genetic pattern (*IBD* or *IBB*), metrics based on allelic frequencies (C_d, B_{cf} or *Fst*) led to similar detection levels (Fig. 3 and Table 2). As expected, *IBD* detection level decreased with increasing dispersal abilities; on the contrary, *IBB* detection level increased with increasing dispersal abilities (Landguth *et al.* 2010). No *IBB* pattern was detected in a low dispersal context (Low_{+barr}).

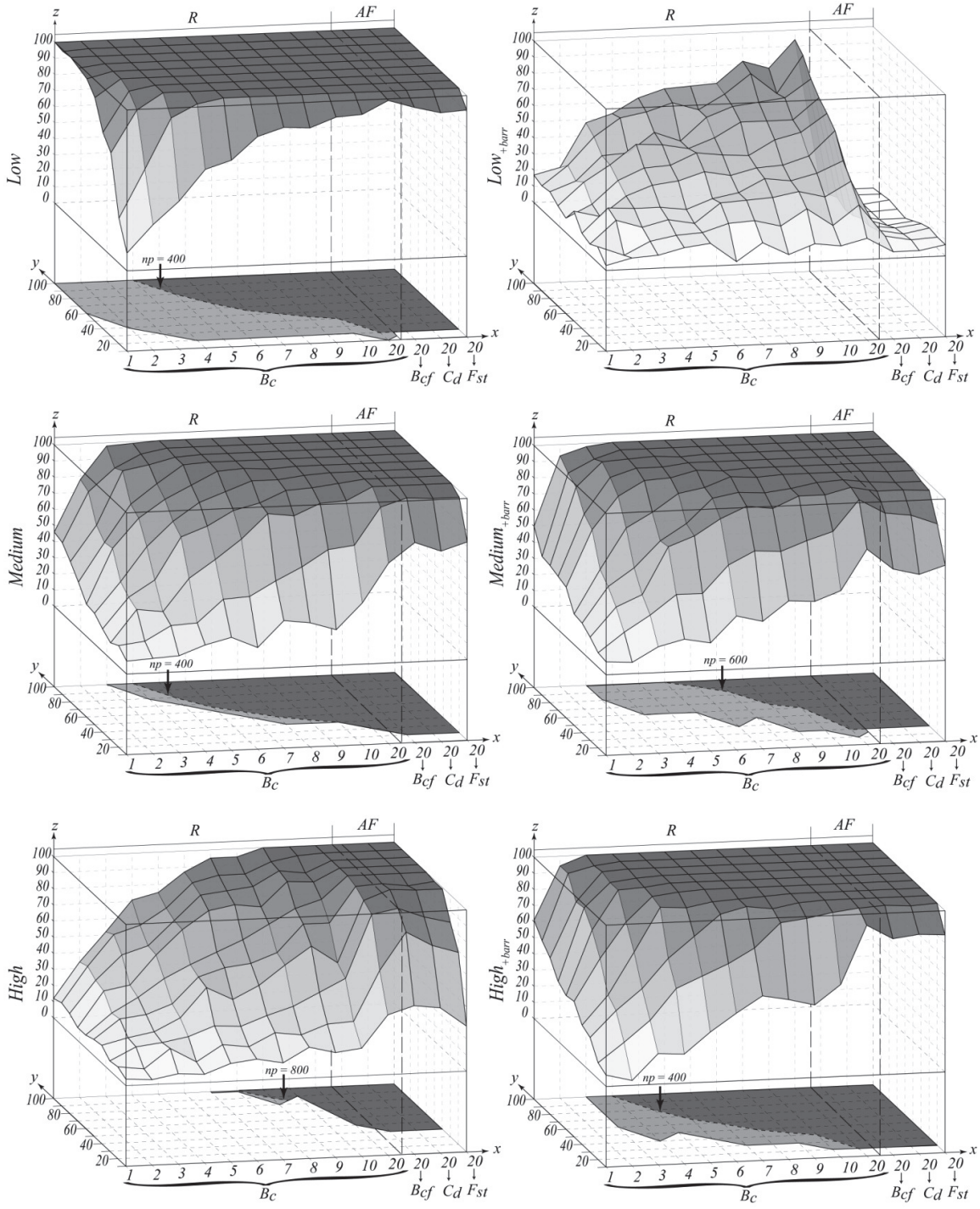
In *IBD* datasets, measures of relatedness between individuals (n = 20) performed roughly as well as metrics based on allelic frequencies in all situations. In *IBB* datasets, the use of measures of relatedness was at least as efficient as the use of allelic frequencies in all situations, and, though non-significant, led to high detection levels in Low_{+barr}, far beyond metrics based on allelic frequencies.

Table 2. Minimum number of sampled genotypes required to significantly detect a spatial genetic pattern (*IBD* or *IBB*) in each dataset, depending on the sampling scheme (Individual *ISS* or Population-based sampling scheme *PSS*) and the kind of measure used to calculate pairwise genetic distances (measures of relatedness between individuals *R* or measures of genetic differentiation between populations based on allelic frequencies *AF*).

Dataset	Pattern	*ISS - R*	*PSS - R*	*PSS - AF*
Low	*IBD*	60	200	400
Medium	*IBD*	300	600	400
High	*IBD*	700	1000	800
Low_{+barr}	*IBB*	/	/	/
$Medium_{+barr}$	*IBB*	240	400	600
$High_{+barr}$	*IBB*	160	400	400
Snail	*IBD*	20	200	200-600
Toad	*IBD*	140	280	280-320
Frog	*IBD*	/	200	/
Salamander	*IBD*	/	180	180

Figure 3 Genetic structure detection in *Low*, *Medium* and *High* (*IBD*) and Low_{+barr}, $Medium_{+barr}$ and $High_{+barr}$ (*IBB*), expressed in percentage of the total number of significant replicates (*z*-axis), according to the number *p* of selected populations (*y*-axis), the number *n* of sub-sampled individuals per population (*x*-axis) and the kind of measure used to calculate pairwise genetic distances (*R*: measures of relatedness between individuals using B_c; *AF*: measures of genetic differentiation between populations based on allelic frequencies using B_{cf}, D_c and *Fst*; see text for details). Gray surfaces at the bottom of graphs enable a better visualization of sampling procedures leading to a significant detection at 95%; in light gray, significant individual-based sampling schemes (*ISS*) requiring a more parsimonious sampling (in terms of the total number *np* of sampled genotypes) than the most parsimonious population-based sampling scheme (*PSS*).

→


Low
Low+barr
Medium
Medium+barr
High
High+barr
R
AF
np = 400
np = 600
np = 800
Bc
Bcf Cd Fst

Relative performance of random sampling schemes.

In all datasets, when using an *ISS* ($n \leq 10$), the use of measures of relatedness provided detection levels (in terms of the number *np* of required genotypes) at least equal to observed detection levels with metrics based on allelic frequencies (Fig. 3 and Table 2). *ISSs* were particularly efficient for *IBB* detection in all situations, and for *IBD* detection in *Low* and *Medium* datasets. Though non-significant in Low_{+barr}, the use of measures of relatedness led to higher detection levels than any metric based on allelic frequencies. As expected, the higher the dispersal abilities in *IBD* datasets and the lower the dispersal abilities in *IBB* datasets, the higher the number *np* of genotypes required for a significant genetic pattern to be detected. Whatever the dataset, there was a trade-off between the sampling coverage of aggregates and the sampling coverage of individuals in each aggregate : detection levels increased either with the increase of *n* (number of sampled genotypes per aggregate) to the expense of *p* (number of selected aggregates) or conversely, with the increase of *p* to the expense of *n*. Except when the spatial genetic pattern was strong (e.g. *Low* dataset), a minimum number of genotypes had to be sampled in each aggregate ($n \geq 3$).

Relative performance of targeted sampling schemes.

When using a targeted sampling procedure (see Fig. 1), the *IBB* detection level was significant as soon as $n \geq 3$, whatever the dataset (Fig. 4). When considering the number *np* of sampled genotypes, the targeted *ISS* was far more efficient than the targeted *PSS*, although both designs outperformed the random sampling schemes (Fig. 3).

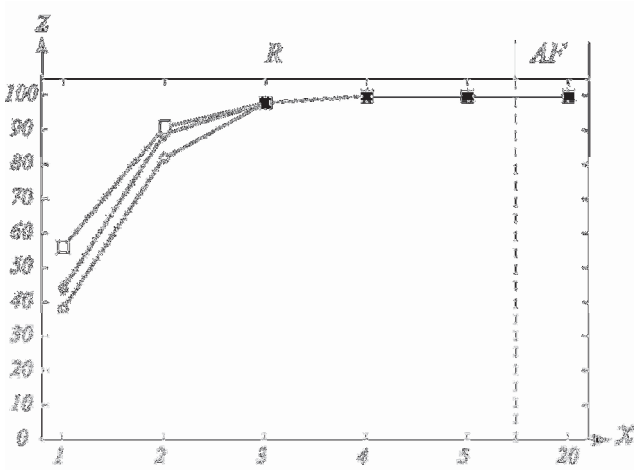


Figure 4 *IBB* detection following a targeted sampling scheme (52 aggregates localized at least than 3000m from the barrier), expressed in percentage of the total number of significant replicates (*z*-axis), according to the number *n* of subsampled individuals per population (*x*-axis). Pairwise genetic distances were based on measures of relatedness between individuals (R) for $n \leq 5$ (using B_c) or on allelic frequencies (AF) for $n = 20$ (using *Fst*). Diamonds: Low_{+barr} dataset; Squares: $Medium_{+barr}$ dataset; Circles: $High_{+barr}$ dataset. A filled symbol indicates that the test was significant in more than 95% of the replicates.

Relative performance of sampling schemes in empirical datasets sampled with a PSS

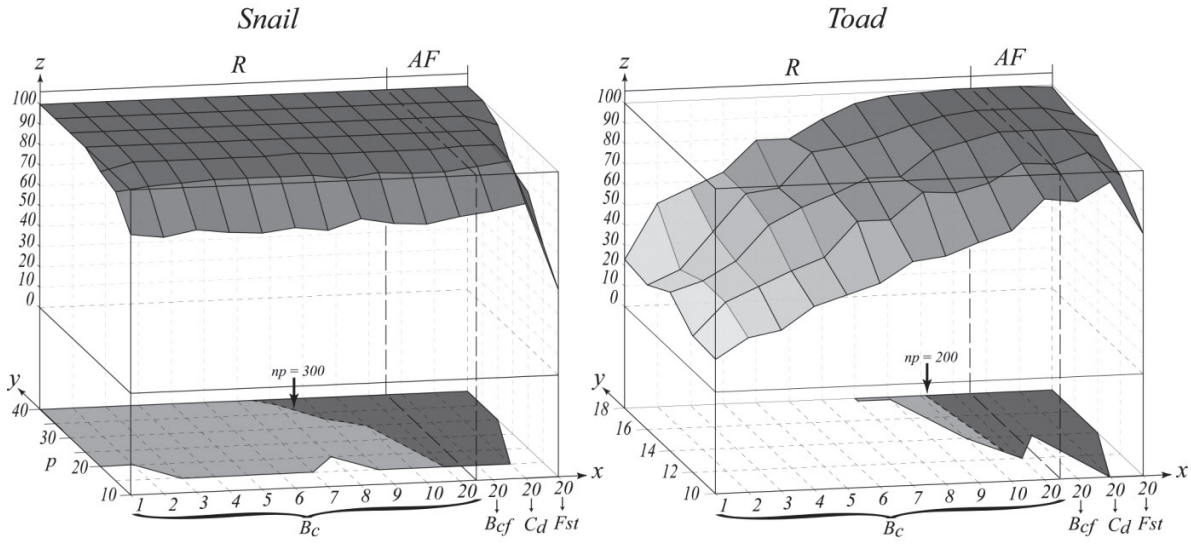
In *Snail*, the use of measures of relatedness between individuals largely outperformed metrics based on allelic frequencies, with 20 as a minimum required number of sampled genotypes ($n = 1$ and $p = 20$) to detect a significant *IBD* pattern. *Fst* was systematically less efficient than any other metric (Fig. 5 and Table 2). In *Toad*, the use of measures of relatedness performed better than metrics based on allelic frequencies for $n \geq 8$, with at least 144 required genotypes for a significant *IBD* pattern to be detected ($n = 8$ and $p = 18$), versus 280 for the most efficient method using allelic frequencies ($n = 20$ and $p = 14$). When using allelic frequencies, C_d performed better than B_{cf}, and *Fst*, and slightly better than B_c. In *Frog*, no *IBD* pattern could be detected, as expected according to Safner *et al.* (2011a), except when using measures of relatedness

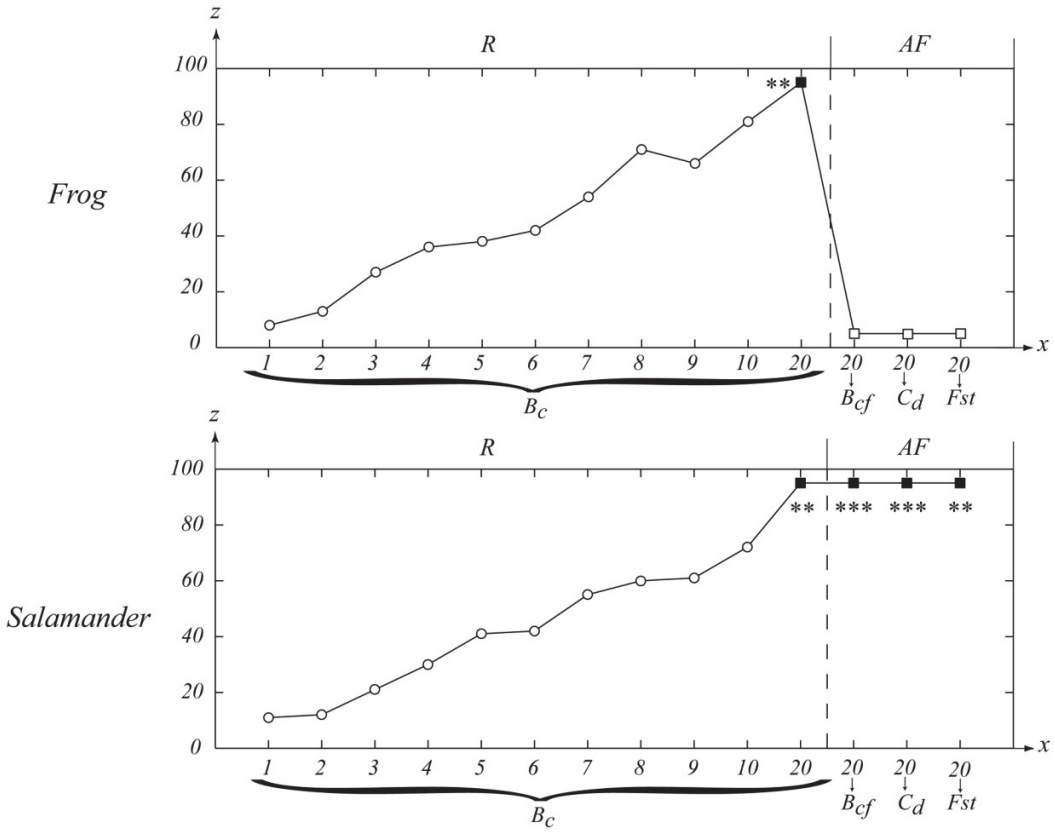
in a *PSS* (n = 20). In *Salamander*, no *IBD* pattern could be detected when using an *ISS* ($n \leq$ 10). For n = 20, measures of relatedness performed as well as metrics based on allelic frequencies.

Relative performance of ISS and PSS to detect a significant IBD pattern in the empirical dataset Newt

When using a *PSS* (six sampled aggregates), no *IBD* pattern (r = 0.2502, p = 0.196) was detected with *Fst* (as in Emaresi *et al.* 2011), while the use of C_d led to the detection of a weak *IBD* (r = 0.5485, p = 0.048). On the contrary, the use of an *ISS* (57 sampled aggregates) led to the detection of a highly significant *IBD* pattern (r = 0.0729, $p = 10^{-4}$). Decreasing the sampling coverage of individuals within populations in favour of a better sampling coverage of aggregates across space decreased the *IBD* detection threshold, while using fewer individuals (159 genotypes in the *PSS* as against 114 in the *ISS*).

Figure 5 *IBD* detection in empirical datasets *Snail*, *Toad*, *Frog* and *Salamander*, expressed in percentage of the total number of replicates (z-axis), according to the number p of selected populations (y-axis, for *Toad* and *Snail* only), the number n of sub-sampled individuals per population (x-axis) and the kind of measure used to calculate pairwise genetic distances. Datasets *Toad* and *Snail*: see legend in figure 2 for more details. Datasets *Frog* and *Salamander*: black circles indicate a significant test in more than 95% of the replicates; black squares indicate, in population-based sampling procedures, a significant test over the whole dataset (∗∗: P < 0.01; ∗∗∗: P < 0.001). Squares were arbitrarily placed at z = 95 for significant tests, at z = 5 otherwise.


Snail
Toad
z
R
AF
y
x
p
np = 300
np = 200
Bc
Bcf
Cd
Fst


Frog
Salamander
z
R
AF
x
**
***
Bc
Bcf
Cd
Fst

Discussion

Relative performance of metrics in PSSs

Metrics based on population allelic frequencies all provided similar *IBD* detection in all simulated datasets. In the empirical datasets *Snail* and *Toad*, the *Fst* metric was less efficient to detect *IBD* than other measures, notably the C_d metric. Strongly relying on equilibrium assumptions, interpreting *Fst* variations in terms of gene flow might actually be misleading (Marko & Hart 2011), because observed population differentiation may be caused by many forces, including gene flow between populations, but also behaviour, demography or mating system (Lowe & Allendorf 2010): *Fst* may thus suffer from high selfing rates, encountered for instance in the *Snail* dataset (Charbonnel *et al.* 2002a, b). On the contrary, the C_d metric, not contingent on any assumption, led to higher *IBD* detection levels in these two datasets. In the context of a population-based sampling scheme *PSS*, the use of the Bray-Curtis metric applied to allelic frequencies (B_{cf}), which proved to perform as well as C_d in all datasets except *Toad*, may deserve further investigation.

Measures of relatedness between individuals used in a conventional population-based sampling scheme *PSS* (number of sampled genotypes per population n = 20) performed as well as the best metric based on allelic frequencies in all *IBD* datasets, whatever the number of populations, the number of loci or the spatial extent of the study area (Tables 1 and 2). In the dataset *Frog*, it was the unique approach for which significant *IBD* was detected. Metrics based on allelic frequencies may suffer from a loss of resolution due to the averaging of genetic information over individuals (Kelly *et al.* 2010). Conversely, metrics of relatedness are founded on the averaging of the genetic information over alleles rather than over individuals, thus overcoming this loss of resolution. They also allow the number of pairwise samples to be increased, thus potentially increasing the inferential power of statistical analyses in detecting significant spatial genetic patterns (Legendre & Fortin 2010).

Relative performance of sampling schemes

The use of measures of relatedness in an individual-based sampling scheme *ISS* ($n \leq 10$) outperformed metrics used in *PSSs*, except when high inter-patch movements led to a weak *IBD* pattern (dataset *High*). Actually, increasing dispersal ability while keeping the extent of the study area unchanged may alter *IBD* detection because the study area becomes too small relative to the scale of gene flow (Anderson *et al.* 2010; Cushman & Landguth 2010a). However, when dispersal rates are too high, methodological and financial constraints might prevent researchers using allelic frequencies to reach the number of genotypes required to detect a significant genetic structure. For the same number of sampled genotypes *np*, increasing *p* to the expense of *n* in an *ISS*, that is decreasing the sampling coverage of individuals within populations in favour of a better sampling coverage of aggregates through space (Jaquiery *et al.* 2011), was at least as efficient as a conventional *PSS*. It was even more efficient in most situations since the same conclusions could be drawn with fewer genotypic data, thus allowing the sampling of aggregates with low densities of individuals. Besides, by decreasing the number of non-sampled aggregates, *ISSs* allow a more regular sampling and hence a better detection of genetic structures (Anderson *et al.* 2010). For instance, the use of an *ISS* in *I. alpestris* (*Newt*) provided a lower *IBD* detection threshold than the use of a *PSS*, although the analysis was performed with fewer genotypes. Finally, by drastically increasing the number *p* of aggregates that may be sampled, the *ISS* also offers a higher flexibility in spatial sampling designs, allowing the analysis of spatial genetic patterns at a wide range of spatial scales. Increasing the number of different distance classes between samples, the spatial extent of the study area and the number of independent replicates across scales and landscapes is of prime importance to make accurate inferences about gene flow (Anderson *et al.* 2010; Manel *et al.* 2010). Nevertheless, a minimum number of individuals must be sampled in each aggregate, especially when genetic patterns are weak: whatever the number *p* of selected aggregates, sampling a unique individual per aggregate may not be efficient enough to detect a significant genetic pattern. However, rather than drastically increasing the number of sampled genotypes per aggregate, which is known to increase the precision of Mantel tests (Landguth *et al.* 2012) but which may alter the spatial sampling scheme flexibility provided by the *ISS*, it may be more relevant

to increase the polymorphism of genetic data by using a higher number of highly variable loci (Landguth *et al.* 2012), a realistic perspective with regard to the rapid development of next-generation sequencing technologies (Segelbacher *et al.* 2010).

The same conclusions may be drawn about the efficiency of measures of relatedness (either in a *PSS* or in an *ISS*) in *IBB* detection (Fig. 2). As expected, the *IBB* signal increased with the increase of inter-patch movements (Landguth *et al.* 2010): metrics based on allelic frequencies, all providing similar *IBB* detection levels in all simulated datasets, totally failed to detect any recent barrier to gene flow when dispersal was low (Low_{+barr}). On the contrary, the use of measures of relatedness outperformed the use of allelic frequencies in all situations, independent of the sampling scheme. *IBB* detection was even better when using a targeted sampling procedure (Fig. 3), that is when sampled units were selected to avoid large gaps relative to a putative barrier to gene flow, as advised in Anderson *et al.* (2010). Provided that a minimum of three genotypes were sampled per population, a significant *IBB* was detected in all situations, including the scenario with low dispersal abilities (Low_{+barr}). Although more efficient than a random *PSS*, the targeted *PSS* was largely outperformed by the targeted *ISS* as the same conclusions were obtained with only three sampled individuals per aggregate. Here again, the *ISS* offers much flexibility in the selection of sampling units, and allows to efficiently assess the effect of localized landscape features on spatial genetic patterns.

Conclusion

The analysis of all the simulated and empirical datasets in this study provides evidence that the use of measures of relatedness between individuals, either based on an *ISS* (provided that a sufficient number *np* of genotypes is sampled) or on a *PSS,* may provide methodological and inferential advantages in landscape genetic studies when compared with the more conventional use of allelic frequencies. Moreover, because of its flexibility, the *ISS* should be considered as a promising alternative, especially for species whose biology and life history are insufficiently known. However, when studying clustered populations in a landscape genetics context, deciding whether the sampling scheme should be performed at the individual-level or at population-level depends on the

scientific objectives and the analysis methods. Some powerful genetic tools allowing the direct estimation on gene flow (such as assignment tests or parentage analyses; Manel *et al.* 2005; Broquet & Petit 2009) do require a *PSS*. Besides, estimations of within-population genetic variability may be a useful parameter in conservation genetics, leading to the identification of critical demographic processes such as bottlenecks, inbreeding or dispersal (Broquet & Petit 2009). The *PSS* thus totally deserves its place in a landscape genetics toolbox. Nevertheless, the use of measures of relatedness between individuals rather than allelic frequencies in *PSSs* should be considered an interesting new way of analysing spatial genetic patterns, provided that Mantel tests are performed with restricted permutations.

On the other hand, the problem of population boundary delimitation still plagues *PSSs* (Manel *et al.* 2003). Besides, a rising number of genetic tools are based on genotypes rather than allelic frequencies (Manel *et al.* 2007), and many others may benefit from an increased number of spatially independent genotypes (Manni *et al.* 2004; Crida & Manel 2007; Jombart *et al.* 2008). Estimations of within-population genetic variability may not always be essential for the analysis of dispersal processes between populations and the emphasis should be rather put on the optimal random or targeted coverage of aggregates through space. Ideally, field sampling in landscape genetic studies may be designed to benefit from the advantages of both strategies (Broquet & Petit 2009). First, genetic samples should be collected according to an *ISS* in a maximum of spatially independent sites in order to ensure the complete coverage of a large landscape area without a priori knowledge about the scale at which genetic structuring occurs. Sample sites may be either randomly distributed or specifically localized to optimize the detection of landscape features effect on genetic patterns. Second, genetic samples should be collected according to a *PSS* in a few relevantly distributed sites in order to enable the use of assignment tests or coalescent methods, which provide insightful complementary information about landscape permeability to gene flow in clustered populations (Holderegger & Wagner 2008; Anderson *et al.* 2010). From an operational point of view, when a landscape genetics study follows an *ISS*, the number *n* of individuals to be sampled per aggregate should be optimized according to the total number *p* of relevant aggregates identified across the study area (maximization of *p* to ensure a random or a

targeted exhaustive scan of the landscape heterogeneity) and the total number *np* of genotypes that might be reasonably analysed.

Analysing the efficiency of the use of measures of relatedness between individuals to detect landscape resistance to gene flow (isolation-by-landscape resistance *IBR;* Cushman *et al.* 2006) was beyond the scope of this study and will deserve further investigation. *IBB* may nevertheless be considered as a specific case of *IBR*, with a unique landscape feature (a barrier) surrounded by a homogeneous matrix. *IBR* may thus also benefit from the better scan of the landscape heterogeneity provided by an *ISS*, since the higher the number of pairwise genetic distances, the higher the number of particular combinations of landscape features that may be analysed, especially when designing a targeted sampling scheme.

References

Anderson CD (2010) Effects of Movement and Mating Patterns on Gene Flow among Overwintering Hibernacula of the Timber Rattlesnake (*Crotalus horridus*). *Copeia*, 54-61.

Anderson CD, Epperson BK, Fortin MJ*, et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology* **19**, 3565-3575.

Austin JD, Gorman TA, Bishop D (2011) Assessing fine-scale genetic structure and relatedness in the micro-endemic Florida bog frog. *Conservation Genetics* **12**, 833-838.

Baguette M (2004) The classical metapopulation theory and the real, natural world: a critical appraisal. *Basic and Applied Ecology* **5**, 213-224.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.03, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Bifolchi A, Picard D, Lemaire C, Cormier JP, Pagano A (2010) Evidence of admixture between differentiated genetic pools at a regional scale in an invasive carnivore. *Conservation Genetics* **11**, 1-9.

Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests, Ghent University.

Brede EG, Rowe G, Trojanowski J, Beebee TJC (2001) Polymerase chain reaction primers for microsatellite loci in the Common Toad *Bufo bufo*. *Molecular Ecology Notes* **1**, 308-310.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics* **8**, 509-511.

Broquet T, Petit EJ (2009) Molecular Estimation of Dispersal for Ecology and Population Genetics. In: *Annual Review of Ecology Evolution and Systematics*, pp. 193-216. Annual Reviews, Palo Alto.

Broquet T, Ray N, Petit E, Fryxell JM, Burel F (2006) Genetic isolation by distance and landscape connectivity in the American marten (Martes americana). *Landscape Ecology* **21**, 877-889.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics* **19**, 233-257.

Charbonnel N, Angers B, Rasatavonjizay R*, et al.* (2002a) The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. *Molecular Ecology* **11**, 2213-2228.

Charbonnel N, Angers B, Rasatavonjizay R, Bremond P, Jarne P (2002b) Evolutionary aspects of the metapopulation dynamics of *Biomphataria pfeifferi*, the intermediate host of *Schistosoma mansoni*. *Journal of Evolutionary Biology* **15**, 248-261.

Crida A, Manel S (2007) WOMBSOFT: an R package that implements the Wombling method to identify genetic boundary. *Molecular Ecology Notes* **7**, 588-591.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation* **128**, 231-240.

Cushman SA, Landguth EL (2010a) Scale dependent inference in landscape genetics. *Landscape Ecology* **25**, 967-979.

Cushman SA, Landguth EL (2010b) Spurious correlations and inference in landscape genetics. *Molecular Ecology* **19**, 3592-3602.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* **168**, 486-499.

Dyer RJ, Nason JD, Garrick RC (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology* **19**, 3746-3759.

Efron B, Tibshirani R (1993) *An Introduction to the Bootstrap* Chapman & Hall, New York, London.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (Mesotriton alpestris) inferred from a strip-based approach. *Conservation Genetics* **12**, 41-50.

Epperson BK (2003) *Geographical Genetics* Princeton University Press, Princeton and Oxford.

Epps CW, Palsboll PJ, Wehausen JD*, et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters* **8**, 1029-1038.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology* **19**, 3650-3663.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). University of Lausanne, Dorigny.

Hanski I (1999) *Metapopulation Ecology* Oxford University Press, Oxford.

Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* **58**, 199-207.

Jaquiery J, Broquet T, Hirzel AH, Yearsley J, Perrin N (2011) Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology* **20**, 692-705.

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

Kalinowski ST (2002) Evolutionary and statistical properties of three genetic distances. *Molecular Ecology* **11**, 1263-1273.

Kalinowski ST (2005) Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* **94**, 33-36.

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity* **101**, 423-436.

Landguth EL, Cushman SA (2010) CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources* **10**, 156-161.

Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**, 4179-4191.

Landguth EL, Fedy BC, Oyler-McCance SJ*, et al.* (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources* **12**, 276-284.

Legendre L, Legendre P (1998) *Numerical Ecology* 2nd English ed. Elsevier, Amsterdam.

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* **10**, 831-844.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038-3051.

Manel S, Berthoud F, Bellemain E*, et al.* (2007) A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology* **16**, 2031-2043.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate techniques. *Trends in Ecology & Evolution* **20**, 136-142.

Manel S, Joost S, Epperson BK*, et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology* **19**, 3760-3772.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**, 173-190.

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

Mayer C, Schiegg K, Pasinelli G (2009) Patchy population structure in a short-distance migrant: evidence from genetic and demographic data. *Molecular Ecology* **18**, 2353-2364.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65.

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

Prunier J, Kaufmann B, Grolet O*, et al.* (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (Ichthyosaura alpestris). *Molecular Ecology Resources* **12**, 524-531.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1-14.

Safner T, Miaud C, Gaggiotti O*, et al.* (2011a) Combining demography and genetic analysis to assess the population structure of an amphibian in a human-dominated landscape. *Conservation Genetics* **12**, 161-173.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011b) Comparison of Bayesian Clustering and Edge Detection Methods for Inferring Boundaries in Landscape Genetics. *International Journal of Molecular Sciences* **12**, 865-889.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.

Segelbacher G, Cushman SA, Epperson BK*, et al.* (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics* **11**, 375-385.

Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology* **14**, 67-73.

Smouse PE, Long JC, Sokal RR (1986) Multiple-regression and correlation extensions of the mantel test of matrix correspondence *Systematic Zoology* **35**, 627-632.

Steinfartz S, Kusters D, Tautz D (2004) Isolation and characterization of polymorphic tetranucleotide microsatellite loci in the Fire salamander *Salamandra salamandra* (Amphibia: Caudata). *Molecular Ecology Notes* **4**, 626-628.

Tischendorf L, Fahrig L (2000) On the usage and measurement of landscape connectivity. *Oikos* **90**, 7-19.

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.

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

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 0097-0159.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.

Acknowledgements

This work was supported by the Agence Nationale de la Recherche et de la Technologie (ANRT, agreement no 2 ⁄ 2009), by two infrastructure managers (Réseau Ferré de France RFF and Autoroutes Paris-Rhin-Rhône APRR) and by a consulting company in ecological engineering (ECOSPHERE). The genetic sampling of alpine newts was conducted in accordance with French laws and with the approval of both the Préfecture de Saône-et-Loire and the Préfecture du Jura. We warmly thank E. Calonnier, A. Plaisance, G. Perez , J. Mangin and O. Grolet for their precious help and support on the field and laboratory. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon/Villeurbanne - France), for providing a significant amount of

the computing resources needed for this work. We also thank E. Landguth for clues as to the use of CDPOP, S. Manel and P. Giraudoux for insightful comments about results, and N. Charbonnel, P. Jarne, T. Safner, and C. Miaud for authorizing the use of their respective published datasets. We warmly thank Dr S. Spear and three anonymous reviewers for insightful suggestions.

Supporting information

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

Appendix S1: Details about dispersal scenarios in simulated datasets.

Appendix S2: Brief description of unpublished empirical datasets *Toad* and *Salamander*.

Appendix S3: Localization of the 57 ponds in which at least two alpine newts were sampled (*Newt* dataset).

Appendix S4: Use of the Bray-Curtis coefficient applied to individual genotypes or to allelic frequencies.

Contribution by authors: BK, JP and JPL designed the study; DP, JP and JPL collected empirical data; JP and SF performed modeling work; BK, FP, JP, JPL and PJ analysed output data; JP and JPL wrote the first draft of the manuscript and all authors contributed substantially to revisions; FP, JPL and PJ ensured the work supervision.

SUPPORTING INFORMATION

Appendix S1: Details about dispersal scenarios in simulated datasets.

With a maximal dispersal distance (*Dmax*) fixed at 10, 20 and 30 % of the maximal extent of the study area, mean dispersal distances were respectively 848±18m, 2071±35m and 3108±77m. Mean dispersal distance for *Dmax* = 10% was inferior to the mean distance between neighbouring populations (1685m, see Appendix S1), leading to high genetic drift. Mean dispersal distance for *Dmax* = 30% was twice as important as the mean distance between neighbouring populations, leading to high gene flow across the study area. Actually, no significant *IBD* pattern could be detected for *Dmax* ≤ 5%, due to an excessive genetic drift, or for ≥ 35%, due to an excessive homogenisation of genotypes among populations (data not shown). As a result, datasets based on maximal dispersal distances of 10%, 20% and 30% represented a realistic gradient in the degree of genetic structuring, from a quasi-metapopulation scenario to a quasi-continuous population scenario.

Appendix S2: Brief description of unpublished empirical datasets *Toad* and *Salamander*.

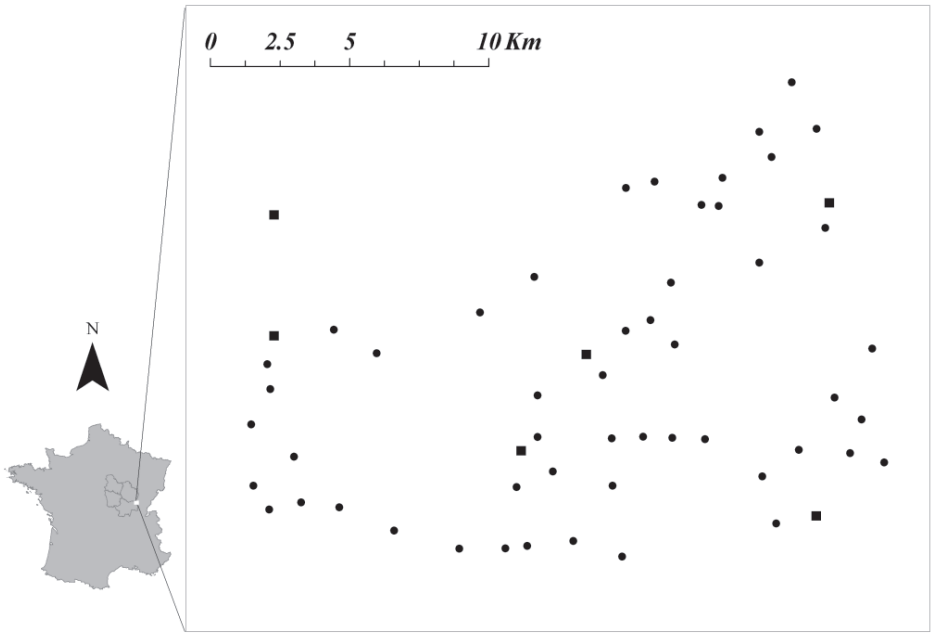
The dataset *Salamander* consisted of 249 genotypes in *Salamandra salamandra*, a widespread pond-breeding amphibian in Seine-et-Marne, France. Fieldwork was conducted during the 2010 breeding season in 9 ponds. The average distance between ponds was 6.51 km, ranging from 0.45 to 12.3 km. In each pond, 19 to 33 adults were trapped and morphologically sexed. Non-destructive genetic samples were then taken from each individual, using non-sterile buccal swabs (Broquet *et al.* 2007). DNA extraction and genotyping were performed as in Bifolchi *et al.* (2010) using 7 microsatellite loci *Sal3*, *SalE5*, *SalE7*, *SalE12*, *SalE14*, *Sale29* and *Sal23* (Steinfartz *et al.* 2004).

The dataset *Toad* consisted of 384 genotypes in *Bufo bufo*, a widespread generalist pond-breeding amphibian in Isère, France. Fieldwork was conducted during the 2010 breeding season in 19 ponds. The average distance between ponds was 23.60 km, with

distances ranging from 2.24 to 59.62 km. In each pond, 12 to 27 adults were trapped and morphologically sexed. Non-destructive genetic samples were then taken from each individual, using non-sterile buccal swabs (Broquet *et al.* 2007). DNA extraction and genotyping were performed using 6 microsatellite loci: *Bbufμ15*, *Bbufμ24*, *Bbufμ39*, *Bbufμ47*, *Bbufμ54* and *Bbufμ65* (Brede *et al.* 2001).

Appendix S3: Localization of the 57 ponds in which at least two alpine newts were sampled (*Newt* dataset).

Localization of the genotypes in dataset *Newt* (46.51°N, 5.17°E, gathered in 2010) in Bourgogne, France. We prospected a total of 223 aquatic sites (man-made or natural ponds, flooded ruts, swamps, etc.). Each site was prospected with a dip net for thirty to sixty minutes (depending on site size and configuration) or until at least two alpine newts *Ichthyosaura alpestris* (a male and a female) were captured. We retained 57 aquatic sites according to an individual-based sampling scheme (a male and a female per aggregate; black circles) and 6 aquatic sites according to a population-based sampling scheme (from 23 to 32 individuals per aggregate, for a total of 69 females and 90 males; black squares); the median distance between neighbouring sampled sites (following a Delaunay triangulation) was 2158m.



Appendix S4: Use of the Bray-Curtis coefficient applied to individual genotypes or to allelic frequencies.

To compute the Bray-Curtis dissimilarity coefficient applied to allelic frequencies (B_{cf}), we produced matrices with *p* rows representing each sampled population, and A columns (see Table 1), each one corresponding to frequencies of an allele in sampled populations. We then applied the Bray-Curtis formula (Legendre & Legendre 1998) for each pair of populations.

CHAPITRE 2

Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*).

Avant-propos.

Dans toute étude génétique, il est indispensable de disposer d'un nombre conséquent de marqueurs génétiques fiables et polymorphes. Dans le cas du triton alpestre *Ichthyosaura alpestris*, servant de modèle biologique pour les études de cas présentées dans les chapitres 3 et 4, sept marqueurs microsatellites ont été développés en 2003 par Garner *et al*. Toutefois, trois de ces marqueurs se sont révélés porteurs d'allèles nuls (Pabijan & Babik 2006; Emaresi *et al.* 2011), c'est-à-dire d'allèles non amplifiables en raison de la présence de mutations au niveau de la séquence complémentaire de l'une ou l'autre des amorces (Pompanon *et al.* 2005). Le développement de nouveaux marqueurs microsatellite était donc un préalable incontournable pour l'étude de l'influence du paysage sur la dispersion de cette espèce.

Par ailleurs, toute analyse génétique doit également porter sur un grand nombre d'individus (plus de 1000 individus ont ainsi été échantillonnés dans le cadre de mes travaux) : la procédure de collecte d'ADN doit donc être rapide, efficace et la moins invasive possible. Employé systématiquement dans le cadre de mes campagnes de terrain, l'échantillonnage buccal (*buccal swabbing ;* Pidancier *et al.* 2003; Broquet *et al.* 2007), s'il s'est révélé très efficace pour la collecte d'ADN, entraînait toutefois très régulièrement un léger saignement des individus échantillonnés (Pidancier *et al.* 2003). La mise au point de nouveaux marqueurs microsatellites a donc été l'occasion de tester l'efficacité d'une procédure de prélèvement d'ADN par frottis de l'épiderme (*skin swabbing),* inspirée des méthodes de détection du champignon pathogène *Batrachochytrium dendrobatidis* (Hyatt *et al.* 2007; Soto-Azat *et al.* 2009) et limitant au maximum la manipulation des animaux.

Cet article a été publié dans *Molecular Ecology Resources* :

Prunier J., Kaufmann B., Grolet O., Picard D., Pompanon F., Joly P. (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*). *Molecular Ecology Resources* **12**, 524-531.

PERMANENT GENETIC RESSOURCES

Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*)

J. PRUNIER[1, 2], B. KAUFMANN [2], O. GROLET [2], D. PICARD [3], F. POMPANON [4] and P. JOLY [2]

[1] *Ecosphère, 3bis rue des Remises, 94100 Saint Maur des Fossés, France*

[2] *Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.*

[3] *Université d'Angers, Laboratoire d'Etudes Environnementales des Systèmes Anthropisés, 2 Boulevard Lavoisier, 49045 Angers Cedex 01, France*

[4] *Université Joseph Fourier, Laboratoire d'Ecologie Alpine, CNRS, UMR 5553, 2233 Rue de la Piscine, 38041 Grenoble Cedex 9, France*

Keywords

454 sequencing, buccal swabs, skin swabs, DNA sampling, *Hyla arborea*, *Ichthyosaura alpestris*

Correspondence: Jérôme Prunier, Fax : +33 145 11 24 37 ;

Email : jerome.prunier@ecosphere.fr

Running title: Skin swabs for DNA sampling in amphibians

Abstract

This study introduces a novel DNA sampling method in amphibians using skin swabs. We assessed the relevancy of skin swabs relevancy for genetic studies by amplifying a set of 17 microsatellite markers in the alpine newt *Ichthyosaura alpestris*, including 14 new polymorphic loci, and a set of 11 microsatellite markers in *Hyla arborea*, from DNA collected with buccal swabs (the standard swab method), dorsal skin swabs and ventral skin swabs. We tested for quality and quantity of collected DNA with each method by comparing electrophoresis migration patterns. The consistency between genotypes obtained from skin swabs and buccal swabs was assessed. Dorsal swabs performed better than ventral swabs in both species, possibly due to differences in skin structure. Skin swabbing proved to be a useful alternative to buccal swabbing for small or vulnerable animals: by drastically limiting handling, this method may improve the trade-off between the scientific value of collected data, individual welfare and species conservation. In addition, the 14 new polymorphic microsatellites for the alpine newt will increase the power of genetic studies in this species. In four populations from France (*n*=19 to 25), the number of alleles per locus varied from 2 to 16 and expected heterozygosities ranged from 0.04 to 0.91. Presence of null alleles was detected in two markers and two pairs displayed gametic disequilibrium. No locus appeared to be sex linked.

Introduction

Although widely distributed in most ecosystems, amphibians are the vertebrate group with the highest proportion of species threatened with extinction (Beebee & Griffiths 2005). Among the identified threats to amphibian populations, habitat fragmentation is considered crucial, impacting both demographic processes and genetic diversity (Cushman 2006). DNA based methods using neutral genetic markers such as microsatellites offer an efficient way to assess the effects of fragmentation on biological processes and long-term population persistence by allowing the estimation of effective population sizes and gene flow (Holderegger & Wagner 2008; Balkenhol *et al.* 2009a; Storfer *et al.* 2010). To increase inferential power, genetic approaches involve both the sampling of many loci across the genome (Selkoe & Toonen 2006), and the sampling of many individuals (Manel *et al.* 2003). In the first case, conservation biologists can rely on new sequencing technologies for identifying and developing new polymorphic microsatellite loci (Perry & Rowe 2011). In the latter case, because genetic sampling may have significant impacts on animals, field studies should be designed according to trade-offs between scientific value of collected data and individual or species welfare (Parris *et al.* 2010). Researchers should use the least destructive and least invasive methods, especially when sampling endangered, vulnerable, or declining amphibian species (Pidancier *et al.* 2003).

There are different methods for genetic sampling in amphibians, each one impacting the animals' welfare in different ways. The toe clipping method (McGuigan *et al.* 1998), which is classically used for individual marking, provide valuable amounts of DNA. This method is rather controversial (Funk *et al.* 2005b) because it may reduce individual survival, especially in climbing species such as tree frogs (McCarthy & Parris 2004). The tail tipping method (Snell & Evans 2006) is usually applied to DNA sampling in anuran tadpoles, but may be used in urodeles at any life-stage (tail- or crest-clipping); however, a proportion of sampled animals may die as a result of stress during sampling, or later because of infection or reduced mobility (Parris *et al.* 2010). Collecting the entire tadpoles may have little impact on species welfare, but it obviously leads to the death of sampled individuals and implies being able to identify species (Parris *et al.* 2010). Buccal swabbing

is an alternative method for genetic sampling that can provide enough DNA for microsatellite genotyping for a range of amphibian species (Pidancier *et al.* 2003; Broquet *et al.* 2007). However, collecting buccal cells with cotton swabs requires levering open the upper and lower jaw with a sterile spatula. Collecting buccal cells and opening the jaw may lead to an amount of bleeding (Pidancier *et al.* 2003). Some species, such as *Hyla arborea*, are easily handled as they tend to keep their jaws opened during sampling, whereas species such as *Triturus cristatus* usually try to keep their mouth closed, and can be easily injured with either rigid tape or cotton swabs. A sampling technique currently used to detect chytrids (*Batrachochytrium dendrobatidis*), a cutaneous pathogenic fungus associated with amphibian mass mortalities and population declines worldwide (Berger *et al.* 1998), relies on sloughing skin removal with skin swabs (Kriger *et al.* 2006; Soto-Azat *et al.* 2009). Although skin swabs can provide fungal DNA, this sampling technique was never used, to our knowledge, to study DNA from the sampled individuals themselves. Nevertheless, skin swabs may minimize stressful handling of animals and may thus overcome ethical issues linked to other DNA sampling methods.

This paper presents an alternative genetic sampling method for amphibians, which meets ethical expectations when studying these vulnerable species. We assessed the efficiency of skin swabs compared to buccal swabs to collect genetic data, in a urodele (*Ichthyosaura alpestris*) and an anuran (*H. arborea*). This work is also an opportunity to present a supplementary set of polymorphic microsatellite loci in the alpine newt *I. alpestris*, a widespread species in France sharing many ecological characteristics with amphibians such as *Lissotriton vulgaris* or the endangered *T. cristatus* (Emaresi *et al.* 2011). We detail the conditions for successful multiplex polymerase chain reaction (PCR) amplifications of the new markers in addition to three markers developed by Garner *et al.* (2003), and present results of cross-species amplifications in two other urodele species: *Lissotriton helveticus* and *T. cristatus*.

Material and methods

Study area and DNA sampling

All samples were collected in France during the breeding season. To develop microsatellite markers, muscle and liver samples were collected from a single *I. alpestris* male found moribund on a road in 2009 and stored in 95% ethanol (code GL in Table 1). All other individuals were trapped in breeding ponds in 2009 or 2010 and sampled using buccal and/or skin swabs. Buccal sampling was performed by opening the animals' mouth with a sterile rigid spatula (Pidancier *et al.* 2003) and swabbing the surfaces of the buccal cavity with an ordinary cotton bud (Poschadel & Moller 2004). The skin swab technique was inspired by Kriger *et al.* (2006) and performed by firmly running an ordinary cotton bud three times over the dorsal (dorsal skin swab) or ventral surface (ventral skin swab).

Skin swab sampling efficiency was assessed using 23 alpine newts (*I. alpestris*), captured in 2010 in site F, and 4 European tree frogs (*H. arborea*), captured in 2010 in site I (Table 1). Each individual was sampled using buccal swabs and dorsal skin swabs. Eight alpine newts and the 4 tree frogs were also sampled using ventral skin swabs. In frogs, the abdomen was too dry to allow a sample, and swabs were not considered for DNA extraction. Individuals were immediately released after the sampling, except 4 alpine newts and the 4 European tree frogs that were kept in observation for a month. All swabs were kept at ambient temperature (around 15-20°C) in a hermetic container with silica gel providing total desiccation.

For the estimation of microsatellite genetic variability in natural populations, we sampled 90 individuals in four distinct breeding ponds, located from 2 to 286 km apart, with 19 to 25 alpine newts at each site (codes CN, CW, X43 and Z23 in Table 1). Each individual was morphologically sexed, sampled using buccal swabs and immediately released. We also tested microsatellite markers for cross-species amplification using 23 individuals of 2 other urodele species: *L. helveticus* (n = 12) and *T. cristatus* (n = 11). The 23 individuals were randomly collected across Bourgogne (France) during an independent fieldwork campaign (Table 1). They were sampled using buccal swabs and immediately released.

Table 1 Sampling characteristics: Type of protocol (DMM: Development of Microsatellite Markers; EGV: Estimation of Genetic Variability; SSE: Skin Sampling Efficiency; CSA: Cross-Species Amplification), region, code and geographical coordinates of each sampled population, species sampled, number of individuals (N) and sample type (T: Tissues; B: Buccal swabs; D: Dorsal swabs; V: Ventral swabs). For the CSA protocol, geographical coordinates (in bold) correspond to the northwestern and southeastern corners of the corresponding sampling area.

Protocol	Region	Code	Latitude	Longitude	Species	N	Sample type
DMM	Rhône-Alpes	GL	45°25'13.3''N	5°24'56.0''E	*I. alpestris*	1	T
EGV	Ile-de-France	CN	48°42'02.7''N	2°43'10.2''E	*I. alpestris*	22	B
EGV	Ile-de-France	CW	48°41'12.3''N	2°42'24.1''E	*I. alpestris*	19	B
EGV	Bourgogne	X43	46°52'07.3''N	5°11'47.8''E	*I. alpestris*	25	B
EGV	Franche-Comté	Z23	46°54'25.1''N	5°27'31.5''E	*I. alpestris*	24	B
SSE	Rhône-Alpes	F	45°46'47.7''N	4°52'04.5''E	*I. alpestris*	23	B, D, V
SSE	Rhône-Alpes	I	45°44'18.0''N	5°21'06.3''E	*H. arborea*	4	B, D
CSA	Bourgogne		**47°39'41.0''N**	**3°02'00.0''E**	*T. cristatus*	11	B
CSA	Bourgogne		**46°46'24.0''N**	**5°29'23.0''E**	*L. helveticus*	12	B

Development of microsatellite markers

Genomic DNA was extracted from tissue samples collected at GL site (Table 1), following Nurnberger *et al.* (2003). To identify alpine newt specific microsatellites, genomic DNA was sequenced using 454 FLX pyrosequencing technology (Abdelkrim *et al.* 2009) at the Savannah River Ecology Laboratory (University of Georgia). Raw DNA sequences were cleaned of remnant vector and screened for appropriate microsatellite motifs using the program MSATCOMMANDER (Faircloth 2008). Among 1015 fragments containing.

Microsatellite genotyping

DNA extractions from swabs were performed using a QIAGEN DNeasy Tissue Kit (QIAGEN), following the manufacturer's instructions. DNA was eluted in a final volume of

100 µL of buffer BE (QIAGEN). DNA concentrations were estimated using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc) and each sample was treated by dilution or evaporation with a Speed-Vac apparatus (*Concentrator plus*, Eppendorf) to reach 20 ng/µL.

For *I. alpestris*, the 14 developed markers were studied together with *Ta1Ca1*, *Ta2Caga3* and *Ta1Caga4* microsatellite loci (Garner *et al.* 2003). Among the 7 primer pairs developed by Garner *et al.* (2003), only those 3 microsatellite loci worked well in our experimental conditions. The 17 primer pairs were used to constitute 4 sets of 4 to 5 markers for multiplex PCR amplification (Table 2). PCR was conducted on PTC-100 Thermal Cycler (MJ Research, Inc.) using Type-it™ Microsatellite PCR Kit (QIAGEN) in a 10 µL volume containing 5 µL QIAGEN multiplex PCR Master Mix, 3 µL of water, 1 µL of primer mix (2 µM of both forward and reverse primers) and 1 µL of genomic DNA at 20 ng/µL. Cycling conditions were the following: 95°C for 5 minutes; 32 cycles of denaturation at 95°C (30 s), annealing at 60°C (90 s), and extension at 72°C (30 s); lastly, 60°C for 30 minutes. The forward of each primer pair was labeled with a fluorescent dye: FAM or HEX (Sigma Aldrich), NED or PET (Applied Biosystems). Cross-species amplifications in *L. helveticus* and *T. cristatus* were attempted using the PCR conditions optimized for *I. alpestris.*

In *H. arborea,* PCR amplifications were performed with 2 sets of 5 and 6 primer pairs (set 1: Ha-B5R3, Ha-A119, WHA1-2, Ha-D3R3, Ha-A11, Ha-A127 ; set 2: D106, H116, Ha-D115, A136, Ha-A130; Arens *et al.* 2000; Berset-Brandli *et al.* 2008a) at the following cycling conditions: 95°C for 15 minutes; 30 cycles of denaturation at 94°C (30 s), annealing at 59°C (90 s) and extension at 72°C (30 s); lastly, 60°C for 30 minutes.

The PCR products were diluted 50 times and discriminated using capillary electrophoresis on a 3730xl DNA Analyser (Applied Biosystems). For each locus, alleles were scored with GeneMarker 1.95 (SoftGenetics), using GS600 LIZ size standard (Applied Biosystems). Ambiguous genotypes were amplified and sized a second time.

We checked the presence of null alleles in the four populations CW, CN, X43 and Z23 (Table 1) by analysing homozygote excess with MICROCHECKER (Van Oosterhout *et al.* 2004) and by calculating the number of individuals apparently homozygous for a null allele, that is which repeatedly failed to amplify any alleles at just one locus while all other loci amplified normally (Selkoe & Toonen 2006). We estimated the number of alleles per

locus, observed and expected heterozygosity, and checked Hardy-Weinberg equilibrium and gametic disequilibrium with GENEPOP 4.0 (Rousset 2008). Tests for Hardy–Weinberg equilibrium (HWE) were conducted using the sequential Bonferroni correction to account for multiple-related tests (Rice 1989).

Table 2 Fourteen microsatellite loci from *I. alpestris* with PCR primers, repeat motifs, size ranges (bp: base pair) and accession number.

Locus	Primer (5'-3')	Repeat motif	Size range (bp)	Accession number
PCR multiplex set 1				
CopTa1	F: (FAM)-CATGAGGGTATGGGTGGTCAGC	(AC)10	85 - 93	JN048427
	R: CCATTGCCCACAATGCAAACTG			
CopTa2	F: (HEX)-ACAGGAACTTACAGCCTACCC	(AC)10	88- 96	JN048428
	R: AGCAGTCCCTGGCTATTTAAG			
CopTa3	F: (NED)-AGTCACTTAAGGCCACAGGG	(AC)11	75 - 127	JN048429
	R: CTAAAGCTCCTTCATGGGAGGC			
CopTa4	F: (PET)-TCTTCCTCCTAGACCCTTGTGG	(AC)12	196 - 206	JN048430
	R: GACACAGTAGATTGGCAAGTGG			
PCR multiplex set 2 (completed with Ta1Caga4)				
CopTa5	F: (HEX)-CTGGCATGAGTGGCCTTGGTTC	(AC)10	48 - 67	JN048431
	R: TGTCTGTTTGGGAAAGCACAG			
CopTa6	F: (FAM)-AACCTACAGATCACCCTTTCC	(AATC)11	70 - 87	JN048432
	R: CCCTCGTGTGCCTTGAGACCC			
CopTa7	F: (FAM)-CGAGCGGAACTACAATTGAAAC	(AGAT)10	188 - 231	JN048433
	R: CTCCCGCAGCACGGAGATACC			
PCR multiplex set 3				
CopTa8	F: (NED)-ACATAAGCTGAGCAAACCATGC	(AC)10	86 - 116	JN048434
	R: GGTCTTACAGTTCGATTTACTGTGG			
CopTa9	F: (HEX)-ACACTACCAATTTCTGAACGC	(AGAT)14	189 - 262	JN048435
	R: TGAAGTGGAAGTTACATCGGG			
CopTa10	F: (FAM)-GCACGAGCGGAACTCCTTCTG	(AGAT)16	185 - 295	JN048436
	R: TTCCCACAGCTTACCTGAGCAC			
CopTa11	F: (PET)-GACTGTCAGAGAACCACTTACC	(ACTC)15	172 - 223	JN048437
	R: GTGTGTGTAGAACTGCCTCC			
PCR multiplex set 4 (completed with Ta1Ca1 and Ta2Caga3)				
CopTa12	F: (HEX)-CTTTGCATGGGAAACAAAGGCG	(AC)12	68 - 72	JN048438
	R: CCCTTGCAAACAGTGTATAGG			
CopTa13	F: (HEX)-GGGACACAGGAAATGAGACAGGC	(AG)11	192 - 204	JN048439
	R: GAACCATTAAGCGTGTCCCTGC			
CopTa14	F: (FAM)-GTGGGATGTATGTTGGATTTCAC	(AC)10	199 - 205	JN048440
	R: TAAGAGGGCTTCAGGGACAGTGG			

Skin sampling efficiency

The inconsistency between genotypes obtained from different swabbing methods was assessed by calculating e_l, the ratio between the number of single-locus genotypes obtained from ventral (e_{lv}) or dorsal (e_{ld}) swabs including at least one allelic difference with genotypes obtained from buccal swabs and the total number of single-locus genotypes. This index is analogous to e_l, the mean error rate per locus (Pompanon *et al.* 2005) using the genotypes obtained from buccal swabs as reference genotypes. It was calculated for each individual collected in populations F and I (Table 1) and averaged over all individuals in each species. We used the free software R (R Development Core Team 2011) to test for the correlation between e_l (pooling e_{ld} and e_{lv}) and the concentrations of DNA in newt skin swab samples, using the Kendall rank correlation test. We also used a Wilcoxon rank sum test to compare error rates when using dorsal and ventral samples. To test for DNA quality and quantity, we compared migration patterns of extracted DNA at 20 ng/µL on a 3 % agarose gel.

Results

Skin sampling efficiency

In *I. alpestris*, buccal swab sampling provided a large quantity of DNA (152.96 ± 66.32 ng/µL, mean ± sd, n = 23) and enabled genotyping of all individuals. Dorsal swab sampling provided a smaller quantity of DNA (18.13 ± 11.88 ng/µL, mean ± sd, n = 23, see fig. 1). The mean difference per locus between genotypes from dorsal and buccal swabs was e_{ld} = 7.69 % (n = 23) after the first amplification. Differences were mainly due to allelic dropouts, PCR failing to amplify some alleles; only 4 differences out of 30 were due to allelic mismatch or contamination. This value decreased to e_{ld} = 0.51 % (n = 23) after a second independent amplification of loci showing inconsistent genotypes. Then, dorsal swabs gave genotypes consistent with that obtained from buccal swabs for all individuals across all markers but CopTa3, the second independent amplification failing to reveal the shorter allele (75 bp) for 2 individuals. Ventral swab skin sampling provided the smallest quantities of DNA (6.48 ± 5.45 ng/µL, mean ± sd, n = 8, see fig. 1) and, after the first

amplification, led to significantly higher inconsistencies between genotypes (e_{Lv} = 52.21 %, n = 8) than dorsal swab skin sampling (e_{ld} = 7.69 %, n = 23, Wilcoxon rank sum test, V = 24, p = 0.0014). Ventral swab sampling was less efficient than dorsal sampling in this species, and thus not considered for new independent amplifications. The correlation between the mean error rates e_l and the concentrations of DNA in skin swab samples was negative and highly significant (Kendall rank correlation test, tau = -0.601, 2-sided p-value < 10^{-5}), indicating that genotyping efficiency notably depends on the quantity of collected DNA in skin swab samples.

In *H. arborea*, buccal and dorsal swab sampling provided identical genotypes for all individuals from the first amplification (e_{ld} = 0%, n = 4) across all 11 loci. Buccal swab sampling tended to provide a higher quantity of DNA than dorsal swab sampling (respectively 43.69 and 10.95 ng/µL, mean, n = 4, see fig. 1). The 8 individuals that had been sampled using both buccal and skin swabs and kept in observation did not show any visible side-effect.

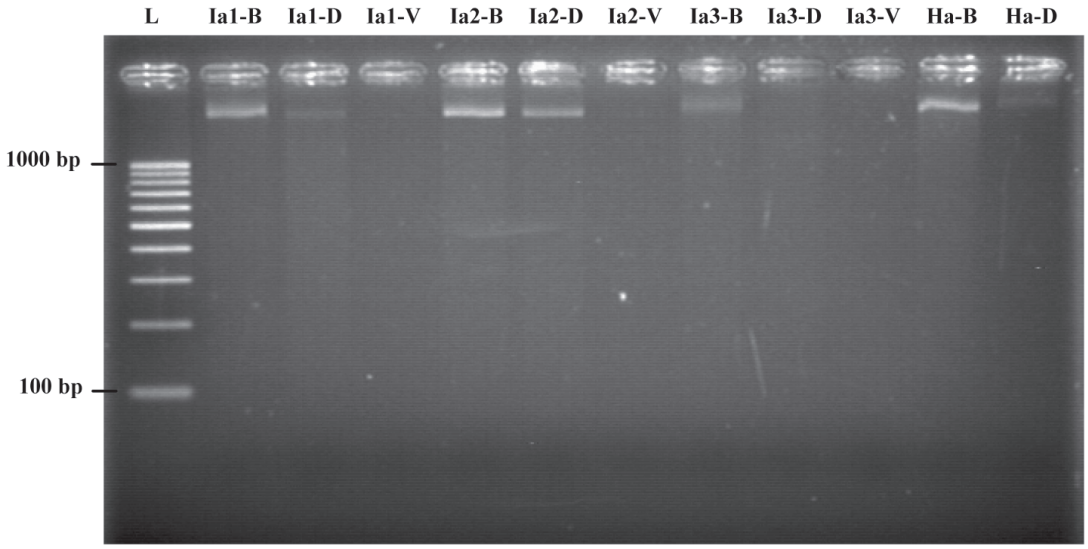


Figure 1 Migration patterns of extracted DNA at 20 ng/µL visualized by ethidium bromide on 3% agarose gel, in *I. alpestris* and *H. arborea*. L: 100 bp DNA ladder; Ia1, Ia2 and Ia3 correspond to samples from three alpine newts (one female and two males respectively); Ha correspond to samples from a male tree frog; -B, -D and -V correspond to Buccal, Dorsal and Ventral samples. Only buccal and dorsal swabs show visible bands, corresponding to high quantity of non-degraded native DNA; lower intensity in dorsal samples indicates lower DNA concentrations of DNA. Absence of a clear band in ventral swabs indicates that observed DNA concentrations are very low. See text for details on concentrations.

Microsatellite markers

None of the 14 developed microsatellite markers (Table 1) appeared to be sex-linked, as heterozygote individuals were found in both sexes at each tested locus. The number of alleles ranged from 2 to 16. Observed heterozygosities (*Ho*) ranged from 0.04 to 0.96 and expected heterozygosities (*He*) ranged from 0.04 to 0.91 (Table 3). No locus departed significantly from Hardy-Weinberg equilibrium, following sequential Bonferroni correction. A trend towards allelic dropout was detected in CopTa3 for the two shorter alleles (75 and 77 bp): In electropherograms, peaks located at 75 or 77 bp are actually systematically smaller than peaks corresponding to longer alleles. For each individual, deciding whether a peak at 75 or 77 bp is an allele or an artifact may be challenging at first sight: any ambiguous genotype, *i.e.* any individual that was not clearly heterozygous for that locus (with two peaks at more than 77 bp), was thus checked with a new amplification that always led to the same electropherogram profile, showing that the DNA concentration (20 ng/µL) enabled a high detection level of these alleles from the first amplification. The presence of null alleles was detected in 2 markers: CopTa5 and CopTa6. Preliminary work on 6 other populations confirmed both the trend towards allelic dropout for the two shorter alleles in CopTa3 and the presence of null alleles in CopTa5 and CopTa6 (data not shown). Significant gametic disequilibrium was detected between CopTa7 and markers CopTa9 and CopTa10.

Using the PCR conditions optimized for *I. alpestris,* cross-species amplification tests revealed only 2 polymorphic markers for *L. helveticus* (CopTa3 and CopTa13), each exhibiting 4 alleles for 8 individuals. No amplification was obtained at any locus for *T. cristatus*.

Table 3 Summary data for twelve microsatellite loci developed for *I. alpestris.* Number of alleles (*A*; in brackets, effective number of alleles), observed and expected heterozygosity (*Ho* and *He*), and fixation index (*f*) are given for each locus for *N* individuals analyzed for 4 populations in France: 2 populations from Seine-et-Marne (CW and CN) and 2 populations from Saône-et-Loire (X43 and Z23). No locus departed significantly from Hardy-Weinberg equilibrium, following sequential Bonferroni correction.

Populations	*N*	*A*	*Ho*	*He*	*f*	*N*	*A*	*Ho*	*He*	*f*
	CopTa1					**CopTa9**				
CW	19	3	0.421	0.496	0.155	19	8	0.842	0.845	0.004
CN	22	2	0.364	0.304	-0.200	22	8	0.864	0.886	0.026
X43	25	3	0.640	0.541	-0.187	25	14	0.960	0.912	-0.054
Z23	24	3	0.417	0.430	0.032	24	13	0.875	0.900	0.028
Total	90	3 (2.76)				90	16 (11)			
	CopTa2					**CopTa10**				
CW	19	4	0.737	0.681	-0.084	19	8	0.632	0.788	0.203
CN	22	4	0.545	0.697	0.221	22	5	0.455	0.673	0.330
X43	25	3	0.800	0.615	-0.310	25	10	0.800	0.811	0.013
Z23	24	4	0.542	0.691	0.219	24	9	0.583	0.801	0.276
Total	90	4 (3.72)				90	15 (8.09)			
	CopTa3					**CopTa11**				
CW	19	8	0.684	0.767	0.110	19	5	0.947	0.781	-0.220
CN	22	6	0.773	0.686	-0.130	22	5	0.636	0.747	0.152
X43	25	7	0.640	0.681	0.061	25	7	0.760	0.673	-0.132
Z23	24	8	0.792	0.769	-0.031	24	4	0.708	0.741	0.045
Total	90	11 (7.23)				90	8 (5.29)			
	CopTa4					**CopTa12**				
CW	19	2	0.158	0.149	-0.059	19	2	0.579	0.508	-0.145
CN	22	2	0.273	0.241	-0.135	22	2	0.455	0.507	0.106
X43	25	3	0.600	0.561	-0.071	25	2	0.160	0.150	-0.067
Z23	24	3	0.500	0.526	0.050	24	2	0.458	0.467	0.019
Total	90	3 (2.54)				90	2 (2)			
	CopTa7					**CopTa13**				
CW	19	10	0.789	0.883	0.109	19	2	0.316	0.478	0.346
CN	22	11	0.818	0.874	0.066	22	3	0.545	0.513	-0.066
X43	25	8	0.760	0.780	0.027	25	4	0.480	0.543	0.118
Z23	24	11	0.833	0.853	0.023	24	4	0.625	0.495	-0.271
Total	90	13 (9.96)				90	4 (3.33)			
	CopTa8					**CopTa14**				
CW	19	3	0.474	0.525	0.100	19	3	0.316	0.317	0.005
CN	22	3	0.727	0.606	-0.207	22	3	0.182	0.172	-0.057
X43	25	2	0.040	0.040	0.000	25	3	0.640	0.575	-0.116
Z23	24	5	0.500	0.483	-0.036	24	3	0.500	0.613	0.188
Total	90	5 (3.26)				90	3 (3)			

Discussion

There are key differences between the skin sampling protocol discussed in that study and the one used to gather and extract chytrid DNA in Kriger *et al.* (2006). In order to optimize the detection of *B. dendrobatidis*, skin has to be swabbed by firmly running the tip of a swab over different body parts (dorsal and ventral surface, but also body sides, underside of thighs and webbing of feet), at least once, but sometimes up to ten times (Kriger *et al.* 2007); swabs have then to be frozen at -20°C. To get host DNA, which can be found obviously on any part of the body, we advise to firmly run the tip of a swab only three times over the skin and to concentrate the sampling on a single surface (dorsal or ventral surface) so as to limit the handling of animals. This protocol also offers practical advantages in the field: as genetic studies often rely on species-specific markers, cotton buds do not necessarily require to be acquired sterile (Poschadel & Moller 2004), as long as inter-individual contaminations are avoided. Besides, samples do not need to be frozen, but can be stored at ambient temperature in a hermetic container with silica gel providing total desiccation.

In both *I. alpestris* and *H. arborea*, dorsal swabs provided satisfactory amounts of DNA, and correct genotypes for all individuals at all loci except one. Allelic dropouts were only observed at locus CopTa3. Collecting a lower quantity of DNA in superficial skin cells may exacerbate amplification difficulties already observed with the two shortest alleles at this locus. Overall, dorsal swab sampling proved to be an efficient method to provide satisfactory DNA quantity, drastically limiting the handling of individuals, but some markers may not work well with the limited amount of DNA recovered with this method.

Ventral swabs in *H. arborea* were not considered for DNA extraction, because their abdomen was too dry to enable recovery of skin cells from gentle sampling. Ventral swabs in *I. alpestris* failed to provide DNA in satisfactory quantities, when compared to dorsal swabs. Cotton swabs may slip over the smooth ventral skin and only collect dead sloughing cells with highly degraded DNA, whereas the grainy dorsal skin may enable better sampling of epidermic cells. Interestingly, although chytrid detection implies swabbing of the abdomen, this kind of sampling did not lead to satisfactory amounts of DNA in our study. Because DNA quantity and quality may vary between amphibian

species according to the localization of the sampled skin (ventral or dorsal), we recommend (1) to sample preferentially grainy skin surfaces, ensuring a better abrasive rubbing, and (2) to use preliminary sampling tests on the study species, to determine the best body region for skin swabs.

Although buccal swabs in amphibians remain the reference method in terms of amount and quality of recovered DNA, necessary in this study for calculation of genotyping efficiency, skin swabs may be an interesting alternative method when sampling small individuals (juveniles) or species likely to be hurt during handling or buccal swab sampling. Skin swabbing makes DNA sampling faster, and might considerably reduce handling and sampling time, which is useful when studying large numbers of individuals. To increase the utility of this protocol, and because the understanding of prevalence and intensity of chytridiomycosis is fundamental for the conservation of amphibians, we also advise researchers to take two skin swabs, the first one intended for host DNA studies, concentrating the swabbing on a grainy skin surface, the second one saved for chytrid detection, using the swabbing on various part of animal, following the standard protocol of Kriger *et al.* (2006). It might also be useful to consider that skin swab samples collected for chytrid detection may be a used to obtain host DNA for microsatellite studies.

The levels of polymorphism found in the new microsatellite markers developed in *I. alpestris* will increase the power of genetic studies in this species. Using the described amplification conditions, these markers will not be useful in *T. cristatus.* Two markers may be used in *L. helveticus*, but more specific PCR conditions may improve their performance. Markers CopTa5 and CopTa6 should be used with caution, as null alleles may lead to genotyping errors and erroneous interpretations of observed homozygote excess in some populations (Pompanon *et al.* 2005; Selkoe & Toonen 2006). Markers with significant gametic disequilibrium should be discarded from analysis in which loci are assumed to be independent samples of the genome (Selkoe & Toonen 2006). However, new descriptive methods, such as spatial principal component analysis (sPCA), do not require linkage equilibrium or Hardy-Weinberg equilibrium (Jombart *et al.* 2008) and may thus benefit from an additional highly polymorphic marker such as CopTa7.

References

Abdelkrim J, Robertson BC, Stanton J-AL, Gemmell NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* **46**, 185-191.

Arens P, Van't Westende W, Bugter R, Smulders MJM, Vosman B (2000) Microsatellite markers for the European tree frog Hyla arborea. *Molecular Ecology* **9**, 1944-1946.

Balkenhol N, Gugerli F, Cushman SA*, et al.* (2009) Identifying future research needs in landscape genetics: where to from here? *Landscape Ecology* **24**, 455-463.

Beebee TJC, Griffiths RA (2005) The amphibian decline crisis: A watershed for conservation biology? *Biological Conservation* **125**, 271-285.

Berger L, Speare R, Daszak P*, et al.* (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9031-9036.

Berset-Brandli L, Jaquiery J, Broquet T, Perrin N (2008a) Isolation and characterization of microsatellite loci for the European tree frog (Hyla arborea). *Molecular Ecology Resources* **8**, 1095-1097.

Berset-Brandli L, Jaquiery J, Broquet T, Ulrich Y, Perrin N (2008b) Extreme heterochiasmy and nascent sex chromosomes in European tree frogs. *Proceedings of the Royal Society B-Biological Sciences* **275**, 1577-1585.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics* **8**, 509-511.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation* **128**, 231-240.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conservation Genetics* **12**, 41-50.

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**, 92-94.

Funk WC, Donnelly MA, Lips KR (2005) Alternative views of amphibian toe-clipping. *Nature* **433**, 193-193.

Garner TWJ, Schmidt BR, Hoeck P, Van Buskirk J (2003) Di- and tetranucleotide microsatellite markers for the Alpine newt (*Triturus alpestris*): characterization and cross-priming in five congeners. *Molecular Ecology Notes* **3**, 186-188.

Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* **58**, 199-207.

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

Kriger KM, Hines HB, Hyatt AD, Boyle DG, Hero JM (2006) Techniques for detecting chytridiomycosis in wild frogs: comparing histology with real-time Taqman PCR. *Diseases of Aquatic Organisms* **71**, 141-148.

Kriger KM, Pereoglou F, Hero JM (2007) Latitudinal variation in the prevalence and intensity of chytrid (Batrachochytrium dendrobatidis) infection in Eastern Australia. *Conservation Biology* **21**, 1280-1290.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.

McCarthy MA, Parris KM (2004) Clarifying the effect of toe clipping on frogs with Bayesian statistics. *Journal of Applied Ecology* **41**, 780-786.

McGuigan K, McDonald K, Parris K, Moritz C (1998) Mitochondrial DNA diversity and historical biogeography of a wet forest-restricted frog (Litoria pearsoniana) from mid-east Australia. *Molecular Ecology* **7**, 175-186.

Nurnberger B, Hofman S, Forg-Brey B*, et al.* (2003) A linkage map for the hybridising toads Bombina bombina and B-variegata (Anura : Discoglossidae). *Heredity* **91**, 136-142.

Parris KM, McCall SC, McCarthy MA*, et al.* (2010) Assessing ethical trade-offs in ecological field studies. *Journal of Applied Ecology* **47**, 227-234.

Perry JC, Rowe L (2011) Rapid Microsatellite Development for Water Striders by Next-Generation Sequencing. *Journal of Heredity* **102**, 125-129.

Pidancier N, Miquel C, Miaud C (2003) Buccal swabs as a non-destructive tissue sampling method for DNA analysis in amphibians. *Herpetological Journal* **13**, 175-178.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-859.

Poschadel JR, Moller D (2004) A versatile field method for tissue sampling on small reptiles and amphibians, applied to pond turtles, newts, frogs and toads. *Conservation Genetics* **5**, 865-867.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Rice WR (1989) Analysing tables of statistical tests. *Evolution* **43**, 223-225.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* (ed. Krawetz S MS), pp. 365-386. Humana Press, Totowa, NJ.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Snell C, Evans IH (2006) Discrimination of moor frog (Rana arvalis) and common frog (Rana temporaria) individuals using a RAPD technique. *Herpetological Journal* **16**, 363-369.

Soto-Azat C, Clarke BT, Fisher MC, Walker SF, Cunningham AA (2009) Non-invasive sampling methods for the detection of Batrachochytrium dendrobatidis in archived amphibians. *Diseases of Aquatic Organisms* **84**, 163-166.

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

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Acknowledgements

This work was supported by the ANRT (Agence Nationale de la Recherche et de la Technologie, agreement n° 2/2009), RFF (Réseau Ferré de France), APRR (Autoroutes Paris-Rhin-Rhône) and the ECOSPHERE company. We are grateful to S. Lance for developing the microsatellite library. We warmly thank Dr T.C. Glenn and three anonymous reviewers for insightful suggestions.

Data Accessibility

DNA sequences: Genbank accessions JN048427- JN048440

Supporting Information

Table S1 Values estimated for the 27 individuals used for the analysis of the skin sampling efficiency.

Table S1 Values estimated for the 27 individuals used for the analysis of the skin sampling efficiency. Ind : individual code; SM: Swabbing method (B, D and V stand for Buccal, Dorsal and Ventral swabs respectively); C: Concentration (ng/µL) of collected DNA for each swabbing method; e_{ld1}: mean error rate per locus between the dorsal genotype and the reference buccal genotype after the first amplification; e_{ld2}: mean error rate per locus between the dorsal genotype and the reference buccal genotype after the second independent amplification; e_{lv}: mean error rate per locus between the ventral genotype and the reference buccal genotype; DL: loci concerned by allelic dropout in dorsal genotypes after the second independent amplification. (/ = No data).

Species	Ind.	Sex	SM	C	e_{ld1}	e_{ld2}	e_{lv}	DL
			B	173.23				
I. alpestris	TaPA1	F	D	5.77	47%	0%	88%	/
			V	3.2				
			B	155.27				
I. alpestris	TaPA2	F	D	4.55	24%	0%	88%	/
			V	3.51				
			B	197.6				
I. alpestris	TaPA3	F	D	4.09	18%	0%	18%	/
			V	2.84				
			B	125.14				
I. alpestris	TaPA4	M	D	9.09	41%	0%	88%	/
			V	4.02				
			B	96.57				
I. alpestris	TaPA5	M	D	5.65	6%	0%	29%	/
			V	14.23				
			B	116.75				
I. alpestris	TaPA6	M	D	15.42	0%	0%	71%	/
			V	3.23				
			B	58.88				
I. alpestris	TaPA7	M	D	6.76	12%	0%	35%	/
			V	4.59				
			B	131.07				
I. alpestris	TaPA8	M	D	12.1	6%	0%	0%	/
			V	16.22				
I. alpestris	TaA1	F	B	170.86	0%	0%	/	/
			D	29.15				
I. alpestris	TaA2	M	B	115.82	0%	0%	/	/
			D	19.37				
I. alpestris	TaA3	F	B	164.06	12%	6%	/	CopTa3
			D	25.28				
I. alpestris	TaA4	F	B	259.65	0%	0%	/	/
			D	31.22				
I. alpestris	TaA5	F	B	294.57	0%	0%	/	/

			D	12.13				
I. alpestris	TaA6	M	B D	247.65 26.26	0%	0%	/	/
I. alpestris	TaA7	M	B D	148.09 48.42	0%	0%	/	/
I. alpestris	TaA8	M	B D	243.96 18.08	0%	0%	/	/
I. alpestris	TaA9	M	B D	230.18 27.79	6%	6%	/	CopTa3
I. alpestris	TaA10	I	B D	99.75 6	6%	0%	/	/
I. alpestris	TaA11	M	B D	99.43 10.95	0%	0%	/	/
I. alpestris	TaA12	F	B D	82.52 16.06	0%	0%	/	/
I. alpestris	TaA13	M	B D	158.21 19.85	0%	0%	/	/
I. alpestris	TaA14	M	B D	62.8 23.22	0%	0%	/	/
I. alpestris	TaA15	F	B D	86.12 39.84	0%	0%	/	/
H. arborea	HA1	M	B D	34.08 2.52	0%	/	/	/
H. arborea	HA2	M	B D	44.31 27.05	0%	/	/	/
H. arborea	HA3	F	B D	24.79 6.29	0%	/	/	/
H. arborea	HA4	F	B D	71.59 7.95	0%	/	/	/

CHAPITRE 3

Permeability of large transport infrastructures to gene flow in the alpine newt *Ichthyosaura alpestris*.

Avant-propos.

Les infrastructures de transport, telles que les autoroutes et les voies ferrées, sont l'une des principales sources de fragmentation du paysage, et leurs impacts potentiels sur la faune et la flore sont largement documentés (Forman 2000; Trombulak & Frissell 2000). De par leurs dimensions exceptionnelles, les grandes infrastructures constituent un enjeu régional (voire national ou international) en terme de conservation : caractériser leur impact sur la biodiversité afin de mettre en place des mesures adéquates de réduction ou de compensation sur les ouvrages actuels, mais également prévoir la « transparence » des infrastructures de demain, est au cœur d'une discipline en plein essor : l'écologie routière (*road ecology* ; Lesbarreres & Fahrig 2012).

L'apport de la génétique du paysage à l'écologie routière (*molecular road ecology* ; Balkenhol & Waits 2009) est substantiel (Holderegger & Di Giulio 2010). En effet, le suivi direct de la dispersion des organismes aux abords des infrastructures de transport peut s'avérer délicat, quand une approche génétique autorise une estimation indirecte et à grande échelle des flux de dispersion. Le premier chapitre de ce manuscrit a permis de mettre en évidence les apports potentiels d'une approche individu-centrée pour la détection de processus d'isolement par barrières aux flux de gènes (*isolation-by-barrier*, IBB): les simulations ont ainsi montré qu'une augmentation de la couverture d'échantillonnage du paysage, couplée à un échantillonnage ciblé, s'avère toute à fait pertinente. Ce nouveau chapitre présente une mise en application concrète d'un tel plan d'échantillonnage, dans le cadre de l'étude des impacts d'une autoroute et d'une ligne à grande vitesse sur la dispersion du triton alpestre *Ichthyosaura alpestris* en Bourgogne (France).

Cet article est en cours de révision pour publication dans <u>*Molecular Ecology*</u> :

Prunier, J., Kaufmann, B., Lena, J.P., Pompanon, F., Joly, P. Permeability of large transport infrastructures to gene flow in the alpine newt *Ichthyosaura alpestris*.

Permeability of large transport infrastructures to gene flow in the alpine newt *Ichthyosaura alpestris*.

J. PRUNIER [1, 2], B. KAUFMANN [2], J.P. LENA [2], F. POMPANON [3] and P. JOLY [2]

[1] *Ecosphère, 3bis rue des Remises, 94100 Saint Maur des Fossés, France*

[2] *Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.*

[3] *Université Joseph Fourier, Laboratoire d'Ecologie Alpine, CNRS, UMR 5553, 2233 Rue de la Piscine, 38041 Grenoble Cedex 9, France*

Keywords: amphibian, Bayesian clustering methods, landscape genetics, local edge detection methods, motorway, railway, sPCA.

Corresponding author:

Jérôme Prunier

Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.

Fax : (+33) 4 72 43 11 41

Email: jerome.prunier@gmail.com

Running title: Molecular road ecology in the alpine newt.

Abstract

Large transport infrastructures (LTI) are known to restrict animal movements through landscape fragmentation and may therefore impact genetic patterns in terrestrial organisms. We assessed the effect of a 40 year-old motorway and a 30 year-old high-speed railway on the spatial genetic structure of the Alpine newt *Ichthyosaura alpestris*, following an individual-based sampling scheme and using a diverse set of statistical approaches. Spatial principal component analyses (sPCA), a method designed to reveal cryptic genetic spatial patterns in high gene flow species, provided an intelligible overview of spatial genetic structure, while other methods only revealed pieces of the puzzle. LTIs never acted as barriers to gene flow, but several genetic boundaries coinciding with transition zones between major landscape entities were detected. The motorway counter-intuitively appeared as a potential dispersal corridor in low-quality habitats, challenging traditional hypotheses on road impacts in amphibians.

Introduction

Transportation infrastructures constitute one of the most widespread forms of land use in contemporary landscapes, and have been the subject of numerous studies aiming to assess their impacts on natural populations (Trombulak & Frissell 2000; Balkenhol & Waits 2009). Roads are known to act as barrier to dispersal and gene flow (isolation-by-barrier, *IBB*), decreasing functional connectivity and increasing the genetic differentiation among individuals on each side of the infrastructure (Balkenhol & Waits 2009; Holderegger & Di Giulio 2010). The larger the size of a transportation infrastructure (in terms of traffic volume and spatial dimensions), the higher is the expected impact on both landscape configuration (Trombulak & Frissell 2000) and dispersal patterns (Holderegger & Di Giulio 2010). Large transport infrastructures (LTIs) such as motorways and high-speed railways are thus of major concern in conservation biology. Their ecological impact on dispersal may be assessed at two complementary spatial scales: first in the immediate vicinity of lanes, and second in a surrounding "road-effect zone" (Forman 2000; Balkenhol & Waits 2009). At the scale of lanes, barrier effects may be explained by mortality due to collisions, which obviously decreases gene exchanges across roads (Holderegger & Di Giulio 2010), and they may also be explained by the modification of effective dispersal patterns (e.g. Riley *et al.* 2006). The presence of 'Jersey barriers' (Lesbarreres & Fahrig 2012) or fences for livestock exclusion (Dodd *et al.* 2004; Grilo *et al.* 2009; Holderegger & Di Giulio 2010; Kuehn *et al.* 2007), as well as also road avoidance due to physical or chemical alteration of the local environment may prevent animals from crossing LTIs and possibly lead them to increase their movement along road verges (Balkenhol & Waits 2009; Trombulak & Frissell 2000; McGregor *et al.* 2008). However, barrier effects are often alleviated by bridges and underpasses that can be used by dispersing animals. Furthermore, specific road crossing structures coupled with guide fences (McGregor *et al.* 2008; Olsson *et al.* 2008; Woltz *et al.* 2008; Grilo *et al.* 2009) can be built to mitigate barrier effects. However, because of important distances between these crossing structures and the specificity of locomotive behavior including orientation capabilities (Joly & Miaud 1993), barrier effects could remain strong, particularly for small

ground-dwelling animals. As mortality due to collisions and road avoidance both depend on traffic volume and infrastructure width, multi-lane motorways might be considered stronger barriers to dispersal than high-speed railways. However, rails may also constitute a physical barrier for small ground-dwelling animals, because, in standard situations, tracks are placed directly against the ballast, leaving no space for animals to easily move underneath, and while in addition, the salient head of tracks (when considering the cross sectional shape of rail) may prevent animals to climb over them.

At a larger scale, modifications of the surrounding landscape configuration consecutive to LTI construction could also impact wildlife dispersal patterns. For instance, LTIs are responsible for the alteration of surface-water habitats, through flow re-routing and wetlands destruction (Trombulak & Frissell 2000), impacting both aquatic and semi-aquatic species. In the particular case of motorways, road connections to urban areas improve rural-urban access and increase farmland value (Drescher *et al.* 2001) therefore often leading to the regrouping of cultivated farmlands, to the expense of suitable landscape features such as wooded patches, extensively pastured meadows and aquatic or hedgerows networks. Conversely, LTIs may also contribute to the creation of large-scale, non-fragmented linear wetlands such as roadside ditches and grasslands corridors such as roadside verges, favouring movement in areas of low-quality or highly fragmented habitats (Tikka *et al.* 2001; Brisson *et al.* 2010). All these potential effects depend on complex interactions between species-, infrastructures- and landscape-specific characteristics, in interaction with public policies and economic context (Woltz *et al.* 2008; Balkenhol & Waits 2009; Grilo *et al.* 2009; Kerth & Melber 2009).

Molecular genetics have been proved to be particularly efficient at detecting road impacts on population structures, even over relatively small temporal scales (Balkenhol & Waits 2009). In close collaboration with local road planners, as recommended in Lesbarreres & Fahrig (2012), we carried out a broad-scale analysis of the potential impacts of two 30 to 40 year-old LTIs on spatial genetic patterns in the alpine newt *Ichthyosaura alpestris*, using 14 highly polymorphic microsatellite loci. This amphibian is a widespread species in central Europe, and strongly depends on permanent fish-free ponds or smaller aquatic sites such as ruts, surrounded by suitable terrestrial habitat such as woods and semi-natural grasslands (Joly *et al.* 2001). It benefits from extensive livestock farming, which allows the preservation of many natural or man-made ponds in

pastures. *I. alpestris* shows highly nomadic behaviour (Perret *et al.* 2003), leading to strong gene flow and weak genetic differentiation among subpopulations (Emaresi *et al.* 2011; Prunier *et al. submitted-a*). This species thus constitutes a pertinent biological model for the study of barrier effects, since individuals with high dispersal abilities may encounter roads at higher rates than less mobile species, and may be more rapidly impacted by anthropogenic barriers (Carr & Fahrig 2001; Cushman 2006). Besides, many landscape genetic studies assessing the impact of roads on amphibians have detected a decrease in genetic diversity and an increase in genetic differentiation in the vicinity of LTIs (Lesbarreres *et al.* 2006; Holderegger & Di Giulio 2010), probably resulting in population fragmentation. We thus expect the two LTIs under study to significantly alter gene flow between populations located on either side. To test this hypothesis, we used a regular individual-based sampling scheme to optimize the number of sampled sites across the study area. This is a powerful alternative to the conventional population-based sampling scheme, particularly efficient at detecting barriers to gene flow when designed as a targeted sampling scheme, that is, when sampled sites are localized in the direct vicinity of a putative barrier (Prunier *et al. submitted-a*). This study also offered the opportunity to compare the efficiency of various statistical tools to detect the impacts of putative barriers on gene flow in an individual-based framework.

Materials and methods

Study area and biological model

The study was carried out in eastern central France (Bourgogne), over an area of approximately 60 x 55 km (Fig. 1a), with elevation ranging from 138 to 857 m (mean elevation: 399 m). The study area comprised several water catchments and three distinct landscape entities (Fig. 1b): (1) In the South-Western quarter of the study area, *Le Morvan*, a granitic massif comprising low elevation fragmented wooded patches and extensively pastured meadows for livestock farming, with a dense hydrographic network (minimum elevation: 400 m). (2) In the east, *L'Auxois,* a cultivated limestone plateau in the continuity of *Le Morvan*, interrupted by two parallel geologic depressions (minimum

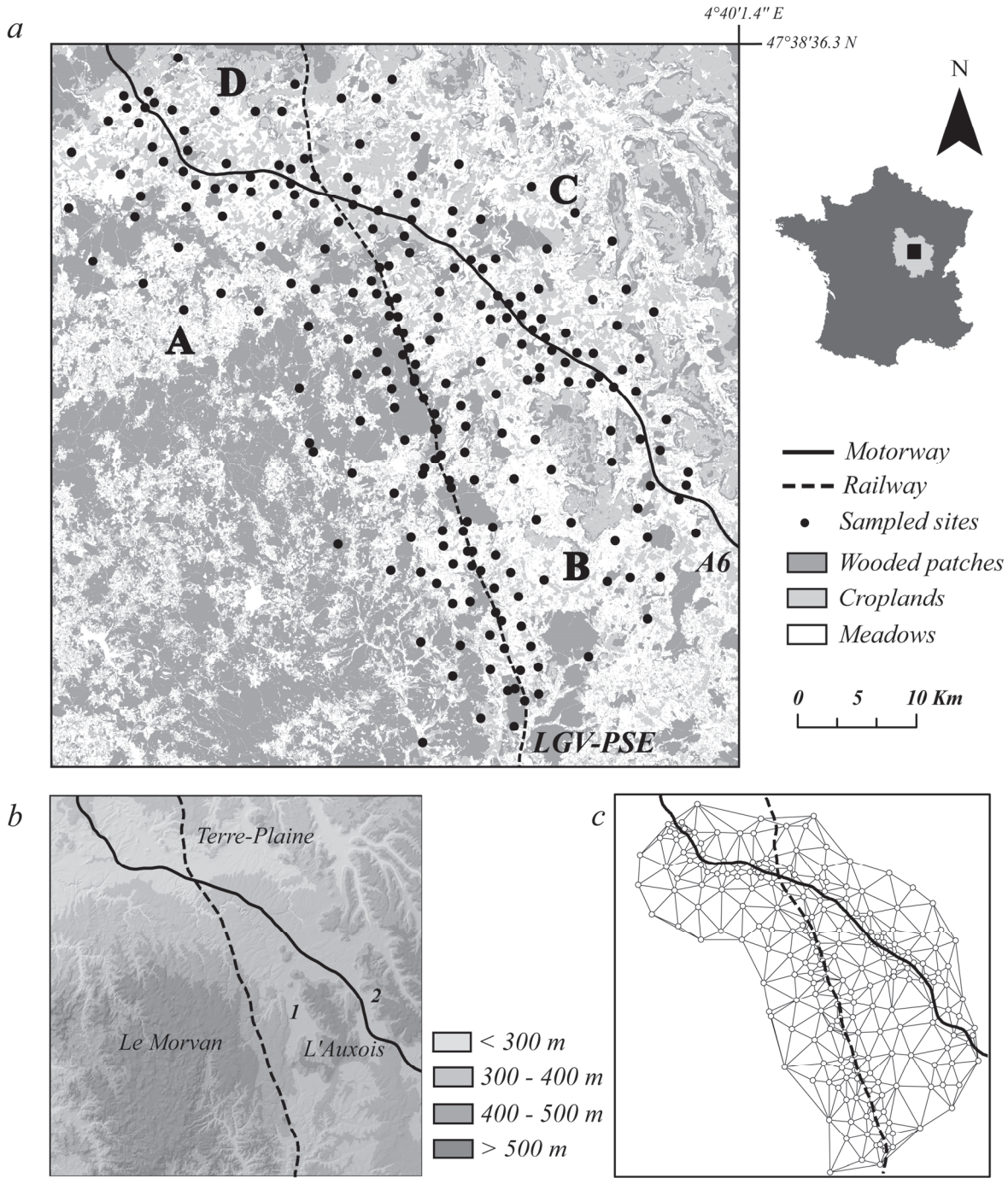
elevation: 300m). (3) In the north, *Terre-Plaine*, a vast low-elevation marly plain (minimum elevation: 187m). Both *Terre-Plaine* and geologic depressions in *L'Auxois* are covered with cultivated fields (mainly cereal crops), extensively pastured meadows and fragmented wooded patches.

The study area is crossed by two large transport infrastructures (LTIs) localized along the Paris-Lyon axis: (1) the *A6*, a 40 year-old dual 2-lane motorway in use since 1969, crosses *Terre-Plaine* and the eastern geologic depression in *L'Auxois*; (2) the *LGV*-PSE, a 30 year-old high-speed railway in use since 1981, runs along *Le Morvan* eastern foothills.

Both LTIs are approximately 25 m wide (50 m wide when considering the fenced area) and are regularly crossed by country roads, tracks and natural streams (approximately one transverse crossing every kilometer). They intersect in *Terre-Plaine* and delimit four distinct sectors (*A*, *B*, *C* and *D*, Fig. 1a). Except for the two LTIs, artificial areas were mainly farms, villages and a few small towns, connected by roads with low to medium traffic.

Figure 1. (***a***) Localisation of sampled ponds, wooded patches and croplands in the vicinity of the two large transport infrastructures (LTI). Intersection of LTIs delimits four distinct sectors (A, B, C and D). (***b***) Relief and main landscape entities in the study area. Western and eastern geologic depressions in *L'Auxois* are respectiviely numbered 1 and 2. (***c***) Neighbour joining tree between sampled sites following a Delaunay triangulation

→


a
4°40'1.4" E
47°38'36.3 N
N
D
C
A
B
A6
LGV-PSE
Motorway
Railway
Sampled sites
Wooded patches
Croplands
Meadows
0 5 10 Km
b
Terre-Plaine
1
2
Le Morvan
L'Auxois
< 300 m
300 - 400 m
400 - 500 m
> 500 m
c

Genetic sampling and genotyping

Fieldwork was conducted during the 2010 and 2011 breeding seasons, and 339 aquatic sites (man-made or natural ponds, flooded ruts, swamps, etc.) were prospected with a dip net for thirty to sixty minutes (depending on site size and configuration) or until a male and a female alpine newt were captured. Alpine newts were found in 225 sites (66% of prospected sites). The median distance between neighbouring sites (computed using a Delaunay triangulation, Fig. 1c) was 2.88 km, ranging from 184 m to 10.79 km. Non-destructive genetic samples using buccal swabs were taken from each captured individual, (Broquet *et al.* 2007). Samples were stored at ambient temperature in an air-tight container with silica gel providing total desiccation. DNA extraction, PCR amplifications and genotyping were performed as described in Prunier *et al.* (2012), using 14 microsatellite loci: *CopTa1*, *CopTa2*, *CopTa3*, *CopTa4*, *CopTa7*, *CopTa8*, *CopTa9*, *CopTa10*, *CopTa11*, *CopTa12*, *CopTa13* and *CopTa14* (Prunier *et al.* 2012), *Ta1Ca1* and *Ta1Caga4* (Garner *et al.* 2003). Significant gametic disequilibrium was detected between *CopTa7*, *CopTa9* and *CopTa10* in a previous study realized at the population-level (Prunier *et al.* 2012). Except for methods that are not contingent on particular genetic models (e.g. Jombart *et al.* 2008), *CopTa7* and *CopTa9* were thus discarded from analyses. Neither null alleles nor linkage disequilibrium were detected with this new combination of 12 markers (Prunier *et al.* 2012). To assess the reliability of our genetic data, we estimated the mean error rate per locus e_l (Pompanon *et al.* 2005) by blind replication of 45 out of 1081 samples (4.2 %) collected from 2009 to 2011 as part of a general research program on *I. alpestris*.

Sex-biased dispersal has been reported by several authors in this species (Joly & Grolet 1996; Perret *et al.* 2003) and transportation infrastructures may thus show distinct impacts on males and females. Since the use of nuclear markers does not prevent the detection of different spatial patterns of relatedness between sexes (Goudet *et al.* 2002), all the following analyses were performed respectively using males (dataset *M*) and females (dataset *F*), except when specified.

Isolation-by-distance IBD

As neighbour mating, leading to patterns of isolation-by-distance *IBD*, and landscape fragmentation, such as isolation-by-barrier *IBB*, may occur simultaneously, it is thus of crucial importance to test for *IBD* prior to any other analysis (Schwartz & McKelvey 2009). We performed spatial autocorrelation analyses through non-directional Mantel correlograms (Smouse & Peakall 1999; Borcard & Legendre 2012) to determine the scale *S* at which *IBD* patterns occur in each dataset (Epperson 2003b). For this purpose, we computed inter-individual pairwise genetic distances using the Bray-Curtis percentage dissimilarity measure (Legendre & Legendre 1998; Cushman *et al.* 2006). Euclidean distance classes were defined every 3000 m (up to 60 km), resulting in twenty binary matrices representing the membership of individuals to the distance class being tested (with 1 for pairs of individuals belonging to the same distance class and 0 otherwise). Each binary matrix was compared to the genetic distance matrix using a simple Mantel test with 1000 permutations. We then plotted Pearson correlation values over distance classes, with a 95% confidence interval determined by bootstrap resampling (1000 iterations). All Mantel tests (used to compute Mantel correlograms or to test for barrier effects) were performed using the MATLAB software coding environment (Mathworks, Inc.). All continuous variables (genetic and Euclidean distances) were log-transformed following the *D=ln(d+1)* formula and standardized to meet linearity assumptions.

Detection of global genetic structures

We performed a spatial analysis of shared alleles (SAShA), a method that has been developed in the MATLAB software coding environment (Mathworks, Inc.) for the detection of subtle geographic subdivisions, notably in high gene-flow species (Kelly *et al.* 2010). This is an allele-based approach, which avoids several limitations linked to population-level genetic analysis such as the a priori delineation of population boundaries, and thus allows a relevant analysis of individual-based genetic datasets. SAShA was used in each dataset, to calculate both the observed (*OM*) and the expected (*EM*) mean Euclidean distance between co-occurrences of an allele under the assumption of panmixia. The significance of the difference *Dg* between *EM* and *OM* was then determined by 1000 permutations. A non-significant *Dg* means that alleles are uniformly

distributed through space, whereas a significant positive *Dg* reveals genetic structure due to restricted gene flow among individuals.

We then used TESS 2.3.1 (Chen *et al.* 2007), a spatial Bayesian clustering method, to identify clusters of individuals. This program computes probabilities that each individual genotype originates from one of *K* panmictic populations. *I. alpestris* is a highly nomadic species, which respectively showed no and little genetic differentiation despite distances of up to 26 km in Emaresi *et al.* (2011) and 160 km in Pabijan & Babik (2006). Accounting for both the putative barrier effect of LTIs segregating the study area in four distinct sectors (Fig. 1a) and the extent of our study area (three times larger than in Emaresi *et al.* 2011), we did not expect to detect more than ten distinct genetic clusters (one cluster in sector *D* and three clusters in each sector *A*, *B* or *C*). Because of such a weak genetic differentiation between populations, genetic data were analyzed using correlated allele frequencies. We used the admixture CAR model starting from a neighbour joining tree based on a Delaunay triangulation (Fig. 1c). To estimate the true number of genetic clusters *K*, we made five runs per *K* for *K* = 2 to 10, with 100000 sweeps and a burn-in period of 10000 sweeps. We then identified which values of *K* produced the highest likelihood runs (lowest values of DIC), and made 100 runs for K_{max}, with the same parameters. The 20 best results (lowest values of DIC) were averaged with the computer program CLUMPP. Other parameters were set to default values. To visualize the results, we used the kriging function provided in the *R*-package 'gstat' (Pebesma 2004; R Development Core Team 2011) to represent the estimated membership of each individual on a grid with a resolution of 500m.

Finally, we performed a spatial principal component analysis (sPCA; Jombart *et al.* 2008), a spatially-explicit multivariate method using individual genotypes to investigate the spatial patterns of genetic variability. This method, seeking principal components that optimize the variance of individual allelic frequencies while taking spatial autocorrelation of data into account, does not require Hardy-Weinberg or linkage equilibrium: sPCA analyses were thus performed using all 14 loci, including *CopTa7* and *CopTa9*. This method provides maps of individual sPCA scores, allowing a visual assessment of the spatial genetic structures. It disentangles global structures, i.e. strong genetic similarity or positive autocorrelation between neighbours, from local ones, i.e. strong genetic differences or negative autocorrelation between neighbours. For each dataset, we used a

distance-based neighbourhood network with a distance threshold *S* consistent with the previously performed spatial correlograms. A global and a local Monte Carlo test were carried out with 1000 permutations to evaluate the significance of detected patterns (Jombart *et al.* 2008).

Detection of local genetic boundaries

While previous analyses look for homogeneous spatial areas, edge detection methods try to identify areas of abrupt genetic discontinuities (Guillot *et al.* 2009). The two LTIs were expected to act as barriers to gene flow, and the spatial individual-based sampling scheme was specifically designed to test this hypothesis. We thus tried to detect such a signal with *a priori* knowledge on the location of putative barriers, using spatial subsets of genotypes selected according to a targeted sampling scheme, that is subsets of genotypes localised in the direct vicinity of putative barriers (Anderson *et al.* 2010). This approach enables a significant improvement of statistical power in barrier detection analyses (Prunier *et al. submitted-a*), while allowing to test for the spatial scale at which gene flow is affected by the barrier (Epps *et al.* 2005). For each dataset, we computed five pairwise genetic distance matrices (using the Bray-Curtis percentage dissimilarity measure) from genotypes localised at maximal distances ranging from 2000 to 10000 m from each putative barrier, every 2000 m. Using partial Mantel tests with 1000 permutations, each pairwise genetic distance matrix was compared to the corresponding effective distance matrix after controlling for the effect of the Euclidian distance matrix (*IBB* detection), and conversely to the corresponding Euclidian distance matrix after controlling for the effect of the effective distance matrix (complementary *IBD* detection). Effective distance matrices were coded with 0 when two individuals were located on the same side of a barrier, or 1 when they were separated by a barrier (Epps *et al.* 2005; Prunier *et al. submitted-a*).

We also evaluated the presence of sharp genetic boundaries with no *a priori* knowledge on the location of putative barriers. Monmonier's Maximum Difference Algorithm (Monmonier 1973) implemented in AIS (Miller 2005) identifies genetic boundaries between pairs of individuals (Dupanloup *et al.* 2002; Manni *et al.* 2004; Kuehn *et al.* 2007), along a connectivity network based on a Delaunay triangulation. We

performed these analyses using residual genetic distances derived from the linear regression of all pairwise genetic distances on Euclidean distances (Manni *et al.* 2004; Miller 2005). For each sex, the number of barriers to detect was set to 4 (the number of distinct LTI sections; Fig. 1). We also used the Wombling method implemented in the R package WOMBSOFT (Cercueil *et al.* 2007; Crida & Manel 2007; Safner *et al.* 2011b), which detects areas of rapid change in allelic frequencies through the estimation of a systemic function (i.e. the averaged gradient in allelic frequencies). A binomial test then allows testing for the significance of the identified boundary elements. For each dataset *M* and *F*, we used a grid size of 3000 m (based on the median distance between neighbouring sampled points) and a bandwidth parameter of 4000 m. The binomial threshold and the statistical significance of the binomial test were set to 0.3 and 0.05 respectively.

Results

Genetic data

A total of 417 individuals were sampled in the study area. Both a male and a female were captured in 192 ponds, from a total of 225 ponds where *I. alpestris* was detected. 206 males (dataset *M*) and 211 females (dataset *F*) were sampled, resulting in a mean number of genotypes per site of 1.85. The genotyping error rate, estimated by blind replications as the mean error rate per locus e_l, was less than 2.4 % in *CopTa1*, *CopTa4*, *CopTa14* and *Ta1Caga4*, and 0% for the 10 other loci.

Isolation-by-distance IBD

In both datasets, Mantel correlograms showed significant genetic relatedness between pairwise individuals for the first 12 km (first 4 bins), and either no or negative autocorrelation as distance increased (Fig. 2). These results indicated the existence of a significant *IBD* process over the study area; the spatial scale of the genetic autocorrelation (*S* = 12 km) was similar in males and females.

Using partial Mantel tests to control for the presence of putative barriers to gene flow, a significant *IBD* was systematically detected whatever the dataset, the LTI or the scale of the targeted subsample ($p \leq 0.05$), except for males located at less than 2000 and 6000 m from the railway (Table 1), probably because of the low number of genotypes. Nevertheless, the detected *IBD* patterns were consistent with autocorrelation analyses.

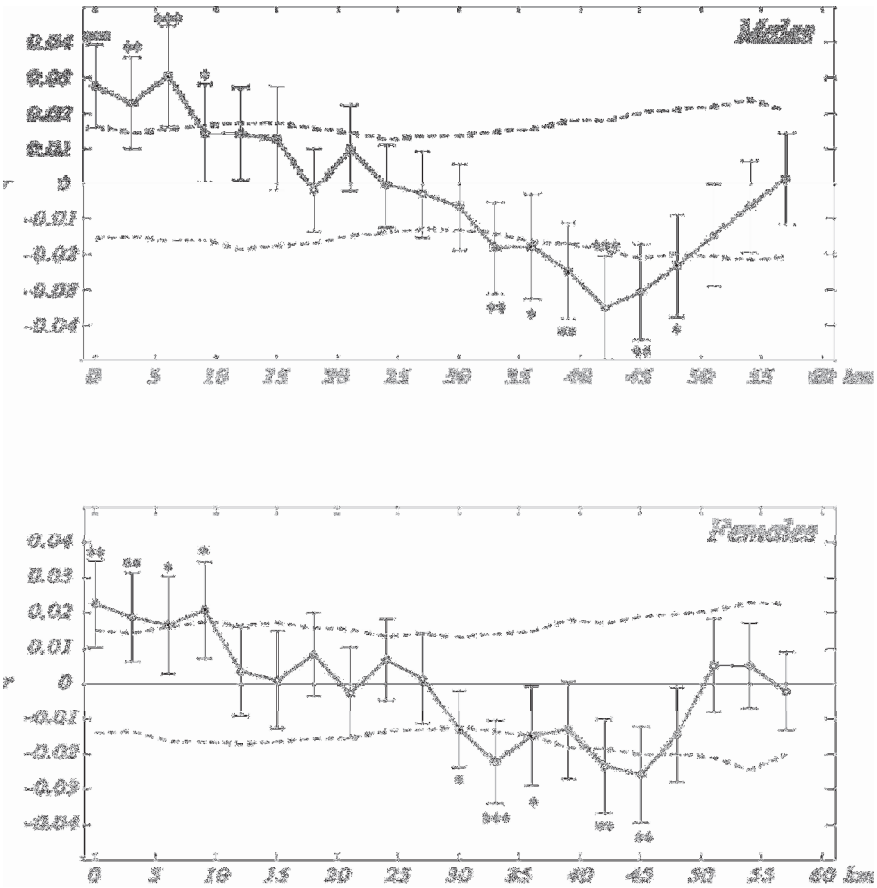


Figure 2. Mantel correlograms showing the relationships between inter-individual genetic distances and Euclidean distance classes (defined every 3000 m) in males (*M*) and in females (*F*). *r*: standard Mantel correlation with 1000 permutations. Error bars bound the 95% confidence interval about *r* as determined by bootstrap resampling. Upper and lower confidence limits (dotted line) bound the 95% confidence interval about the null hypothesis of no spatial structure as determined by permutation.
∗: *p*-value < 0.05; ∗∗: *p*-value < 0.01; ∗∗∗: *p*-value < 0.001. (see text for details).

Detection of global genetic structures

The use of SAShA highlighted the lack of genetic structure in all datasets: the spatial arrangement of alleles in *I. alpestris* was not statistically different from the expectation under panmixia (males: *OM* = 22840 m, *EM* = 22941 m, *NS*; females: *OM* = 23077 m, *EM* = 23057 m, *NS*), indicating that there was no limitation to the spread of alleles over space.

When the analyses were performed in males or in females only, TESS systematically failed to detect any genetic structure (data not shown), which may be explained by the low number of genotypes in each dataset. We thus used the whole dataset, combining males and females. The best estimate of the number of clusters using the averaged DIC criterion over the five runs performed per *K* was 3 (Fig. 3a). However, runs performed for K_{max} = 3 showed high variability in DIC values. Furthermore, for K_{max} > 2, the posterior estimates of cluster membership for each individual systematically displayed only one or two clusters, the additional clusters containing negligible proportions. We thus performed 100 runs with K_{max} = 2; this number of inferred clusters allowed the use of a simple kriging interpolation function. The 20 best runs averaged using CLUMPP led to a weak discrimination of individuals in two distinct clusters (Fig. 3b): individuals with an estimated membership in cluster 1 higher than 40% were located in *Terre-Plaine*, while those with a lower estimated membership were located in *Le Morvan* and *L'Auxois* (Fig. 3c). High levels of admixture were encountered at the interface of these two clusters, confirming the existence of an *IBD* pattern. Here again, the detected boundary was not related to the presence of LTIs.

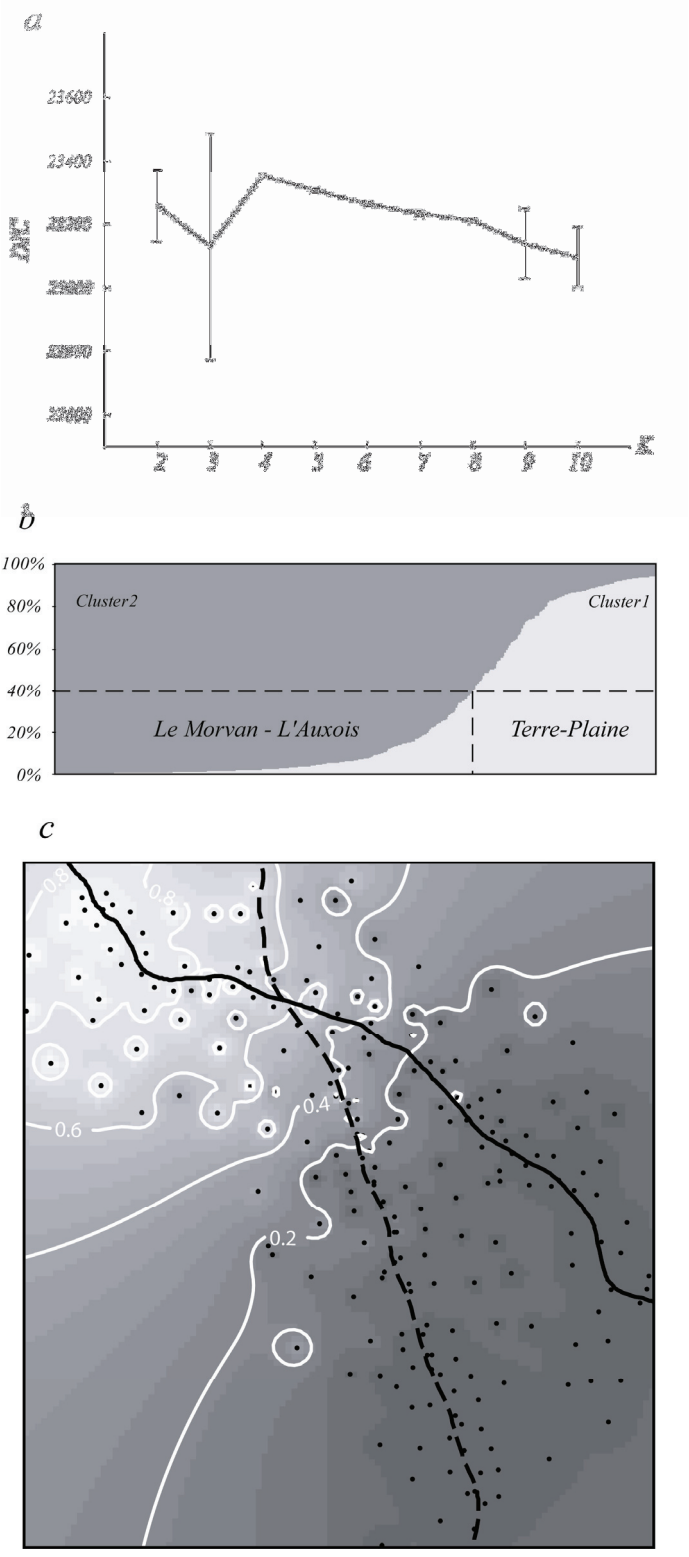
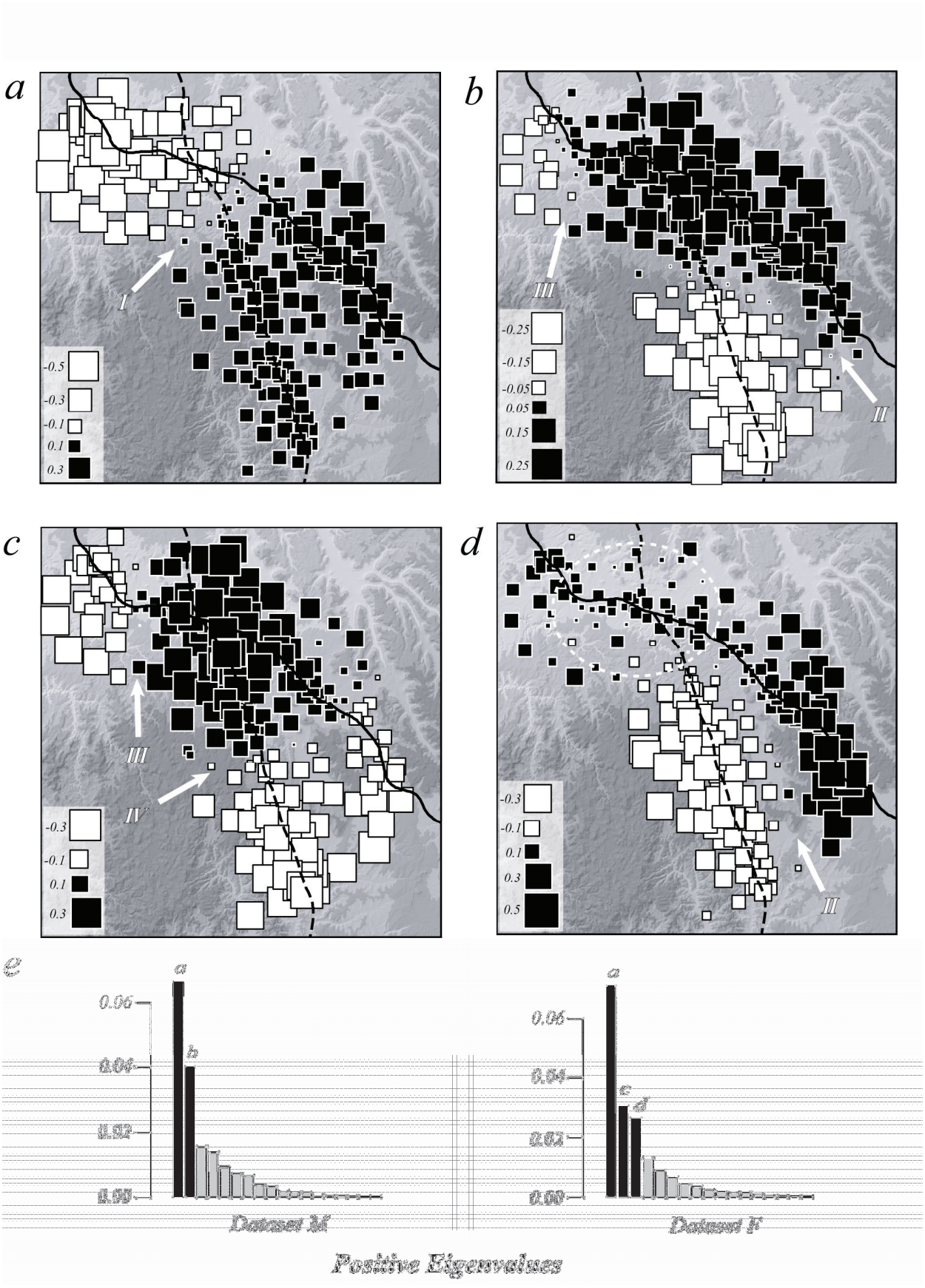


Figure 3. Spatial genetic structures inferred from TESS. The figure provides estimates of the true number *K* of clusters using the DIC criterion (a), the estimated membership (*y*-axis) of each individual (*x*-axis) in clusters 1 and 2 (b) and the spatial distribution of clusters 1 and 2 in the study area (c). Black dots stand for the sampling points.

All global Monte-Carlo tests performed in sPCA were significant (dataset *M*, max(*t*) = 0.0109, *p* = 0.001; dataset *F*, max(*t*) = 0.011, *p* = 0.001), indicating the presence of significant global genetic structures in all datasets. On the contrary, local Monte-Carlo tests did not detect any significant local structure (dataset *M*, max(*t*) = 0.007, *NS*; dataset *F*, max(*t*) = 0.006, *NS*). The analyses revealed four kinds of global genetic patterns (Fig. 4e). In both datasets, scores of individuals along the first sPCA axis distinguished *Terre-Plaine* from the rest of the study area, as did TESS (Fig. 4a, Boundary *I*). A second pattern was revealed by second sPCA scores in datasets *M* (Fig. 4b): this structure clearly segregated individuals located in *Le Morvan* from the rest of the study area (Boundary *II*); a third group, identified in the west of *Terre-Plaine* (Boundary *III*) corresponded to the westernmost water catchment in *Terre-Plaine*, although further analyses, performed over a more extended study area, may be required to confirm this pattern. Scores of females along the second sPCA axis (dataset *F*) revealed a slightly different pattern in the south, with higher genetic similarity among individuals located in southern hilly landscapes from *Le Morvan* and *L'Auxois* (Fig. 4c, southern white squares, Boundary *IV*). Finally, third sPCA scores in females (Fig. 4d) displayed a pattern highly similar to the one obtained from the second sPCA scores in dataset *M* (Boundary *II*). However, females from *Terre-Plaine* located next to the motorway showed higher scores than females located more distantly (area delimited with white dashes), indicating higher genetic similarity among individuals in the direct vicinity of the motorway. In all cases, these global structures did not show any sharp boundaries between groups, but rather progressive changes from one cluster to the other, as suggested by low individual scores in transition zones. LTIs never separated inferred groups.

Figure 4. Analyses of alpine newt data using sPCA. Large white squares stand for highly negative scores whereas large black squares stand for highly positive scores. Small squares stand for lows sPCA scores. (a) Map of the first global sPCA scores in dataset *M;* analyses led to highly similar patterns in dataset *F* (data not shown). (b) Map of the second global sPCA scores in dataset *M;* (c) Map of the second global sPCA scores in dataset *F*. (d) Map of the third global sPCA scores in dataset *F*; white dashes delimit an area in which females located in the direct vicinity of the motorway show high genetic similarity. (e) Screeplots of sPCA eigenvalues in datasets *M* and *F*; retained structures are filled in black; letters in italics refer to maps a to d.

→


a
I
-0.5
-0.3
-0.1
0.1
0.3
b
III
II
-0.25
-0.15
-0.05
0.05
0.15
0.25
c
III
IV
-0.3
-0.1
0.1
0.3
d
III
-0.3
-0.1
0.1
0.3
0.5
e
a
b
0.06
0.04
0.02
0.00
Dataset M
a
c
d
0.06
0.04
0.02
0.00
Dataset F
Positive Eigenvalues

Detection of local genetic boundaries

All partial Mantel tests failed to detect any *IBB* pattern, whatever the LTI, the dataset or the scale of the targeted subsample (Table 1).

Barriers inferred using the Monmonier's algorithm were mainly located close to one individual or only a few neighbours (all barriers in males and the fourth one in females, Fig. 5). These barriers actually indicated individuals distinct from their direct neighbours and were not interpretable as boundaries. However, the three first barriers detected in females (Fig. 5b) were connected and thus consistent with the presence of a landscape barrier to gene flow. They segregated south-eastern females from the rest of the study area, as did the combination of boundaries *II* and *IV* detected with second and third sPCA scores in females (Fig. 4c-d).

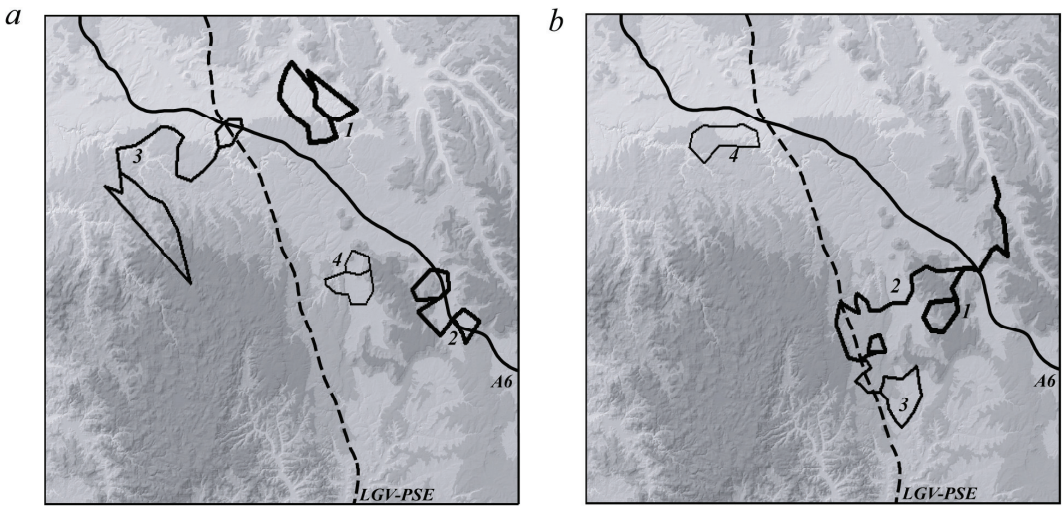


Figure 5. Four genetic barriers detected by Monmonier's algorithm implemented in AIS (Miller 2005), numbered from 1 (bold black lines) to 4 (thin black line) and projected over Figure 1b; (a) in males (dataset *M*); (b) in females (dataset *F*).

Apart from artifactual gradients detected on the edge of the study area, WOMBSOFT only detected sharp genetic boundaries along the eastern geologic depression from *L'Auxois* and along the eastern foothills of *Le Morvan* (Fig. 6). These two zones coincided with local portions of LTIs, but also with areas of rapid change in landscape configuration especially elevation and proportion of wooded patches between *Le Morvan* and the western geologic depression from *L'Auxois*, as well as elevation and proportion of croplands between the central limestone plateau and the eastern geologic depression in *L'Auxois*.

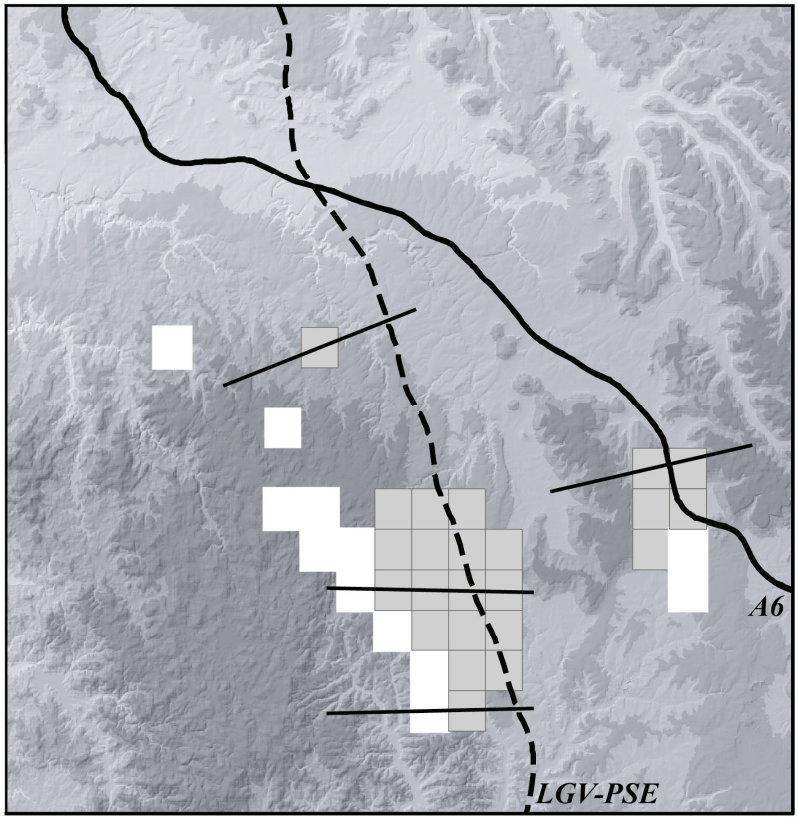


Figure 6. Identification of areas of local variation of allele frequencies (light grey squares) using Wombling in dataset *M*. Black lines represent gradient directions. White areas are artefactual boundaries detected on the edge of the study area. Boundaries were highly similar in females (dataset *F*; data not shown).

Table 1. Isolation-by-distance (*IBD*) and Isolation-by-barrier (*IBB*) detection through partial Mantel tests using all genotypes or targeted subsets of genotypes. Dataset: Males *M*; Females *F*. Barrier: Motorway (*A6*) or railroad (*LGV-PSE*). X: extent of the targeted subset on each side of the barrier. *n*: number of genotypes in the subset. *r*: Mantel correlation. *DG.B*: partial Mantel test between the pairwise genetic distance matrix *G* and the Euclidean distance matrix *D*, with the barrier matrix *B* partialed out (*IBD* detection).
BG.D: partial Mantel test between the pairwise genetic distance matrix *G* and the barrier matrix *B*, with the Euclidean distance matrix *D* partialed out (*IBB* detection).

Dataset	Barrier	X	*n*	*IBD* (*DG.B*)		*IBB* (*BG.D*)	
				r	*p*-value *	*r*	*p*-value
M	*A6*	2000	58	0.0647	0.013	0.0035	NS
M	*A6*	4000	75	0.0949	0.001	-0.0237	NS
M	*A6*	6000	92	0.0988	< 0.001	-0.0077	NS
M	*A6*	8000	112	0.111	< 0.001	-0.0157	NS
M	*A6*	10000	129	0.1103	< 0.001	-0.0234	NS
M	*LGV-PSE*	2000	53	0.0341	NS	0.0274	NS
M	*LGV-PSE*	4000	84	0.0414	0.044	0.0361	NS
M	*LGV-PSE*	6000	107	0.0276	NS	0.1029	NS
M	*LGV-PSE*	8000	124	0.0427	0.017	0.0154	NS
M	*LGV-PSE*	10000	139	0.0416	0.022	0.0252	NS
F	*A6*	2000	75	0.1494	< 0.001	0.0019	NS
F	*A6*	4000	84	0.1186	< 0.001	-0.0203	NS
F	*A6*	6000	102	0.1096	< 0.001	-0.0243	NS
F	*A6*	8000	121	0.1049	< 0.001	-0.0248	NS
F	*A6*	10000	137	0.0918	< 0.001	-0.0272	NS
F	*LGV-PSE*	2000	57	0.081	0.01	-0.0298	NS
F	*LGV-PSE*	4000	87	0.0484	0.028	-0.0041	NS
F	*LGV-PSE*	6000	110	0.0395	0.032	0.0056	NS
F	*LGV-PSE*	8000	124	0.0407	0.022	0.0021	NS

* NS : p-value > 0.05

Discussion

A significant *IBD* pattern was detected in both sexes, with a higher genetic relatedness among individuals less than 12 km apart. The detection of a similar long-distance *IBD* pattern in both males and females is explained by the fact that any species is necessarily affected by *IBD* when considered at a large spatial extent (Guillot *et al.* 2009). In Mantel correlograms, each distance class (corresponding to a set of pairwise genetic distances) is confronted to all other distance classes, including the largest ones: as a result, the larger the study area, the higher the relative genetic similarity among low to medium distant individuals when compared to highly distant ones, thus leading to positive autocorrelation at larger scales than expected in a less extended study area. This also weakens the differences that may be observed at a finer scale between sexes. The detection of this long-distance *IBD* pattern in *I. alpestris* is consistent with previous studies (Perret *et al.* 2003; Pabijan & Babik 2006; Emaresi *et al.* 2011) and reinforced the expectation of a significant barrier effect of linear transport infrastructures (LTI) on gene flow in this species (Carr & Fahrig 2001; Cushman 2006). Highly nomadic individuals may actually encounter roads at higher rates than less mobile species, and may thus be more rapidly affected by anthropogenic barriers. However, inferred spatial genetic patterns did not support this hypothesis. Four main genetic boundaries were detected, which coincided with transition zones between major landscape entities in the study area (Fig. 1b and Fig. 4) rather than with the motorway or the railway. Boundaries *I* and *II* respectively distinguished *Terre-Plaine* and *Le Morvan* as two distinct homogeneous clusters. In females, boundary *IV* assigned individuals located in southern higher elevation areas from *Le Morvan* and *L'Auxois* to the same group. The observed genetic patterns may be explained by subtle modifications in landscape characteristics such as elevation (Boundaries *II* and *IV*), land use (Boundaries *I*, *II* and *IV*), proportions of unfragmented wooded patches (Boundary *II*), hydrologic networks (Boundary *III*), or most probably a combination of these factors.

The most informative analysis was the sPCA, which provided an intelligible overview of spatial genetic structure in the study area, while other methods only revealed pieces of the puzzle. The major genetic boundary (*I*), detected with the sPCA (Fig. 4a) as well as

with TESS (Fig. 3), was not a sharp but a progressive boundary. This first boundary explained most of the genetic variability in all datasets (Fig. 4e). Boundaries *II* and *IV* (Fig. 4b, c and d) were not detected by TESS but were identified as sharp boundaries by Monmonier's algorithm applied to females, and this pattern coincided with rapid changes in elevation and in land use between the three main landscape entities (Fig. 1a and b). A similar observation may be drawn from WOMBSOFT results (Fig. 6). In individual-based genetic studies, Safner *et al.* (2011b) recommended favouring Bayesian clustering methods to the expense of local edge detection methods, since the performance of the latter may be weakened by high within-individual genetic variability. Considering our own results, we argue that, taking into account this potential bias, their use may contribute to the evidence needed to correctly interpret observed genetic patterns. We also highly recommend the use of sPCA, combined with spatial autocorrelation analyses to determine the scale of the required neighbouring tree.

Although the alpine newt was suspected to be particularly impacted by landscape fragmentation due to the motorway and the railway, both LTIs were systematically embedded in inferred clusters and the SAShA method did not detect any limitation to the spread of alleles. Furthermore, partial Mantel tests did not support the hypothesis of a barrier effect due to LTIs, although the use of a targeted sampling scheme, that is the use subsets of genotypes located in the direct vicinity of LTIs so as to avoid large gaps relative to a putative barrier to gene flow, is known to be particularly efficient in barrier detection (Anderson *et al.* 2010; Prunier *et al. submitted-a*).

The absence of any detectable barrier effect of the two LTIs may be due to several factors. A major factor may be the time lag between the processes that caused the formation of a spatial genetic structure and the observed spatial genetic structure itself (Anderson *et al.* 2010). In our case, the two LTIs were more than 30 years-old, which may be sufficient for a significant *IBB* to appear (Lesbarreres *et al.* 2006; Landguth *et al.* 2010). Although the alpine newt is a long-lived species (Joly & Grolet 1996; Wagner *et al.* 2011), its generation time at low elevations is rather low (age at maturity ranging from 1 to 3 years at altitudes lower than 500 m; Miaud *et al.* 2000), and individuals show high dispersal rates (Emaresi *et al.* 2011; Prunier *et al. submitted-a*), which might counteract longevity (Landguth *et al.* 2010).

A second important factor explaining why LTIs do not act as barrier in our study is large effective population sizes (Gauffre *et al.* 2008): populations separated by a barrier may still be functionally connected with populations in their hinterland, rendering genetic barrier effects difficult to detect. This factor may not be discarded in the present study, as SAShA results indicated that all alleles were randomly distributed over the study area, suggesting a high genetic homogeneity on both sides of LTIs.

A third, often overlooked factor may also explain the absence of a significant *IBB*: LTIs may not be impermeable barriers to gene flow, and may even be totally permeable to dispersal. If barriers were only partially permeable to gene flow, methods such as sPCA designed to reveal cryptic genetic patterns (Jombart *et al.* 2008) may have detected a signal. If transportation infrastructures were totally permeable, no statistical approach would ever succeed in detecting any 'ghost' barrier. While our study cannot evaluate the importance of each factor in the case of the railway (field observations suggested that newts may be able to move underneath rail tracks), a clue concerning the case of the motorway may be found in figure 4d. The sPCA analysis displayed an informative pattern, with higher genetic similarity among females located in the direct vicinity of the motorway. This may suggest that this infrastructure is not only permeable to gene flow, but also that it may serve as a longitudinal dispersal corridor in a low-quality matrix such as *Terre-Plaine*, characterized by high proportions of fragmented wooded patches and cultivated crops (Fig. 1a; Meunier *et al.* 1999; Tikka *et al.* 2001). This pattern is consistent with sex-biased dispersal abilities observed in the alpine newt: females being less mobile than males (Joly & Grolet 1996; Perret *et al.* 2003), they may benefit from the existence of both evenly-spaced transverse passages (e.g. culverts, extended stream crossings or dirt roads; Lesbarreres *et al.* 2004; Corlatti *et al.* 2009; Lesbarreres & Fahrig 2012) and favorable roadside structures (e.g. ditches for rainwater drainage) to disperse through and along the motorway (as several exotic species; Trombulak & Frissell 2000; Brown *et al.* 2006; Jodoin *et al.* 2008), notably when it crosses highly resistant agricultural sectors (Joly *et al.* 2001; Janin *et al.* 2009). This counterintuitive genetic pattern was only detected with the third sPCA axis, suggesting that it may be the result of recent landscape changes when compared to the main genetic structure coinciding with major historical landscape entities. sPCA therefore emerges as a promising tool to detect cryptic genetic

patterns caused by temporal lags in genetic response (Anderson *et al.* 2010; Manel *et al.* 2010).

More generally, our study indicates that LTIs are not univocal conservation issues that necessarily lead to impeded dispersal of native species, if they are equipped with transverse and longitudinal structures allowing the movement of animals. Further investigations are thus required to define species-specific benchmarks in terms of permeability of road structures to movement (Mata *et al.* 2008; Woltz *et al.* 2008; Lesbarreres & Fahrig 2012), and to clarify the impact of land use conversion in the vicinity of various LTIs (or sections of LTIs).

REFERENCES

Anderson CD, Epperson BK, Fortin MJ*, et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology* **19**, 3565-3575.

Balkenhol N, Waits LP (2009) Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology* **18**, 4151-4164.

Borcard D, Legendre P (2012) Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology* **93**, 1473-1481.

Brisson J, de Blois S, Lavoie C (2010) Roadside as Invasion Pathway for Common Reed (Phragmites australis). *Invasive Plant Science and Management* **3**, 506-514.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics* **8**, 509-511.

Brown GP, Phillips BL, Webb JK, Shine R (2006) Toad on the road: Use of roads as dispersal corridors by cane toads (Bufo marinus) at an invasion front in tropical Australia. *Biological Conservation* **133**, 88-94.

Carr LW, Fahrig L (2001) Effect of road traffic on two amphibian species of differing vagility. *Conservation Biology* **15**, 1071-1078.

Cercueil A, Francois O, Manel S (2007) The Genetical Bandwidth Mapping: A spatial and graphical representation of population genetic structure based on the Wombling method. *Theoretical Population Biology* **71**, 332-341.

Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* **7**, 747-756.

Corlatti L, Hacklander K, Frey-Roos F (2009) Ability of Wildlife Overpasses to Provide Connectivity and Prevent Genetic Isolation. *Conservation Biology* **23**, 548-556.

Crida A, Manel S (2007) WOMBSOFT: an R package that implements the Wombling method to identify genetic boundary. *Molecular Ecology Notes* **7**, 588-591.

Croteau EK, Heist EJ, Nielsen CK (2010) Fine-scale population structure and sex-biased dispersal in bobcats (*Lynx rufus*) from southern Illinois. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **88**, 536-545.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation* **128**, 231-240.

Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology* **19**, 3592-3602.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* **168**, 486-499.

Dodd CK, Barichivich WJ, Smith LL (2004) Effectiveness of a barrier wall and culverts in reducing wildlife mortality on a heavily traveled highway in Florida. *Biological Conservation* **118**, 619-631.

Drescher K, Henderson J, McNamara K (2001) Farmland Prices Determinants

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571-2581.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conservation Genetics* **12**, 41-50.

Epperson BK (2003) *Geographical Genetics*, Princeton University Press.

Epps CW, Palsboll PJ, Wehausen JD*, et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters* **8**, 1029-1038.

Forman RTT (2000) Estimate of the area affected ecologically by the road system in the United States. *Conservation Biology* **14**, 31-35.

Garner TWJ, Schmidt BR, Hoeck P, Van Buskirk J (2003) Di- and tetranucleotide microsatellite markers for the Alpine newt (*Triturus alpestris*): characterization and cross-priming in five congeners. *Molecular Ecology Notes* **3**, 186-188.

Gauffre B, Estoup A, Bretagnolle V, Cosson JF (2008) Spatial genetic structure of a small rodent in a heterogeneous landscape. *Molecular Ecology* **17**, 4619-4629.

Grilo C, Bissonette JA, Santos-Reis M (2009) Spatial-temporal patterns in Mediterranean carnivore road casualties: Consequences for mitigation. *Biological Conservation* **142**, 301-313.

Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Molecular Ecology* **18**, 4734-4756.

Hagerty BE, Nussear KE, Esque TC, Tracy CR (2011) Making molehills out of mountains: landscape genetics of the Mojave desert tortoise. *Landscape Ecology* **26**, 267-280.

Holderegger R, Di Giulio M (2010) The genetic effects of roads: A review of empirical evidence. *Basic and Applied Ecology* **11**, 522-531.

Janin A, Lena JP, Ray N*, et al.* (2009) Assessing landscape connectivity with calibrated cost-distance modelling: predicting common toad distribution in a context of spreading agriculture. *Journal of Applied Ecology* **46**, 833-841.

Jodoin Y, Lavoie C, Villeneuve P*, et al.* (2008) Highways as corridors and habitats for the invasive common reed *Phragmites australis* in Quebec, Canada. *Journal of Applied Ecology* **45**, 459-466.

Joly P, Grolet O (1996) Colonization dynamics of new ponds, and the age structure of colonizing Alpine newts, *Triturus alpestris*. *Acta Oecologica* **17**, 599-608.

Joly P, Miaud C (1993) How does a newt find its pond - the role of chemical cues in migrating newts (*Triturus alpestris*). *Ethology Ecology & Evolution* **5**, 447-455.

Joly P, Miaud C, Lehmann A, Grolet O (2001) Habitat matrix effects on pond occupancy in newts. *Conservation Biology* **15**, 239-248.

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

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity* **101**, 423-436.

Kerth G, Melber M (2009) Species-specific barrier effects of a motorway on the habitat use of two threatened forest-living bat species. *Biological Conservation* **142**, 270-279.

Kuehn R, Hindenlang KE, Holzgang O*, et al.* (2007) Genetic effect of transportation infrastructure on roe deer Populations (Capreolus capreolus). *Journal of Heredity* **98**, 13-22.

Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**, 4179-4191.

Legendre L, Legendre P (1998) *Numerical Ecology* 2nd English ed. Elsevier., Amsterdam.

Lesbarreres D, Fahrig L (2012) Measures to reduce population fragmentation by roads: what has worked and how do we know? *Trends in ecology & evolution (Personal edition)*.

Lesbarreres D, Lode T, Merila J (2004) What type of amphibian tunnel could reduce road kills? *Oryx* **38**, 220-223.

Lesbarreres D, Primmer CR, Lode T, Merila J (2006) The effects of 20 years of highway presence on the genetic structure of Rana dalmatina populations. *Ecoscience* **13**, 531-538.

Manel S, Joost S, Epperson BK*, et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology* **19**, 3760-3772.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**, 173-190.

Mata C, Hervas I, Herranz J, Suarez F, Malo JE (2008) Are motorway wildlife passages worth building? Vertebrate use of road-crossing structures on a Spanish motorway. *Journal of Environmental Management* **88**, 407-415.

McGregor RL, Bender DJ, Fahrig L (2008) Do small mammals avoid roads because of the traffic? *Journal of Applied Ecology* **45**, 117-123.

Meunier FD, Verheyden C, Jouventin P (1999) Bird communities of highway verges: Influence of adjacent habitat and roadside management. *Acta Oecologica-International Journal of Ecology* **20**, 1-13.

Miaud C, Guyetant R, Faber H (2000) Age, size, and growth of the alpine newt, *Triturus alpestris* (Urodela : Salamandridae), at high altitude and a review of life-history trait variation throughout its range. *Herpetologica* **56**, 135-144.

Miller MP (2005) Alleles In Space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* **96**, 722-724.

Monmonier M (1973) Maximum-difference barriers: An alternative numerical regionalization method. *Geographical Analysis* **5**, 245-261.

Olsson MPO, Widen P, Larkin JL (2008) Effectiveness of a highway overpass to promote landscape connectivity and movement of moose and roe deer in Sweden. *Landscape and Urban Planning* **85**, 133-139.

Pabijan M, Babik W (2006) Genetic structure in northeastern populations of the Alpine newt (Triturus alpestris): evidence for post-Pleistocene differentiation. *Molecular Ecology* **15**, 2397-2407.

Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* **30**, 683-691.

Perret N, Pradel R, Miaud C, Grolet O, Joly P (2003) Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology* **72**, 567-575.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-859.

Prunier J, Kaufmann B, Fenet S*, et al.* (*submitted*) Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling scheme.

Prunier J, Kaufmann B, Grolet O*, et al.* (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*)

R Development Core Team (2011) R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Riley SPD, Pollinger JP, Sauvajot RM*, et al.* (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* **15**, 1733-1741.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011) Comparison of Bayesian Clustering and Edge Detection Methods for Inferring Boundaries in Landscape Genetics. *International Journal of Molecular Sciences* **12**, 865-889.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Tikka PM, Hogmander H, Koski PS (2001) Road and railway verges serve as dispersal corridors for grassland plants. *Landscape Ecology* **16**, 659-666.

Trombulak SC, Frissell CA (2000) Review of ecological effects of roads on terrestrial and aquatic communities. *Conservation Biology* **14**, 18-30.

Wagner A, Schabetsberger R, Sztatecsny M, Kaiser R (2011) Skeletochronology of phalanges underestimates the true age of long-lived Alpine newts (*Ichthyosaura alpestris*). *Herpetological Journal* **21**, 145-148.

Woltz HW, Gibbs JP, Ducey PK (2008) Road crossing structures for amphibians and reptiles: Informing design through behavioral analysis. *Biological Conservation* **141**, 2745-2750.

Acknowledgements

This work was supported by the *Agence Nationale de la Recherche et de la Technologie* (ANRT, agreement 2⁄2009), *Réseau Ferré de France* (RFF), *Autoroutes Paris-Rhin-Rhône* (APRR) and the ECOSPHERE company. It was conducted in accordance with French laws and with the approval of the *Préfectures de Saône-et-Loire*, *de Côte-d'Or* and *de l'Yonne*. We warmly thank T. Jombart for helpful comments about sPCA results. In addition, we thank C. De Fazio, G. Perez, F. Revollon, F.

Ribette, J. Mangin and O. Grolet for their precious help and support in the field and laboratory. We warmly thank Dr T. Vines and three anonymous reviewers for insightful suggestions.

Contribution by authors: BK, JP, JPL and PJ designed the study; JP collected empirical data; BK, JP and JPL analyzed output data; JP wrote the first draft of the manuscript and all authors contributed substantially to revisions; BK, FP, JPL and PJ ensured the work supervision.

CHAPITRE 4

Migration and dispersal patterns in a complex landscape inferred from occurrence and genotypic data in the alpine newt *Ichthyosaura alpestris*.

Avant-propos.

L'isolement par barrière aux flux de gènes (IBB ; Chapitre 3) peut être considéré comme un cas particuliers d'isolement par la résistance des habitats (*isolation-by-resistance*, IBR) : dans ce cas, l'hypothèse est que la barrière potentielle considérée est, au sein de la matrice, le seul élément paysager susceptible de perturber la dispersion des organismes. Or tous les éléments naturels et anthropiques composant un paysage peuvent présenter des degrés de perméabilité différents à la dispersion des organismes. Dans ce contexte, une amélioration de la couverture d'échantillonnage de la zone d'étude par une approche individu-centrée, assurant une meilleure prise en compte de l'hétérogénéité du paysage, peut être déterminante. Or cet aspect n'a pu être traité directement par simulations (Chapitre 1).

Ce dernier chapitre s'intéresse donc à une mise en application concrète d'un échantillonnage individu-centré dans le cadre de l'étude de la connectivité fonctionnelle des habitats : il porte sur l'influence de la composition et de la configuration d'un paysage complexe sur la dispersion du triton alpestre, par la combinaison de données génétiques et démographiques.

Cet article est préparation, en vue d'une publication dans *Biological conservation* :

Prunier, J., Kaufmann, B., Pompanon, F., Joly, P., Lena, J.P. Migration and dispersal patterns in a complex landscape inferred from occurrence and genotypic data in the alpine newt *Ichthyosaura alpestris*.

Migration and dispersal patterns in a complex landscape inferred from occurrence and genotypic data in the alpine newt *Ichthyosaura alpestris*.

Jérôme PRUNIER [a, b], Bernard KAUFMANN [b], François POMPANON [c], Pierre JOLY [b] and Jean-Paul LENA [b]

[a] *Ecosphère, 3bis rue des Remises, 94100 Saint Maur des Fossés, France*

[b] *Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.*

[c] *Université Joseph Fourier, Laboratoire d'Ecologie Alpine, CNRS, UMR 5553, 2233 Rue de la Piscine, 38041 Grenoble Cedex 9, France*

Emails:

jerome.prunier@gmail.com

bernard.kaufmann@univ-lyon1.fr

francois.pompanon@ujf-grenoble.fr

pierre.joly@univ-lyon1.fr

jean-paul.lena@univ-lyon1.fr

Corresponding author:

Jérôme Prunier

Université de Lyon; UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés; Université Lyon 1; ENTPE; CNRS; Villeurbanne, F-69622, France.

Tel: (+33) 4 72 43 29 53

Fax: (+33) 4 72 43 11 41

Email: jerome.prunier@gmail.com

Keywords: individual-based sampling scheme; principal component analysis; collinearity; landscape genetics; multimodel inference; agricultural practice; complementation

Abstract

Species conservation relies on the understanding of ecological requirements at all life-stages and on a fine comprehension of local landscape characteristics. Using occurrence and genotypic data gathered in the framework of an individual-based sampling scheme, this study aimed at identifying the influence of various landscape features on migration and dispersal patterns in the alpine newt *Ichthyosaura alpestris* in a complex landscape. Principal component analyses identified multicollinearity among habitat proportions, pointing out that crops were mainly located on plateaus while permanent meadows were rather located in sloping sectors such as dales and ridges. Integrating these landscape characteristics in information theoretic model selection, we found that although probability of occurrence decreased with increasing proportions of crops, dispersal mainly occurred within agricultural plateaus, sloping sectors acting as potential barriers to gene flow. Furthermore, genotypic data showed high rates of dispersal among neighbouring ponds, suggesting low breeding site fidelity, notably in males: we argue that agricultural practices such as crop rotation, that lead to spatiotemporal unpredictability in quality of breeding sites may explain these patterns. Our study emphasizes the need for amphibian conservation to be planned on a case-by-case basis, migration and dispersal patterns intimately depending on local topography and agricultural practices.

1. Introduction

The viability of local populations strongly depends on the permeability of the landscape matrix to individual movement between resource patches (Dunning et al. 1992; Taylor et al. 1993). However, in anthropogenic landscapes, habitat loss, land-use conversion and fragmentation are known to alter matrix permeability and constitute critical concerns for the conservation of most organisms (Ricketts 2001; Fahrig 2003; Fischer & Lindenmayer 2007). This is especially true for amphibians, known to suffer from dramatic decline worldwide (Stuart *et al.* 2004; Beebee & Griffiths 2005). Actually, pond-breeding amphibians have to undertake various migratory and dispersal movements between distinct aquatic and terrestrial resource patches that expose them to cross more or less harsh landscape matrix (Cushman 2006; Denoel & Ficetola 2008).

Migrations are obligatory within-population movements occurring between breeding aquatic sites and non-substitutable terrestrial resources such as summer foraging habitats or overwintering patches, a process also known as landscape complementation (Dunning *et al.* 1992; Denoel & Lehmann 2006; Semlitsch 2008). These movements are seasonal and concern all terrestrial life stages (Rothermel 2004). Individuals must have access to, but also to reach, favorable terrestrial resource patches, leading to a tradeoff between benefits of migration and survival costs related to crossing inhospitable features such as roads or crops (Fahrig *et al.* 1995; Mazerolle & Desrochers 2005; Rittenhouse *et al.* 2009; Janin *et al.* 2011). At the scale of migration, landscape connectivity is thus a key factor for the persistence of local populations. Besides migration, dispersal may be defined as between-population unidirectional movements from natal sites to distant breeding sites, and is a non-obligatory process for juveniles and adults (Perret *et al.* 2003; Palo *et al.* 2004; Semlitsch 2008). At this scale, by determining dispersal success, landscape connectivity contributes to the long-term regional persistence of populations, through rescue effect and (re-)colonization of empty sites (Trenham et al. 2001). Migration and dispersal are antagonistically related. Philopatric behaviour, i.e. fidelity to the breeding site, may result from a trade-off between breeding requirements and dispersal cost: when breeding habitat remains stable and suitable over

time, philopatry may prevail since a random dispersal behaviour may yield additional costs for individuals lacking knowledge on location, accessibility and quality of distant ponds; on the contrary, deterioration of habitat quality through ecological succession, increasing inbreeding or anthropogenic activities (dredging, fish introduction, pollution, increasing livestock) may favor dispersal behaviour (Semlitsch 2008; Kopecky et al. 2012).

Understanding how the landscape matrix influences these two kinds of movements is thus critical for the management and conservation of amphibian populations (Van Dyck & Baguette 2005; Semlitsch 2008). Indirect methods are generally used to assess the success of these movements: variables such as occurrence, abundance, genetic diversity and inter-site data (or pairwise data, e.g. indirect estimators of gene flow) are confronted to habitat and landscape matrix characteristics to identify the drivers of functional connectivity at both spatial scales (Manel et al. 2003; Pellet et al. 2004; Janin et al. 2009; Emaresi et al. 2011). However, several issues may complicate analysis of land use data, such as low sample size (financial or operational constraints may limit the number of sampled aggregates in pairwise genetic analyses; Prunier et al. *submitted-a*), and collinearity between variables that may hinder statistical and inferential interpretation (Graham 2003; Vaughan & Ormerod 2005). When individuals are aggregated, such as in pond-breeding amphibians, the number of sampled aggregates can be increased by using an individual-based sampling scheme, i.e. sampling only a few individuals per aggregate. This scheme provides a better spatial spreading out of pairwise distances across a given landscape, and therefore a better coverage of landscape heterogeneity than population-based sampling. As it proved to be particularly efficient at detecting spatial genetic structures at the scale of dispersal (Prunier et al. *submitted-a*), it may also provide accurate and numerous occurrence data through the study area for further analyses at the scale of complementation. Collinearity between variables may plague landscape analyses at both spatial scales (Graham 2003; Cushman & Landguth 2010b), but the use of principal component analyses (PCA) helps prevent this bias. PCAs have been widely used to reduce and analyze complex ecological data: they assume that multicollinearity among land use variables reflects the existence of a smaller number of underlying ecological factors, such as geographical gradients. Resulting principal components are then used as uncorrelated predictor variables (Graham 2003; Vaughan & Ormerod 2005; Manel *et al.* 2009).

In this study, we focused on migration and dispersal in the alpine newt *Ichthyosaura alpestris*, a widespread species in western and central Europe. At the scale of complementation, this species strongly depends on permanent fish-free ponds or small aquatic sites such as ruts, surrounded by suitable terrestrial habitat such as woods and semi-natural grasslands (Joly et al. 2001; Kopecky et al. 2012). It benefits from extensive livestock farming, which allows the preservation of many natural or man-made ponds in pastures. High dispersal rates have been reported in adults (breeding dispersal; Perret et al. 2003; Kopecky et al. 2012) and in male juveniles (Joly & Grolet 1996). This nomadic behaviour may explain high gene flow and weak genetic differentiation among subpopulations observed in previous studies (Emaresi et al. 2011; Prunier et al. *submitted-a*). Our objective was to identify landscape features, or combination of landscape features, driving the distribution pattern and the genetic structure of the alpine newt in a complex landscape using both occurrence and genotypic data gathered in the framework of an individual-based sampling scheme (Prunier et al. *submitted-a*). First using PCA to detect and integrate correlated habitat variables, and then using information theoretic model selection and multimodel inference (Burnham & Anderson 2002) at both scales, we looked into the following questions: (1) Are migration and dispersal occurring at distinct spatial scales and possibly concerning different life stages both driven by the same landscape features? (2) As sex-biased dispersal was reported in this species, do males and females perceive the matrix differently at the scale of migration and at the scale of dispersal? Answering these questions may provide insightful information for accurate management and conservation of the alpine newt, which coexists and shares many ecological characteristics with threatened amphibians such as *Lissotriton helveticus* or the endangered *Triturus cristatus* (Joly et al. 2001; Emaresi et al. 2011).

2. Material and methods

2.1. Study area and landscape features

The study was carried out over an area of approximately 20 x 25 km in *Puisaye*, a farmland region in Bourgogne (France). It is crossed by several streams running through small dales seperating low-elevated hilly plateaus along a southeast-northwest axis (Fig. 1). For the last 40 years, the *Puisaye* has been undergoing important landscape modifications, with the regrouping and the drainage of cultivated farmlands and the digging out of many hedgerows networks. However, sandy-clay soils being suitable for livestock farming and many natural obstacles (notably woods and standing waters) impeding an efficient regrouping of agricultural plots, these modifications were less drastic in *Puisaye* than in *Forterre*, the south-eastern limestone neighbouring region, which is mainly devoted to intensive farming (Pinton & Le Caro 2008).

In *Puisaye*, the land-cover is characterized by a complex assemblage of more or less fragmented wooded patches (woods, groves and hedgerows; 39.5%), extensively pastured meadows (22.7%) and plots of crop rotation (mainly cereals, often alternating with temporary pastured meadows from year to year; 33.5%). The study area also comprises a few small towns connected by roads with low to medium traffic, several reservoirs and fishing ponds, all these features representing less than 4.5 % of the study area (Fig.1).

Eight elements were retained to describe landscape characteristics: the three main land-cover features (wooded areas, permanent meadows and plots of crop rotation, hereafter denoted simply as crops), four secondary land-cover features (urbanized areas (towns), roads, standing waters and running waters) and one topographic variable (slope). Wooded patches, urbanized areas, roads, standing and running waters were extracted from national maps (BD Topo from National Geographic Institute, France, 1/25000). Semi-natural opened areas (meadows and crops) were identified by automatic classification from satellite images (ASTER /TERRA images with 15-m resolution) using the ENVI software (ITT, Boulder, CO, USA), completed by the 2008 French RPG (Registre parcellaire graphique), providing shape and cover-type for a large number of agricultural plots. Slope, measured in percent, was derived from a 50-m SRTM-Digital Elevation Model

(Shuttle Radar Topography Mission). We used ArcGIS 9.3 and its extension Spatial Analyst to manage, rasterize and compile these elements, respectively in a land-cover map and a topographic map. All the following map treatments, aiming to produce explanatory variables, were performed using the Matlab software coding environment (Mathworks, Inc.).

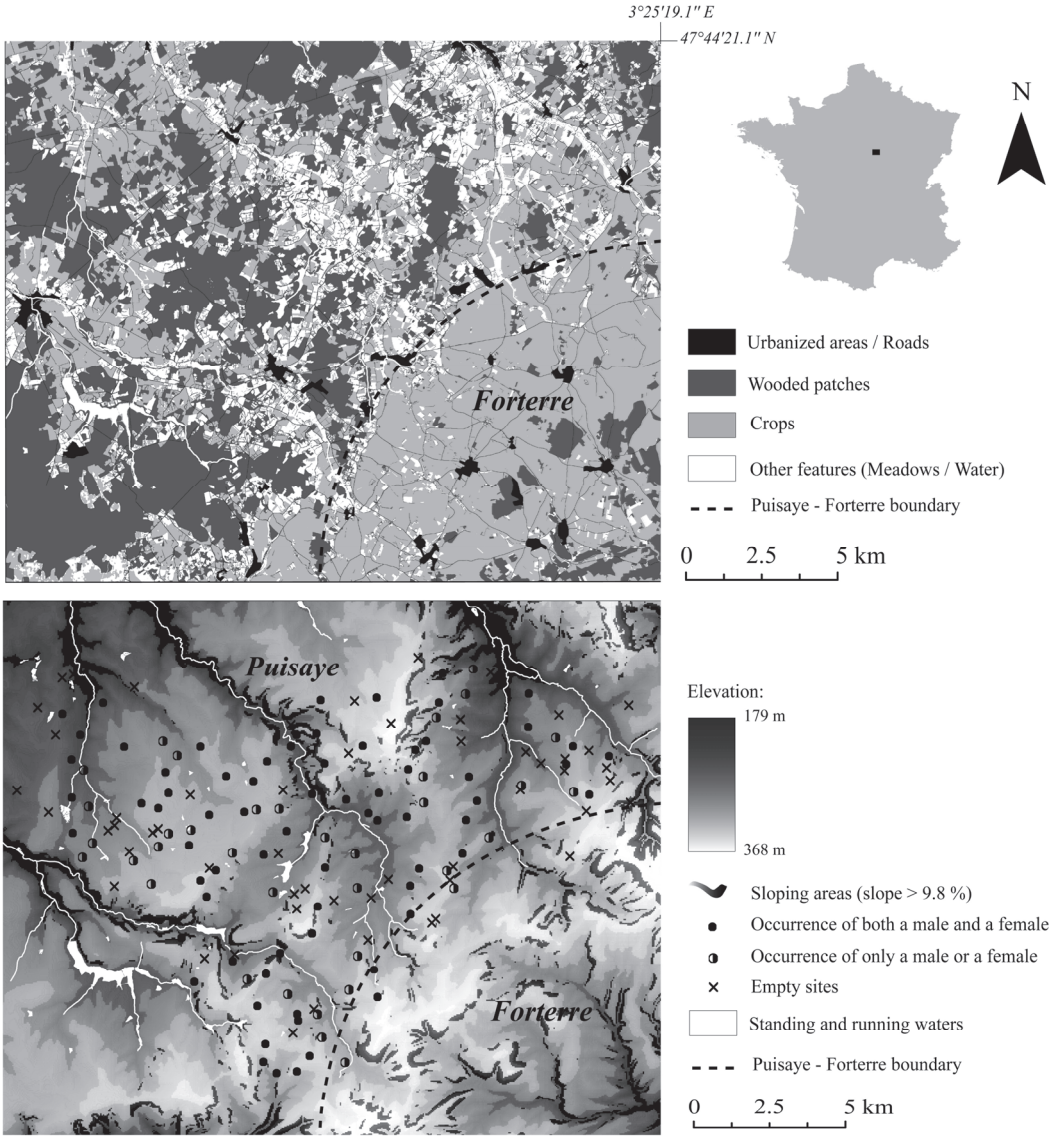


Figure 1. Main characteristics of the study area: Land-use cover (a) and topography along with distribution of *Ichthyosaura alpestris* among prospected fishless aquatic sites (b).

2.2. Field sampling, occurrence and genotypic data

Fieldwork was conducted during the 2009 breeding season. We prospected a total of 186 aquatic sites (man-made or natural ponds, flooded ruts, swamps, etc.). Each site was prospected with a dip net for thirty to sixty minutes (depending on site size and configuration) or until a male and a female alpine newt were captured. *I. alpestris* being highly sensitive to the presence of fish (Joly et al. 2001), the detection of fish immediately triggered the end of site prospection. This allowed an efficient genetic sampling coverage of the study area in a minimum amount of time (Prunier et al. *submitted-a*), while yielding reliable occurrence data for males (*occM*) and females (*occF*).

Non-destructive genetic samples were taken from each captured individual using buccal swabs (Broquet *et al.* 2007). Samples were stored at ambient temperature in an air-tight container with silica gel providing total desiccation. DNA extraction, PCR amplifications and genotyping were performed as described in Prunier *et al.* (2012), using 12 microsatellite loci: *CopTa1*, *CopTa2*, *CopTa3*, *CopTa4*, *CopTa8*, *CopTa10*, *CopTa11*, *CopTa12*, *CopTa13* and *CopTa14* (Prunier *et al.* 2012), *Ta1Ca1* and *Ta1Caga4* (Garner *et al.* 2003). Neither null alleles nor linkage disequilibrium were detected with this combination of markers (Prunier *et al.* 2012). The mean error rate per locus e_l (Pompanon *et al.* 2005), assessed by blind replication of 45 out of 1081 samples (4.2 %) collected from 2009 to 2011 as part of a general research program on *I. alpestris*, was less than 2.4 % in *CopTa1*, *CopTa4*, *CopTa14* and *Ta1Caga4*, and 0% for the 8 other loci (Prunier et al. *submitted-b*). For each sex, inter-individual pairwise genetic distances (B_c) were calculated using the Bray-Curtis percentage dissimilarity measure (Legendre & Legendre 1998; Cushman *et al.* 2006; Prunier *et al. submitted-a*).

2.3. Local analysis at the scale of migration

We first assessed the relative importance of landscape features on the occurrence of males and females. Landscape features were extracted from both the land-cover and the topographic maps within a 1000-m circular zone centered on each prospected site (Pellet *et al.* 2004; Denoel & Lehmann 2006). The 1000-m radius was chosen to encompass the potential migration area of *I. alpestris* (Joly et al. 2001; Kovar et al. 2009). This classical approach was preferred to the use of friction-based migration zones (Ray et

al. 2002; Janin et al. 2009) for simplicity, as the latter did not provide higher predictive power than the classical circular approach for *I. alpestris* in Ray et al. (2002). From the land-cover map and for each site, we first calculated the relative proportions of the three main land-cover features (woods, meadows and crops) once secondary land-cover features had been removed from the circular zone. In the same way, we then calculated the relative proportions of secondary land-cover features (urbanized areas, roads, standing and running waters) once main land-cover features have been removed: this procedure ensured that weak proportions of secondary features were correctly represented in regard to the proportions of main features. Finally, we calculated from the topographic map the proportions S_{75} and S_{95} of pixels with a value respectively higher than the 75th and the 95th percentiles extracted from slope histogram (respectively 6.1 and 9.8%; see Fig. S1 in Supplementary Material). Land-cover proportions were then condensed with either S_{75} or S_{95} in principal component analyses (PCA) using the R-package FactoMineR. Individual scores on each of the *p* retained principal components (principal components for local data L_p) were used as explanatory variables in a set of candidate logistic models [$Occ \sim \Sigma L_p$]. For each candidate model, we checked for spatial autocorrelation in residuals using Moran's I correlograms (R-package spdep; Dormann *et al.* 2007; Denoel & Ficetola 2008) and used information theoretic model selection and multimodel inference (Burnham & Anderson 2002) to explore how the additive relationships between principal components L_p may explain males and females occurrence (see section 2.5).

2.4. Pairwise analysis at scale of dispersal

To characterize the genetic structure in our datasets, we first assessed isolation-by-distance pattern (IBD) for each sex through non-directional Mantel correlograms using Genalex 6 (Smouse & Peakall 1999; Peakall & Smouse 2006; Borcard & Legendre 2012). Euclidean distance classes were defined every 2000 m (up to 20 km), resulting in ten binary matrices representing the membership of individuals to the distance class being tested (with 0 for pairs of individuals belonging to the same distance class and 1 otherwise). For each distance class, the binary matrix was compared to genetic distance matrices using simple Mantel tests with 1000 permutations and a 95% confidence interval was determined by bootstrap resampling (1000 iterations).

Second, we tried to determine how gene flow may be influenced by landscape features during dispersal. Landscape features were extracted from both the land-cover and the topographic maps using straight-line strips between each pair of sampled site, a procedure proposed by Emaresi et *al.* (2011) and initially used in *I. alpestris*. This method relies on the assumption that the alpine newt has a low perceptual range and may need wide dispersal corridors. A length-width ratio of 1:3 was thus chosen, as it proved to provide higher explanation power than constant-width strips or smaller ratio strips in this species. We calculated for each strip the relative proportions of the three main land-cover features, the relative proportions of the four secondary land-cover features and the slope proportions S_{75} and S_{95}. For each sex, land-cover proportions were then condensed with either S_{75} or S_{95} in PCA. Individual scores on each of the *p* retained principal components (principal components for pairwise data P_p) were used as explanatory variables in a set of candidate linear mixed models using SAS PROC MIXED (SAS Institute 2011) with two random terms denoting the pair of sites associated with each pairwise genetic distance and a Toeplitz(1) covariance structure, as proposed in Selkoe *et al.* (2010): this procedure presents an advantage when compared to other approaches (Cushman *et al.* 2006; Goldberg & Waits 2010; Emaresi *et al.* 2011) as it takes into account the non-independency of pairwise genetic data through the direct estimate of local site-effects. Following Emaresi et al. (2011), we hypothesized that Euclidean distance alone (δ) may not be sufficient to explain the genetic structure in our data: we designed a set of nested models, from the null model based on the IBD assumption [$B_c \sim \delta$] to the full model [$B_c \sim \delta + \Sigma P_p + \delta.(\Sigma P_p)$]. As previously, we then used information theoretic model selection and multimodel inference to explore how both the additive relationships between principal components P_p and their non-linear relationships with Euclidean distance (δ) may explain genetic distances among males and among females (see section 2.5). Euclidean distances were log-transformed following the $\delta = ln(d+1)$ formula and all continuous variables were standardized, as recommended by Schielzeth (2010).

2.5. Multimodel inference

For each local or pairwise candidate model, we calculated Akaike's information criterion (*AIC*) as a measure of model fit (Burnham & Anderson 2002). AIC differences (Δ_{AIC}) were then used for model ranking and *AIC* weights (w_i) were calculated as a weight

of evidence for the candidate model *i* being the best fitting model in each full set. We finally estimated the relative importance w_{j+} of each predictor *j* (L_p, P_p or $\delta.P_p$) by summing the *AIC* weights w_i across all the models where the predictor *j* occurred. Whenever $w_{j+} \geq 0.5$, we used model averaging to calculate the model averaged estimate and the unconditional variance estimator of the predictor *j*, while taking into account both the relative support of candidate models and uncertainty in model selection (Burnham & Anderson 2002). Model averaging was performed from subsets of all models including the predictor *j*, except models including the interactive predictor $\delta.P_p$ when only the additive predictor P_p was supported by w_{j+}.

3. Results and discussion

3.1. Principal component analyses and landscape characteristics

For each principal component analysis (PCA), we only retained the first three principal components, explaining approximately 72 % of the variation in landscape composition in circular zones and in strips (Fig. 2). In both local and pairwise analyses, PCA yielded very similar patterns and revealed multicollinearity among various variables, providing an accurate description of landscape characteristics.

The first principal components (L_1 for local data and P_1 for pairwise data) summarized the gradient associated with vegetal cover, from high proportions of woods (negative values) to high proportions of semi-natural open-areas (meadows and crops). They also revealed high collinearity between woods and standing waters: this pattern is consistent with landscape modifications which have occurred in *Puisaye* for the last decades: the regrouping of cultivated plots being impeded by natural obstacles, areas occupied by wooded patches and standing waters (reservoirs and fishing ponds) were preserved (Pinton & Le Caro 2008). The second principal components (L_2 and P_2) revealed collinearity between semi-natural opened-areas and slope: they segregated high proportions of crops located in flatter sectors (negative values) from high proportions of meadows located in sloping sectors. This latter characteristic is peculiar to farmland

landscapes, flat plots being reserved to mechanized agriculture while sloping plots are rather dedicated to pastures. Principal components being used as uncorrelated predictor variables, this latter landscape characteristic was directly incorporated in all landscape analyses (Graham 2003). Finally, the third principal components (L_3 and P_3) segregated high proportions of urbanized areas (positive values for L_3 and for P_3 in females, negative values for P_3 in males) from high proportions of roads. The contribution of running waters to these principal components was low, suggesting that this secondary land-cover feature was either under-represented when compared to others, or showed too little variability among sites (or strips). Very similar patterns were obtained when using S_{75} (data not shown). We thus only used S_{95} in the following analyses.

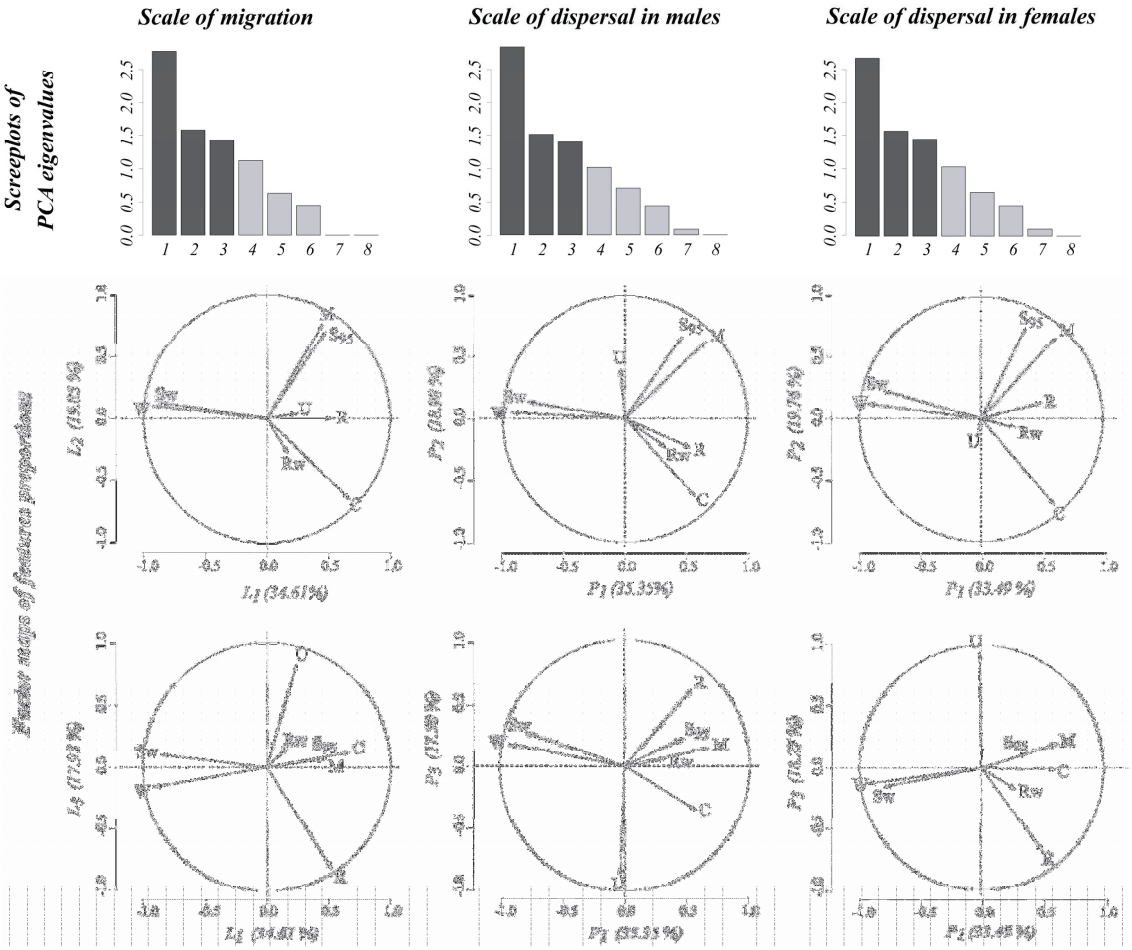


Figure 2. Results of principal component analyses at the scale of complementation and at the scale of dispersal in males and females. The figure shows (1) screeplots of PCA eigenvalues with retained principal components filled in dark grey ; (2) factor maps of features proportions with W: wooded patches; M: permanent meadows; C: plots of crop rotation; U: urbanized areas; R: roads; Rw: running waters; Sw: standing waters; S_{95}: sloping sectors (slope > 9.8%).

3.2. Landscape complementation

Fish were found in 23 out of 186 prospected sites. From the 163 remaining fishless ponds, 108 turned out to be occupied by alpine newts (66%). We detected 89 occurrences of males and 91 occurrences of females, with 72 sites occupied by both sexes (Fig. 1).

Given the three retained local principal components L_p, for each sex we designed a set of 7 candidate logistic models. We did not find evidence of autocorrelation in any model's residuals (data not shown). The best candidate model explaining occurrence data selected was [$Occ \sim L_2$] in both sexes, despite low AIC weights suggesting high uncertainty in model selection (Table 1). Nevertheless, the relative importance of this predictor L_2 was higher than 0.97 while it was pretty low for other predictors, indicating that the relative proportions of crops and meadows (along with slope) were the most important drivers for probability of occurrence. Model averaging provided positive averaged estimates in both sexes (Table 2): probability of occurrence was low when high proportions of crops were found in the direct vicinity of aquatic sites and increased with increasing proportions of meadows (along with slope). This result is consistent with previous studies (Joly *et al.* 2001; Denoel & Ficetola 2008) which reported a negative influence of cultivated land on abundance in this species. Permanent meadows located in sloping sectors were preserved from the regrouping of agricultural plots and thus still present important hedgerow networks connecting groves and low-fragmented wooded patches, along with many natural or man-made watering-places for livestock. Although hedgerows may not constitute substitutive terrestrial habitats for the alpine newt (Joly et al. 2001), they may however provide temporary protection against predation or desiccation (Semlitsch 2008; Cosentino et al. 2011). On the contrary, cultivated sectors may present higher risks of mortality, especially for emigrating juveniles (Rothermel & Semlitsch 2002; Cosentino *et al.* 2011).

The proportion of woods (L_1) was not retained as an important driver of pond occupancy (tables 1 and 2), contrary to previous studies which highlighted the importance of wooded patches in the alpine newt (Joly *et al.* 2001; Denoel & Ficetola 2007, 2008). This may come from the use of gradients (principal components) as explanatory variables: when explaining probability of occurrence, the contrast between crops and meadows along L_2 may be far more marked than the contrast between woods and semi-natural

opened-areas along L_1, since newts are likely to be encountered in both woodlands and pastures (see above). Last, neither urbanized areas nor roads (L_3) seemed to impact probability of occurrence. However, it is worth noting that occupancy does not mean viability (Denoel & Ficetola 2007): we may have detected transient individuals (Perret et al. 2003; see section 3.3), as may be suggested by the high proportion of sites where only one individual was captured (33 % of occupied sites, Fig. 1), or we may also have prospected declining populations.

Table 1. Predictor variables included in the top 4 local or pairwise models selected by AIC in each sex; *k* is the number of parameters in the model and w_i is the *AIC* weight, that is the weight of evidence for the model being the best fitting model in the full set of models; the numbers into brackets show, for each analysis, the number of candidate models in the full set of models. *L* = principal components for local variables; *P* = principal components for pairwise data.

Analysis	Sex	Model	k	AIC	Δ_{AIC}	w_i
Local	Males (7)	L_2	1	220.478	0.000	0.372
		$L_1 + L_2$	2	220.721	0.243	0.330
		$L_2 + L_3$	2	222.388	1.910	0.143
		$L_1 + L_2 + L_3$	3	222.637	2.160	0.126
	Females (7)	L_2	1	216.836	0.000	0.495
		$L_2 + L_3$	2	218.514	1.678	0.214
		$L_1 + L_2$	2	218.650	1.814	0.200
		$L_1 + L_2 + L_3$	3	220.324	3.488	0.086
Pairwise	Males (27)	$\delta + P_2$	2	9151.069	0.000	0.198
		$\delta + P_2 + P_3$	3	9151.622	0.553	0.150
		$\delta + P_2 + \delta * P_2$	3	9152.548	1.479	0.095
		$\delta + P_1 + P_2$	3	9152.777	1.709	0.084
	Females (27)	$\delta + P_2 + P_3 + \delta * P_3$	4	10202.094	0.000	0.145
		$\delta + P_1 + P_2 + P_3 + \delta * P_3$	5	10202.373	0.279	0.126
		$\delta + P_2$	2	10202.649	0.555	0.110
		$\delta + P_1 + P_2$	3	10203.049	0.955	0.090

3.3. Sex-biased dispersal

As stated in section 3.2, 89 males and 91 females were captured and genotyped. Using pairwise genetic distances as indirect estimators of gene flow, we detected several differences between sexes at the scale of dispersal. Mantel correlograms showed significant genetic relatedness among males up to 6 km while no correlation was observed for the two first kilometers (Fig. 3). This suggests high landscape connectivity at the scale of dispersal, with male being highly mobile and displaying a nomadic rather than philopatric behaviour (Liebgold et al. 2011). Although a similar pattern was observed in females, genetic autocorrelation occurred up to 4 km, while a trend for positive genetic relatedness was observed between sites localised less than 2 km apart (p-value: 0.068), suggesting that females may be slightly less mobile and more philopatric than males, probably because their investment in reproduction is higher than in males (Perret et al. 2003). These results are consistent with previous studies which detected or predicted high gene flow (Emaresi et al. 2011) and sex-biased dispersal in *I. alpestris* (Joly & Grolet 1996; Perret *et al.* 2003), while confirming that the organization of aggregated individuals may be better described using a patchy population model than a metapopulation model (Perret *et al.* 2003; Smith & Green 2005; Emaresi *et al.* 2011).

Given the three retained pairwise principal components P_p, for each sex we designed a set of 27 candidate mixed linear models. In both sexes, the selected best candidate model explaining genetic distances showed low AIC weights, here again suggesting high uncertainty in model selection (Table 1). Nevertheless, the relative importance of the predictor P_2 was higher than 0.95, indicating that the relative proportions of crops and meadows (along with slope) were likely to affect gene flow in both sexes. Model averaging provided positive averaged estimates (Table 2): surprisingly, the higher the proportion of meadows and slope (and respectively the lower the proportion of crops), the higher the genetic distances between sampled sites. This counter-intuitive result suggests that sloping sectors (dales, ridges) may act as barriers to gene flow or, in other words, that dispersal occurs preferentially within more or less flat sectors (plateaus). A similar observation had already been observed in this species, with genetic discontinuities detected along sides of low-elevated mountain massifs (Prunier et al. *submitted-b*). Although *I. alpestris* was observed climbing steep slopes during breeding migration (Vilter & Vilter 1962), that is during obligatory movements up to breeding

ponds, the energetic costs of moving up slopes (Funk *et al.* 2005a) may impede movement rates in the case of random dispersal between sites separated by dales or ridges. However, we could expect that the low costs of downward movements would compensate for the high costs of upwards movements, thus leading to suppress any effect of slope on gene flow. Another hypothesis may be proposed: suitable conditions in a landscape dominated by meadows may lead to higher phylopatry (i.e. low dispersal), whereas small populations in landscapes dominated by crop could experience higher inbreeding and stronger pollution by fertilizers and pesticides (Janin *et al.* 2009), leading to higher dispersal.

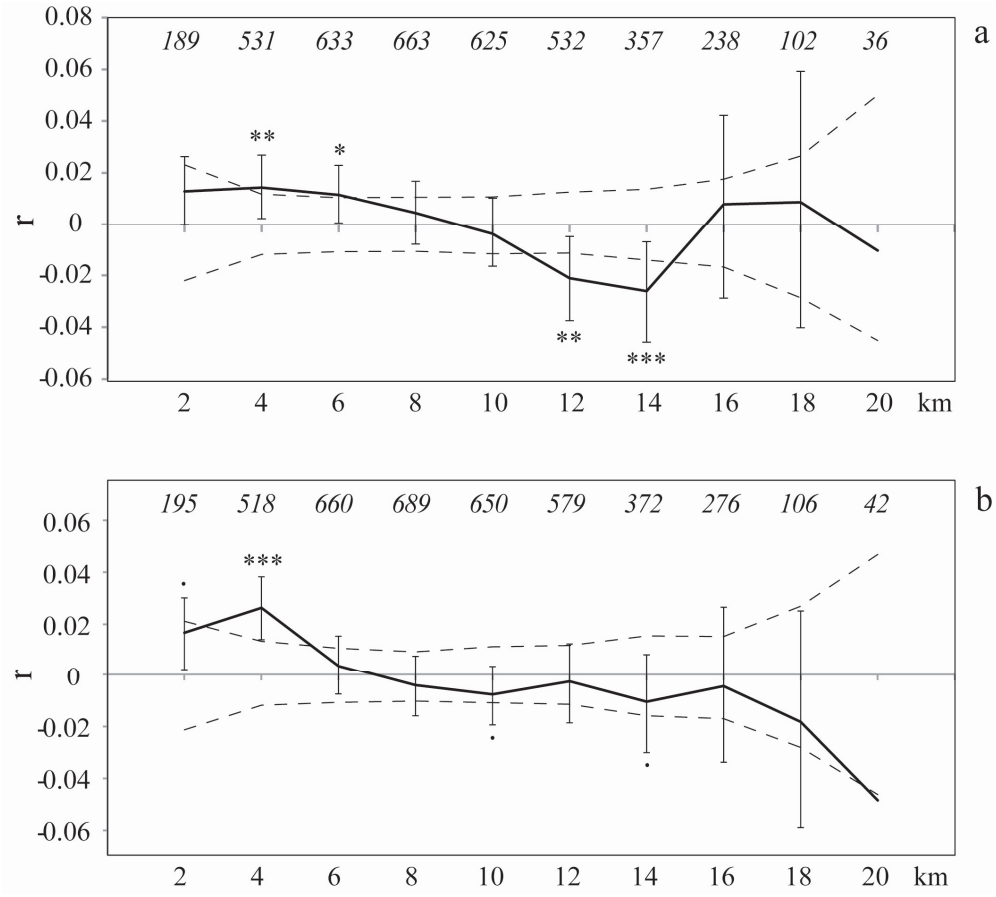


Figure 3. Mantel correlograms showing the relationships between inter-individual genetic distances and Euclidean distance classes (defined every 2000 m) in males (a) and in females (b). r: standard Mantel correlation with 1000 permutations. Error bars bound the 95% confidence interval about r as determined by bootstrap resampling. Upper and lower confidence limits (dotted line) bound the 95% confidence interval about the null hypothesis of no spatial structure as determined by permutation. ˙: p-value < 0.1; *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001.

While P_1 was not retained as a pertinent predictor in males, it showed a relative weight of evidence support of 0.57 in females and a positive averaged estimate, suggesting that females may be more sensitive to matrix permeability than males while confirming that wooded patches may contribute significantly to landscape connectivity in *I. alpestris* (Emaresi et al. 2011). Despite low to medium road-traffic in the study area, higher genetic distances were associated with high proportions of roads in both sexes (P_3), suggesting that increased mortality when crossing roads may impede dispersal rates (Holderegger & Di Giulio 2010). In males, P_3 showed a relative importance of 0.52 and a positive averaged estimate, actually suggesting that high proportions of roads may affect gene flow. P_3 was also retained as an important predictor in females, showing a relative importance of 0.65; however, among candidate models including P_3, models also including the interactive predictor $\delta.P_3$ were the most supported (relative importance of 0.78 for $\delta.P_3$, Table 2), revealing that variation in genetic distances among females may be explained by an interactive relationship between P_3 and Euclidean distance: while the relationship between genetic distances and proportions of roads was constant in males, it was only revealed for distant sites in females (Fig. 4). This result is consistent with the hypothesis of males being more mobile than females, the former being more likely to encounter roads than the latter (Carr & Fahrig 2001).

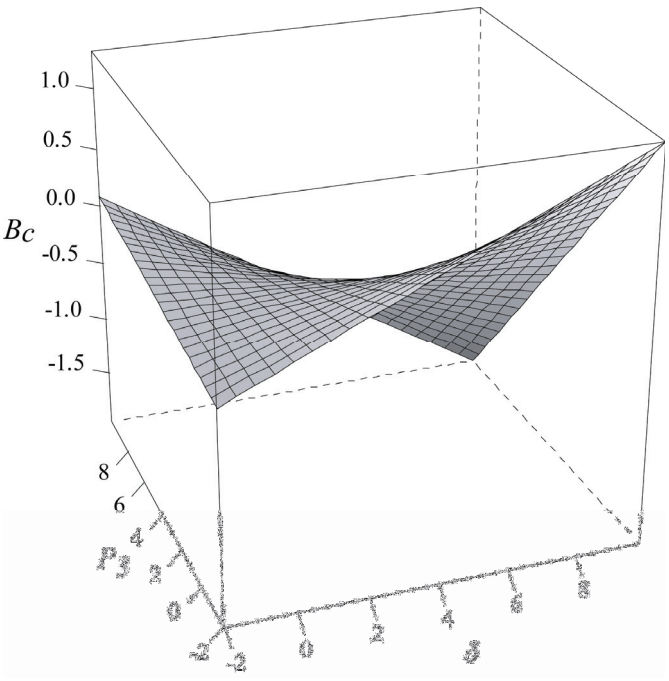


Figure 4. Genetic distances *Bc* (calculated using the Bray-Curtis percentage dissimilarity measure) predicted in function of the interaction between the Euclidean distance δ and the third pairwise principal component P_3 in females. Each parameter was estimated by model averaging (see section 2.5 and table 2). Ranges of values refer to the standardized values of *Bc*, δ and P_3.

Table 2. For each analysis, relative importance w_{j+} of each local or pairwise predictor (obtained by summing the AIC weights across all the models in a set where the predictor occurred); in bold, predictors with a relative importance higher than 0.5. Averaged estimates are given with their unconditional variance estimator (see section 2.5 for details)

Analysis	Sex	Variable	w_{j+}	Averaged estimate
Local	Males	*L1*	0.478	/
		L2	**0.972**	0.366 ± 0.132
		L3	0.282	/
	Females	*L1*	0.289	/
		L2	**0.994**	0.430 ± 0.135
		L3	0.304	/
Pairwise	Males	*P1*	0.374	/
		P2	**0.960**	0.049 ± 0.018
		P3	**0.518**	0.021 ± 0.017
		*δ*P1*	0.270	/
		*δ*P2*	0.325	/
		*δ*P3*	0.270	/
	Females	*P1*	**0.574**	0.027 ± 0.020
		P2	**0.955**	0.053 ± 0.019
		P3	**0.646**	-0.027 ± 0.023[a]
		*δ*P1*	0.347	/
		*δ*P2*	0.281	/
		*δ*P3*	**0.785**	-0.025 ± 0.015[a]

[a] Both the additive and the interactive predictors showed high relative importance in females pairwise analysis: the corresponding predictive equation is [Bc = 0.002 + 0.075 δ - 0.027 P_3 - 0.025 (δ*P_3)], with Bc the predicted genetic distance calculated using the Bray-Curtis coefficient. See Figure 4 for a representation.

3.4.Conservation implications

At the scale of dispersal, results revealed substantial movements among ponds, notably in males, as long as aquatic sites were not separated by sloping areas (dales or ridges), giving insight into the most appropriate geographical management units (i.e. plateaus) necessary for alpine newts' persistence at large scale (Funk *et al.* 2005a). Given

that a large amount of prospected ponds were well-kept watering-places for livestock, thus remaining stable and favorable over time despite crop rotation, we may have predicted high breeding site fidelity (Joly & Miaud 1989; Semlitsch 2008). Both this low apparent philopatry and the negative impact of crops on probability of occurrence rather suggest that the unpredictable deterioration of habitat quality, known to trigger dispersal behaviour, is likely to occur at the scale of complementation. Actually, the agricultural practice in the studied area (Pinton & Le Caro 2008), characterized by an alternation of cereal crops and temporary meadows, may lead to a regular turn-over in the quality and/or accessibility of terrestrial habitats surrounding breeding sites. In such a situation, a nomadic behaviour, with high rates of breeding dispersal or transience (Perret et al. 2003), may enable to deal with the dynamic spatial distribution and temporal rotation of crops (Cosentino et al. 2011) and ensure a better exploitation of suitable breeding patches over time, as long as landscape connectivity remains favorable at the scale of dispersal. Our results confirmed however that the widespread alpine newt could be affected by habitat loss and landscape fragmentation (Emaresi et al. 2011): while wooded patches may constitute dispersal corridors, anthropogenic features and unsuitable agricultural practices are likely to impede migration and dispersal movements of juveniles and breeding adults (notably females), highlighting the need for a better comprehension of both individuals requirements at all life-stages and local landscape characteristics in amphibian management and conservation. Further work is nevertheless needed to investigate the determinants of dispersal propensity that could be related to the spreading of intensive agriculture.

4. Conclusion

Our study provided an insightful overview of landscape connectivity at the scales of complementation and dispersal in the alpine newt *I. alpestris*, consistent with several demographic and genetic studies (Joly & Grolet 1996; Perret *et al.* 2003; Emaresi *et al.* 2011). Our results highlighted the inferential advantages to combine demographic and genetic data for assessing population connectivity (Lowe & Allendorf 2010), although we are conscious that occurrence is the most elementary demographic data that may be gathered in the field. Nevertheless, the use of an individual-based sampling scheme

turned out to provide valuable information for the assessment of matrix permeability to gene flow (Prunier et al. *submitted-a*), notably when combined with principal component analyses, taking into account the unique characteristics of the studied landscape, and with the use of mixed models as described in Selkoe *et al.* (2010), dealing with the non-independency of pairwise genetic distances.

References

Beebee TJC, Griffiths RA (2005) The amphibian decline crisis: A watershed for conservation biology? *Biological Conservation* **125**, 271-285.

Borcard D, Legendre P (2012) Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology* **93**, 1473-1481.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics* **8**, 509-511.

Burnham KP, Anderson DR (2002) *Model selection and multimodel inference - A practical information-theoric approach*, 2nd Edition. Springer-Verlag, New York, New York, USA.

Carr LW, Fahrig L (2001) Effect of road traffic on two amphibian species of differing vagility. *Conservation Biology* **15**, 1071-1078.

Cosentino BJ, Schooley RL, Phillips CA (2011) Connectivity of agroecosystems: dispersal costs can vary among crops. *Landscape Ecology* **26**, 371-379.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation* **128**, 231-240.

Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology* **19**, 3592-3602.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* **168**, 486-499.

Denoel M, Ficetola GF (2007) Landscape-level thresholds and newt conservation. *Ecological Applications* **17**, 302-309.

Denoel M, Ficetola GF (2008) Conservation of newt guilds in an agricultural landscape of Belgium: the importance of aquatic and terrestrial habitats. *Aquatic Conservation-Marine and Freshwater Ecosystems* **18**, 714-728.

Denoel M, Lehmann A (2006) Multi-scale effect of landscape processes and habitat quality on newt abundance: Implications for conservation. *Biological Conservation* **130**, 495-504.

Dormann CF, McPherson JM, Araujo MB*, et al.* (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609-628.

Dunning JB, Danielson BJ, Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* **65**, 169-175.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conservation Genetics* **12**, 41-50.

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

Fahrig L, Pedlar JH, Pope SE, Taylor PD, Wegner JF (1995) Effect of Road Traffic on Amphibian Density. *Biological Conservation* **73**, 177-182.

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography* **16**, 265-280.

Funk WC, Blouin MS, Corn PS*, et al.* (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-496.

Garner TWJ, Schmidt BR, Hoeck P, Van Buskirk J (2003) Di- and tetranucleotide microsatellite markers for the Alpine newt (*Triturus alpestris*): characterization and cross-priming in five congeners. *Molecular Ecology Notes* **3**, 186-188.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology* **19**, 3650-3663.

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809-2815.

Holderegger R, Di Giulio M (2010) The genetic effects of roads: A review of empirical evidence. *Basic and Applied Ecology* **11**, 522-531.

Janin A, Lena JP, Joly P (2011) Beyond occurrence: Body condition and stress hormone as integrative indicators of habitat availability and fragmentation in the common toad. *Biological Conservation* **144**, 1008-1016.

Janin A, Lena JP, Ray N*, et al.* (2009) Assessing landscape connectivity with calibrated cost-distance modelling: predicting common toad distribution in a context of spreading agriculture. *Journal of Applied Ecology* **46**, 833-841.

Joly P, Grolet O (1996) Colonization dynamics of new ponds, and the age structure of colonizing Alpine newts, *Triturus alpestris*. *Acta Oecologica* **17**, 599-608.

Joly P, Miaud C (1989) Fidelity to the breeding site in the alpine newt *Triturus alpestris*. *Behavioural Processes* **19**, 47-56.

Joly P, Miaud C, Lehmann A, Grolet O (2001) Habitat matrix effects on pond occupancy in newts. *Conservation Biology* **15**, 239-248.

Kopecky O, Vojar J, Denoel M (2012) Sex-specific effect of pool desiccation on the movement of Alpine Newts, *Mesotriton alpestris* (LAURENTI, 1768), among breeding sites (Caudata: Salamandridae). *Herpetozoa* **24**, 127-134.

Kovar R, Brabec M, Vita R, Bocek R (2009) Spring migration distances of some Central European amphibian species. *Amphibia-Reptilia* **30**, 367-378.

Legendre L, Legendre P (1998) *Numerical Ecology* 2nd English ed. Elsevier., Amsterdam.

Liebgold EB, Brodie ED, Cabe PR (2011) Female philopatry and male-biased dispersal in a direct-developing salamander, *Plethodon cinereus*. *Molecular Ecology* **20**, 249-257.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038-3051.

Manel S, Conord C, Despres L (2009) Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. *Bmc Evolutionary Biology* **9**.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.

Mazerolle MJ, Desrochers A (2005) Landscape resistance to frog movements. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **83**, 455-464.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Pellet J, Guisan A, Perrin N (2004) A concentric analysis of the impact of urbanization on the threatened European tree frog in an agricultural landscape. *Conservation Biology* **18**, 1599-1606.

Perret N, Pradel R, Miaud C, Grolet O, Joly P (2003) Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology* **72**, 567-575.

Pinton F, Le Caro Y (2008) Protection de l'eau et remembrement en Puisaye-Forterre. *Programme Dialog. AgroPariTech and Université Rennes 2.* *www.afip.asso.fr/IMG/pdf/Monographie_Puisaye_Forterre.pdf*.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-859.

Prunier J, Kaufmann B, Fenet S*, et al.* (*submitted-a*) Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling scheme.

Prunier J, Kaufmann B, Grolet O*, et al.* (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (*Ichthyosaura alpestris*). *Molecular Ecology Resources* **12**, 524-531.

Prunier J, Kaufmann B, Lena JP, Pompanon F, Joly P (*submitted-b*) Permeability of large transport infrastructures to gene flow: individual-based sampling scheme combined to sPCA reveal counterintuitive patterns in the Alpine newt *Ichthyosaura alpestris*.

Ray N, Lehmann A, Joly P (2002) Modeling spatial distribution of amphibian populations: a GIS approach based on habitat matrix permeability. *Biodiversity and Conservation* **11**, 2143-2165.

Ricketts TH (2001) The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist* **158**, 87-99.

Rittenhouse TAG, Semlitsch RD, Thompson FR (2009) Survival costs associated with wood frog breeding migrations: effects of timber harvest and drought. *Ecology* **90**, 1620-1630.

Rothermel BB (2004) Migratory success of juveniles: A potential constraint on connectivity for pond-breeding amphibians. *Ecological Applications* **14**, 1535-1546.

Rothermel BB, Semlitsch RD (2002) An experimental investigation of landscape resistance of forest versus old-field habitats to emigrating juvenile amphibians. *Conservation Biology* **16**, 1324-1332.

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103-113.

Selkoe KA, Watson JR, White C*, et al.* (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* **19**, 3708-3726.

Semlitsch RD (2008) Differentiating migration and dispersal processes for pond-breeding amphibians. *Journal of Wildlife Management* **72**, 260-267.

Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography* **28**, 110-128.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Stuart SN, Chanson JS, Cox NA*, et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783-1786.

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. *Oikos* **68**, 571-573.

Trenham PC, Koenig WD, Shaffer HB (2001) Spatially autocorrelated demography and interpond dispersal in the salamander Ambystoma californiense. *Ecology* **82**, 3519-3530.

Van Dyck H, Baguette M (2005) Dispersal behaviour in fragmented landscapes: Routine or special movements? *Basic and Applied Ecology* **6**, 535-545.

Vaughan IP, Ormerod SJ (2005) Increasing the value of principal components analysis for simplifying ecological data: a case study with rivers and river birds. *Journal of Applied Ecology* **42**, 487-497.

Vilter A, Vilter V (1962) Migration de reproduction chez le Triton alpestre des Alpes vaudoises. *Comptes Rendus de l'Académie des Sciences* **156**, 2005-2007.

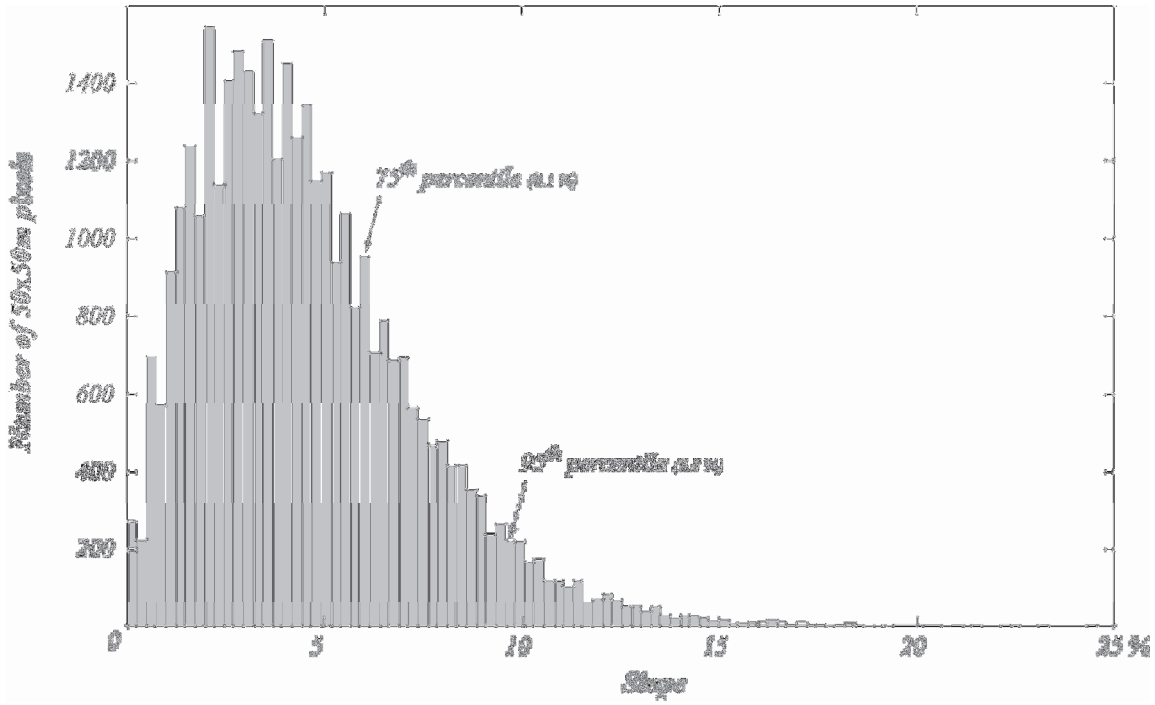
Acknowledgements

This work was supported by the ANRT (Agence Nationale de la Recherche et de la Technologie, agreement n° 2/2009), RFF (Réseau Ferré de France), APRR (Autoroutes Paris-Rhin-Rhône) and the ECOSPHERE company. It was conducted in accordance with French laws and with the approval of the Préfecture de l'Yonne. We warmly thank M. Clasquin, J. Mangin and O. Grolet for their precious help and support in the field and laboratory.

Supporting information

Figure S1. Histogram of slope values extracted from the topographic map.

Figure S1. Histogram of slope values extracted from the topographic map.



CONCLUSIONS ET PERSPECTIVES

A l'origine, mon travail de thèse avait un objectif à visée clairement opérationnelle : employer une méthode basée sur la génétique du paysage pour mesurer l'impact des grandes infrastructures de transport sur les déplacements du triton alpestre, et étendre cette méthode à l'étude des impacts des éléments paysagers surfaciques. Les contraintes financières et la distribution spatiale semi-agrégée des animaux m'ont rapidement amené à examiner de quelle façon l'information génétique pouvait être collectée dans l'espace de façon optimale pour mesurer l'impact d'un élément paysager. Dans ce cadre, je me suis dans un premier temps intéressé aux performances de l'échantillonnage génétique individu-centré. Les analyses corrélatives en génétique du paysage se faisant à partir de paires de points, j'ai été amené à considérer dans un second temps les différents problèmes de non-indépendance des données qui en résulte et à envisager de quelle façon ces problèmes pouvaient être résolus. L'ensemble de ces travaux m'ont permis de mieux dégager les limites et les avantages de l'approche individu-centrée par rapport aux approches populationnelles, et d'envisager comment concilier les avantages de chacune de ces approches. Dans cette discussion, j'aborderai successivement ces différents aspects.

Apport de l'échantillonnage individu-centré en génétique du paysage

L'approche individu-centrée est basée sur des mesures de distances génétiques interindividuelles, obtenues entre paire de points échantillonnés dans l'espace. Si cette approche est naturelle dès lors que l'on s'intéresse à un peuplement continu (p. ex. Broquet *et al.* 2006b; Cushman *et al.* 2006), ce n'est pas le cas lorsque les individus sont distribués de façon plus ou moins agrégée dans l'espace. Dans ce cas de figure, les agrégats sont généralement considérés comme unité d'échantillonnage, et la différentiation génétique entre paires de points est estimée à partir de modèles génétiques théoriques. Résumant l'information génétique à une simple mesure statistique, l'approche individu-centrée ne permet pas d'étudier finement les processus génétiques conduisant à la structuration observée, mais offre par contre beaucoup plus de flexibilité lorsque l'on s'intéresse spécifiquement à l'influence des éléments paysagers.

Plus de flexibilité en génétique du paysage

L'approche individu-centrée permet tout d'abord de s'affranchir du problème de la délimitation des populations *a priori* (Manel *et al.* 2003), le traitement de l'information génétique étant individuel plutôt que populationnel (comparaison des génotypes plutôt que des fréquences alléliques). L'utilisation d'un échantillonnage individu-centré s'est ainsi révélé tout à fait pertinent dans le cadre de l'étude de la structuration génétique du triton alpestre *Ichthyosaura alpestris*. Caractérisée par un fort nomadisme assurant une forte probabilité de recolonisation de sites éventuellement laissés vacants (Perret *et al.* 2003), cette espèce ne suit pas un modèle de peuplement métapopulationnel, mais semble plutôt se répartir dans l'espace selon un modèle de peuplement hétérogène (*patchy populations* ; Harrison 1991; Smith & Green 2005). Dans ce contexte, la délimitation *a priori* des populations étant particulièrement problématique, l'utilisation de mesures de différentiation génétique entre agrégats sur la base de l'estimation de fréquences alléliques (p. ex. Emaresi *et al.* 2011) ne semble pas être l'approche la plus pertinente pour l'étude de la dispersion chez cette espèce.

L'approche individu-centrée offre par ailleurs une grande flexibilité dans la conception des plans d'échantillonnage. Puisque seuls quelques individus sont échantillonnés en chaque site et que les durées de prospection sont réduites, elle permet de démultiplier, à moindre coût (tant financier que temporel), le nombre de points d'échantillonnage dans l'espace, autorisant par là une meilleure couverture d'échantillonnage, aléatoire ou ciblé, du paysage. Cette couverture d'échantillonnage permet de mieux caractériser l'influence de la composition et de la configuration de la matrice paysagère, mais également d'autres variables environnementales (p. ex. climatiques ; Manel *et al.* 2012) sur les flux de gènes (Anderson *et al.* 2010 ; Chapitre 1), que ce soit par des approches corrélatives ou *ad hoc*.

Dans le cas des deux infrastructures linéaires de transport étudiées dans ce mémoire (Chapitre 3), les approches corrélatives n'ont pas permis de mettre en évidence un quelconque effet barrière, malgré l'utilisation d'un échantillonnage intensif et ciblé, et malgré le fait que d'autres infrastructures, relativement récentes et d'origine humaine telles que le réseau de transport secondaire, aient été détectées comme responsables d'une structuration génétique des peuplements chez l'alpestre (Chapitre 4). L'ensemble de ces résultats suggère qu'autoroutes et lignes à grande vitesse ne constituent pas des barrières totales à la dispersion chez cette espèce, mais permettraient, de par leurs caractéristiques, une circulation des organismes de part et d'autres des voies de circulation. Des analyses complémentaires (par exemple comportementales) restent toutefois nécessaires pour confirmer cette perméabilité des infrastructures. Par ailleurs, des approches *ad hoc*, telles que des analyses en composantes principales spatialisées (ACP spatialisées ; Jombart *et al.* 2008), ont permis de détecter une structuration génétique liée à la topographie de la zone d'étude et, pour les femelles, un potentiel effet « corridor longitudinal » de l'autoroute A6 au sein des zones les plus fragmentées (zones d'agriculture intensive). Cette dernière observation doit être considérée avec prudence, car elle correspond à une structuration génétique relativement ténue (troisième axe ACP) : elle mériterait donc également une analyse plus approfondie, par exemple par suivi télémétrique. Quoiqu'il en soit, elle soulève la question d'une perméabilité de la matrice paysagère différente en fonction du sexe chez le triton alpestre. Bien qu'un biais sexuel dans le patron de dispersion soit un phénomène courant, cet aspect n'est pas (si) souvent considéré en génétique du paysage (hormis chez les angiospermes, où la dispersion des

gamètes est inévitablement sexe-biaisée). Cela peut pourtant conduire à sous-estimer l'impact des éléments paysagers en termes de flux démographiques. De plus, la structuration paysagère variant en fonction de l'échelle spatiale, on peut s'attendre à ce que les contraintes exercées par les structures paysagères soient différentes en fonction de la distance moyenne parcourue par les animaux. Chez le triton alpestre, une dispersion biaisée en faveur des mâles a déjà été mise en évidence par plusieurs études démographiques (Joly & Grolet 1996; Perret *et al.* 2003) bien que jamais révélée ou prise en compte en génétique du paysage (Pabijan *et al.* 2005; Pabijan & Babik 2006; Emaresi *et al.* 2011). Dans le cas où la dispersion est biaisée en faveur d'un sexe, on peut s'attendre, pour le sexe dispersant, à un effet *Wahlund* au sein des agrégats, qu'une approche populationnelle peut permettre de détecter (par mesure de *Fis* ou de scores d'assignement ; Goudet *et al.* 2002; Palo *et al.* 2004; Liebgold *et al.* 2011). Dans le cadre de mes travaux, j'ai examiné si un échantillonnage individu-centré pouvait également permettre de détecter un tel biais de dispersion au sein d'un paysage agricole complexe (Chapitre 4). Les résultats de ces travaux montrent que tel est effectivement le cas, et suggèrent donc que l'approche individu-centrée est tout aussi pertinente qu'une approche populationnelle classique pour tester ce genre d'hypothèses. Mes résultats suggèrent ainsi que les femelles sont moins mobiles et plus sensibles à la fragmentation de l'habitat forestier que les mâles, et soulignent également l'importance des zones boisées pour le maintien de la connectivité fonctionnelle du paysage chez le triton alpestre. Il serait néanmoins intéressant de déterminer à quel point l'augmentation de la couverture d'échantillonnage de la zone d'étude, autorisée dans le cadre d'une approche individu-centrée, peut contribuer à la détection d'un biais de dispersion chez une espèce moins mobile : le degré de mobilité des organismes et l'intensité du biais de dispersion sont en effet des facteurs-clés pour la détection d'un tel phénomène (Goudet *et al.* 2002).

Il est également à noter que, chez *Ichthyosaura alpestris*, les densités d'individus observées au niveau des sites de reproduction peuvent être relativement faibles. Ceci s'explique sans doute par l'existence de nombreux sites favorables à la reproduction sur la zone d'étude, une simple ornière suffisant par exemple à accueillir un petit nombre d'individus reproducteurs (Kopecky *et al.* 2012). Dans le cadre de mes travaux, ces densités étaient bien souvent trop faibles pour permettre l'obtention de fréquences alléliques non-biaisées (Kalinowski 2005; Broquet & Petit 2009). Si la présence d'agrégats

non échantillonnés au sein d'une zone d'étude (*ghost populations* ; Waples & Gaggiotti 2006; Marko & Hart 2011) est susceptible de biaiser l'estimation des paramètres d'intérêt tels que les taux de dispersion entre sites, l'approche individu-centrée autorise la collecte et l'exploitation de données génétiques en des sites qui auraient été écartés de tout plan d'échantillonnage populationnel.

L'ensemble de ces résultats souligne l'intérêt de l'approche individuelle par rapport à une approche populationnelle, notamment lorsque l'on s'intéresse à une espèce dont la biologie est méconnue (Anderson *et al.* 2010): l'échantillonnage individu-centré permet de récolter des données génétiques fiables, facilement exploitables et suffisamment pertinentes, sans avoir à connaître par avance la dynamique de peuplement ou le comportement de dispersion de l'espèce-cible.

Un dernier élément de flexibilité offert par l'échantillonnage individu-centré mérite d'être mis en lumière : la démultiplication du nombre de points d'échantillonnage, qui, comme discuté plus haut, peut se décliner dans l'espace, peut également se décliner dans le temps. Cette approche permet donc d'envisager l'obtention de séries temporelles de la structuration génétique spatiale, assurant par là une plus grande fiabilité dans l'interprétation des résultats (Anderson *et al.* 2010).

Maximisation du rapport signal / bruit (signal-to-noise ratio)

Dans le cadre d'un échantillonnage individu-centré, si la multiplication du nombre de génotypes dans l'espace permet une meilleure couverture d'échantillonnage de l'hétérogénéité du paysage, chaque agrégat d'individus n'est malgré tout que très partiellement échantillonné. Cette perte d'information à l'échelle locale, couplée à la démultiplication du nombre de mesures appariées, est susceptible d'entraîner une augmentation du « bruit de fond » par rapport au signal génétique que l'on cherche à détecter. Cette augmentation du bruit de fond peut être compensée de plusieurs manières.

Tout d'abord à l'échelle locale : l'échantillonnage en chaque agrégat d'un nombre suffisant d'individus permet de limiter le taux de collecte d'individus non représentatifs de leur site de capture, le risque d'échantillonner un individu « de passage » étant par exemple d'autant plus élevé que l'espèce-cible est mobile (Waples 1998; Perret *et al.*

2003). L'augmentation de ce nombre d'échantillons à l'échelle locale ne doit pas pour autant limiter la couverture d'échantillonnage de la zone d'étude et dépend donc notamment des ressources financières dont peuvent disposer les chercheurs : échantillonner entre 3 et 10 individus par agrégat (au lieu de 20 ou 30 dans le cadre d'une approche populationnelle) semble raisonnable (Chapitre 1 ; Figure M4). Par ailleurs, l'utilisation d'un grand nombre de marqueurs génétiques hautement polymorphes peut permettre d'améliorer la résolution de l'information génétique récoltée localement (Landguth *et al.* 2012) : à ce titre, l'utilisation des SNPs (*single-nucleotide polymorphisms*), dont l'obtention par centaines ou par milliers est désormais facilitée par les nouvelles techniques de séquençage existantes, s'avère particulièrement intéressante (Hedrick 2001; Morin *et al.* 2004; Balkenhol *et al.* 2009a; Storfer *et al.* 2009).

La maximisation du rapport signal/bruit peut également être assurée par la mise en place d'un échantillonnage ciblé (*targeted sampling scheme*), permettant de concentrer l'échantillonnage de la variabilité génétique au niveau d'éléments paysagers dont on chercherait à connaître spécifiquement l'influence sur les flux de gènes (Chapitre 1).

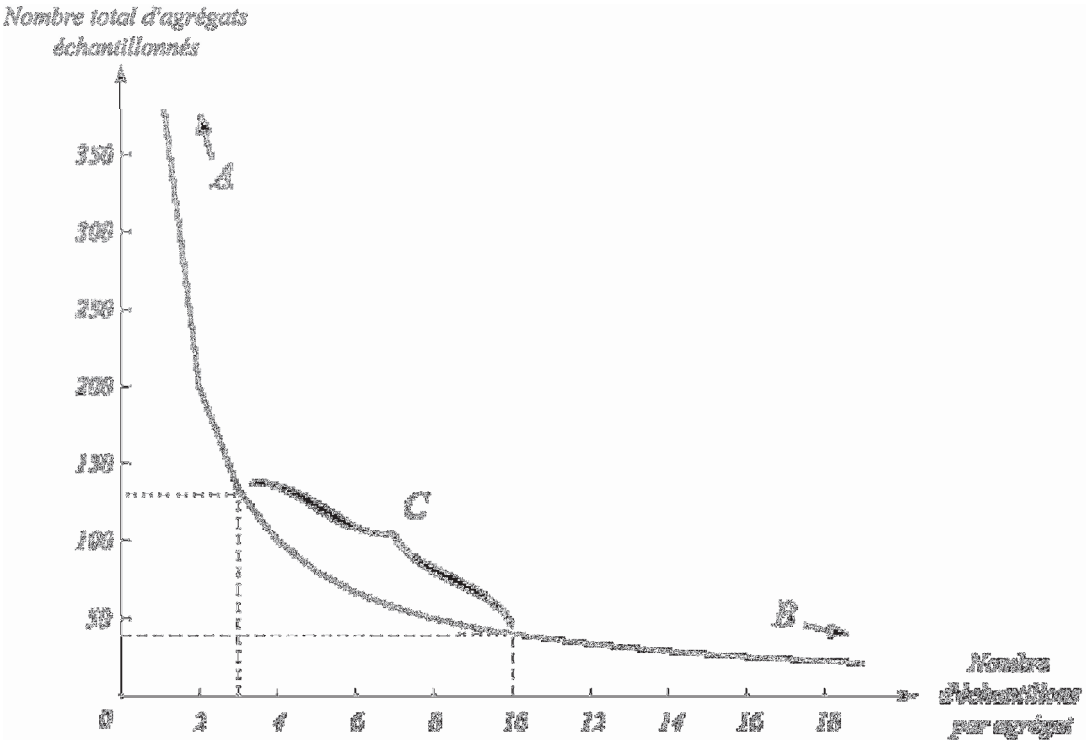


Figure M4 (complétant la Figure M3 présentée en introduction du manuscrit): Compromis entre couverture d'échantillonnage du paysage (A : accent mis sur la variabilité paysagère) et exhaustivité d'échantillonnage des agrégats (B : accent mis sur la variabilité génétique dans le cadre d'une approche populationnelle). L'exemple présente le cas théorique de la répartition spatiale de 400 échantillons : dans le cadre d'une approche individu-centrée (C), diminuer l'échantillonnage des agrégats à un nombre de prélèvements compris entre 3 et 10 permet d'envisager la prospection de 40 et 130 agrégats, contre au maximum 20 dans le cadre d'une approche classique (échantillonnage de 20 individus par agrégat).

Prise en compte de la non-indépendance des données

Quelle que soit la procédure d'échantillonnage utilisée, les études en génétique du paysage sont confrontées à des problèmes de non-indépendance des données génétiques et paysagères. Le premier problème provient du fait que les données génétiques et paysagères utilisées correspondent à des mesures appariées et ne peuvent donc pas être considérées comme indépendantes ; le second, du fait que les structures paysagères sont très souvent multi-colinéaires.

Non-indépendance des données appariées

Le premier niveau de non-indépendance résulte de l'utilisation de données appariées (ou mesures de distances : distances génétiques, distances euclidiennes, distances de moindre coût, etc. ; Slatkin 1985; Adriaensen *et al.* 2003). Les mesures de distances sont par nature non-indépendantes les unes des autres car un même point d'échantillonnage peut être utilisé pour calculer plusieurs mesures appariées. Cette non-indépendance statistique peut induire une covariance positive au sein des résidus, contaminant les estimations de pentes et d'ordonnées à l'origine lors de l'utilisation de régressions linéaires et entraînant un risque accru d'erreurs de type 1 (Selkoe *et al.* 2010). Ce problème est généralement contourné par l'utilisation de tests statistiques non paramétriques permettant de comparer des matrices de distances sur la base de permutations (tests de Mantel simples ou partiels ; Smouse *et al.* 1986). Bien que leur pertinence ait fait l'objet de nombreux débats (Raufaste & Rousset 2001; Castellano & Balletto 2002; Legendre & Fortin 2010; Borcard & Legendre 2012), ces analyses de corrélation matricielle semblent tout à fait intéressantes lorsqu'elles sont utilisées dans le cadre d'une modélisation de type causale (*causal modeling framework* ; Cushman *et al.* 2006). Toutefois, les tests de Mantel ne peuvent être utilisés dans un contexte de sélection de modèles (Burnham & Anderson 2002) car la non-indépendance des données est seulement prise en compte par permutations lors du test de significativité, tandis que la mesure de corrélation entre variables reste basée sur une régression linéaire classique

(mais voir Roach *et al.* 2001). Pour s'affranchir de ce problème de non-indépendance des données, il est possible de ne travailler que sur des paires de points indépendantes les unes des autres : cette indépendance peut être testée après échantillonnage (Goldberg & Waits 2010) ou bien obtenue par construction d'un plan d'échantillonnage au sein duquel chaque point d'échantillonnage ne serait impliqué que dans une seule mesure de distance (Figure M5).

Toutefois, cette dernière approche implique de pouvoir disposer d'un grand nombre de sites d'échantillonnages (à chaque donnée correspondrait en effet deux sites d'échantillonnage) et ne peut donc être sérieusement envisagée que dans le cadre d'un échantillonnage individu-centré. Ce pourrait être d'ailleurs l'occasion de cibler directement des éléments paysagers d'intérêt par de multiples réplicats de données appariées indépendantes, combinant dès lors indépendance et puissance statistique (Anderson *et al.* 2010).

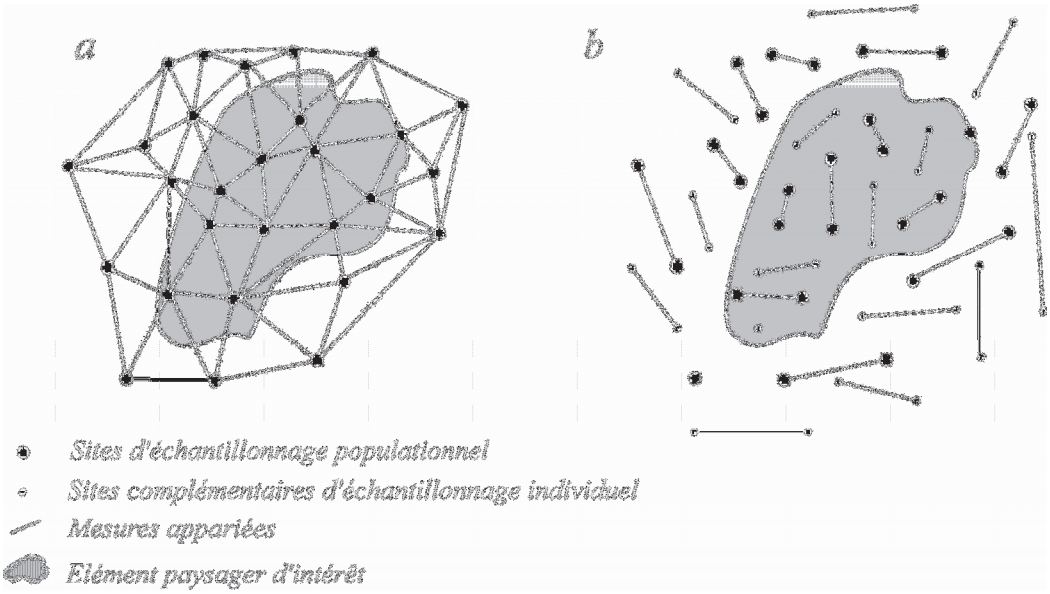


Figure M5 : (a) Plan d'échantillonnage populationnel aléatoire classique : les données appariées, ici obtenues sur la base d'une triangulation de Delaunay, ne sont pas indépendantes, et l'élément paysager d'intérêt n'est pas directement ciblé.
(b) Plan d'échantillonnage envisageable lors d'une augmentation du nombre de sites d'études grâce à une approche individu-centrée : les paires indépendantes de points pourraient être utilisées pour comparer la perméabilité de l'élément paysager d'intérêt par rapport à la perméabilité du reste de la matrice (ou par rapport à la perméabilité d'autres éléments paysagers) par simple ANOVA, après prise en compte de la distance euclidienne entre paires de points.

L'utilisation de modèles linéaires mixtes, permettant de traiter la corrélation entre paires de données appariées par le biais d'un effet aléatoire associé au point d'échantillonnage, semble toutefois encore plus prometteuse. D'une part, le temps de calcul est plus réduit comparé aux méthodes basées sur les permutations, et d'autre part la possibilité de spécifier les effets des variables environnementales par le biais d'une combinaison linéaire offre un cadre beaucoup plus souple pour tester l'influence de ces variables. Proposée par Selkoe *et al.* (2010), cette approche traite chaque point d'échantillonnage comme un effet aléatoire additif s'exprimant sur la distance entre paires de points d'échantillonnage. L'utilisation d'une structure de covariance Toeplitz d'ordre 1 contraint l'ensemble des effets aléatoires à être estimé par le biais d'un seul paramètre de covariance. Bien qu'initialement présentée pour étudier des différentiations génétiques à l'échelle populationnelle (Fst), j'ai employé cette approche pour analyser l'influence des éléments paysagers sur les distances génétiques entre paire d'individus (chapitre 4). Cela m'a ainsi permis d'utiliser les méthodes de sélection de modèles basées sur l'approche de vraisemblance (Burnham & Anderson 2002), afin de mieux ségréger l'influence des différents éléments paysagers.

Un second niveau de non-indépendance des données appariées peut résulter de l'agencement des mesures de distance les unes par rapport aux autres. Dans le cas de l'utilisation d'un graphe de voisinage immédiat (*nearest neighbouring graph*, p. ex. triangulation de Delaunay, minimum-spanning tree ; Arnaud 2003; Goldberg & Waits 2010 ; Figure M6-a), les données appariées ne se « chevauchent pas ». Toutefois, ce type de graphe n'est pas toujours adéquat pour une représentation réaliste des capacités de déplacement du modèle biologique considéré. Lorsque toutes les combinaisons sont prises en compte (absence de graphe de voisinage), ce qui est généralement le cas lors de l'utilisation classique des tests de Mantel (mais voir Arnaud (2003) pour un exemple d'utilisation des tests de Mantel sur données issues de graphes de voisinage immédiat), ou bien lors de l'utilisation d'un graphe de voisinage par la distance (*distance-based neighbouring graph* ; Jombart *et al.* 2008), les données appariées peuvent être chevauchantes, c'est-à-dire qu'un même élément paysager peut se retrouver associé à plusieurs mesures de distances génétiques et paysagères (Figure M6-b). Ce second niveau de non-indépendance des données appariées n'a, à ma connaissance, jamais été pris en compte, mais est également susceptible de perturber la significativité des tests

statistiques. Là encore, l'utilisation des modèles mixtes pourrait être une réponse à ce problème de pseudo-réplication : la part de variabilité expliquée par ce degré de chevauchement entre données appariées pourrait en effet être prise en compte par un second effet aléatoire (codant pour un « effet chevauchement »), complétant la prise en compte des « effets sites ».

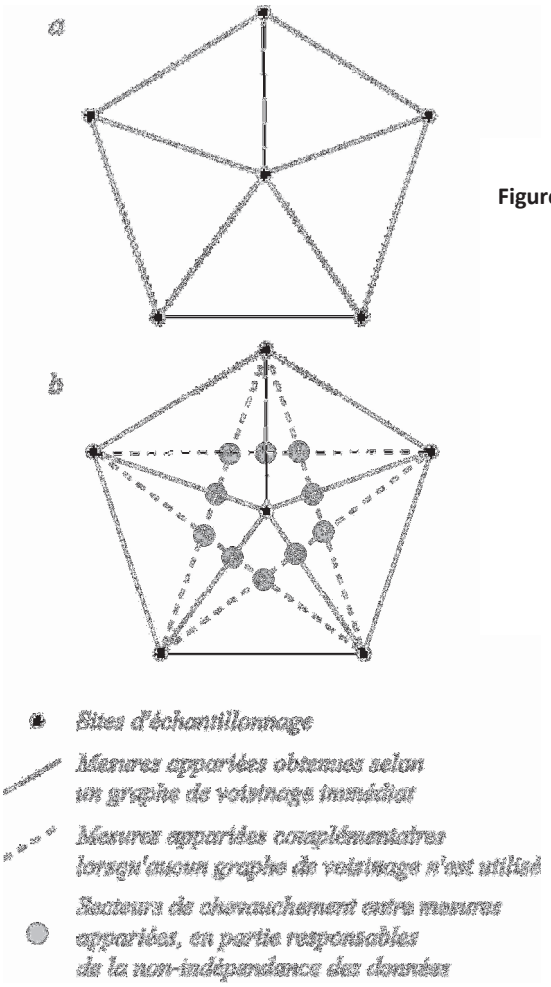


Figure M6 : (a) Données appariées récoltées selon un graphe de voisinage immédiat (triangulation de Delaunay) : bien que les mesures soient non-chevauchantes, un tel graphe n'est pas toujours à même de décrire au mieux le comportement de dispersion de l'espèce-cible.
(b) Toutes les combinaisons de points sont considérées pour l'obtention des données appariées (absence de graphe de voisinage) : les mesures ne sont pas indépendantes car présentent des secteurs communs de chevauchement.

Colinéarité des variables paysagères

Un problème récurrent en écologie du paysage provient de la corrélation positive ou négative qui peut exister entre différents éléments paysagers (Graham 2003; Cushman & Landguth 2010b). C'est le cas par exemple de la topographie et du réseau hydrographique (les cours d'eau étant systématiquement associés aux fonds de vallées), ou encore, comme mis en évidence dans le cadre de mes travaux, de la topographie et de l'occupation du sol : la ligne TGV-PSE longeant les contreforts du massif du Morvan (Chapitre 3) et les zones agricoles se concentrant principalement en zone peu vallonnées (Chapitre 4). Ignorer cette colinéarité entre variables paysagères peut conduire à des erreurs d'interprétation. L'utilisation de tests de Mantel partiels dans le cadre d'une modélisation de type causale peut permettre d'éviter ces erreurs (Cushman & Landguth 2010b). Il est également possible d'identifier les variables colinéaires par le biais d'analyses en composantes principales (ACP ; Graham 2003) et d'intégrer cette non-indépendance dans les analyses statistiques par une utilisation des composantes principales comme variables explicatives (Chapitre 4). Enfin, considérant la flexibilité offerte par l'échantillonnage individu-centré, la meilleure approche pourrait être d'augmenter le nombre de zones d'études indépendantes : les éléments paysagers pouvant présenter des degrés différents de colinéarité entre réplicats, comparer les résultats obtenus en chaque zone d'étude pourrait permettre de ségréger l'importance des différents éléments paysagers sur les flux de gènes (Anderson *et al.* 2010; Fahrig *et al.* 2011). Dans tous les cas, il s'avère indispensable d'intégrer dans les analyses l'ensemble des variables environnementales qui sont susceptibles d'influer sur les flux de gènes de l'espèce considérée.

Vers une meilleure intégration des différentes approches en génétique du paysage

Quelle que soit l'approche génétique utilisée (populationnelle ou individu-centrée), l'objectif généralement recherché est d'inférer des flux démographiques à partir de données génétiques. Néanmoins, il apparait clairement que les données génétiques à elles seules sont insuffisantes pour répondre à cet objectif, et doivent être complétées par des données démographiques. Par ailleurs, s'il est clair que l'approche individu-centré offre de multiples avantages en génétique du paysage, elle ne permet pas d'estimer certains paramètres clef relatifs aux processus génétiques et démographiques, qui sont mieux appréhendés par une approche populationnelle.

Intégration de données démographiques

La connectivité génétique renvoie à la diffusion des allèles dans l'espace (flux géniques) et à leur influence sur les processus évolutifs locaux tels que dépression de consanguinité, dérive génétique ou variation du potentiel adaptatif ; la connectivité démographique fait quant à elle référence aux flux d'individus entre populations (immigration et émigration) et à leur influence sur les paramètres démographiques locaux tels que taux de croissance ou persistance des peuplements (Lowe & Allendorf 2010). Si ces deux processus sont intimement liés, des données génétiques seules n'apportent, dans la plupart des cas, que peu d'information sur la connectivité démographique (Bohonak 1999; Spear *et al.* 2010).

Les taux actuels de dispersion entre peuplements ne peuvent pas toujours être estimés directement par des approches démographiques (capture-marquage-recapture, télémétrie, etc. ; e.g. Schtickzelle *et al.* 2002; Perret *et al.* 2003; Broquet *et al.* 2006a) ou génétiques (méthodes d'assignement ou analyses de lignées, ces deux approches impliquant un échantillonnage le plus exhaustif possible des agrégats ; Jones & Ardren 2003; Paetkau *et al.* 2004; Manel *et al.* 2005; Holderegger & Wagner 2008; Lowe & Allendorf 2010). Dans ce cas, les estimations indirectes de flux de gènes par le biais de mesures de distances génétiques doivent pouvoir être interprétées en termes de flux

démographiques (Palsboll *et al.* 2007; Lowe & Allendorf 2010). Cette interprétation repose sur l'hypothèse suivante : de fortes distances génétiques correspondent à de faibles flux démographiques, et inversement, de faibles distances génétiques correspondent à d'importants flux démographiques (p. ex. Vos *et al.* 2001). Or cette interprétation n'est pas toujours évidente.

Les mesures de différentiation génétique entre populations (p. ex. Fst) reposent sur des modèles démographiques théoriques et sont donc censées permettre de faire un lien direct entre connectivité génétique et connectivité démographique. Pour autant, les nombreuses hypothèses sous-jacentes telles que l'équilibre évolutif entre dérive génétique et flux de gènes ou l'égalité des tailles efficaces de populations, sont rarement valides en conditions naturelles (Waples 1998). Or s'écarter de ces conditions peut conduire à des biais importants dans l'interprétation des flux de gènes (p. ex. Bjorklund *et al.* 2010). Ainsi, dans une dynamique métapopulationnelle, des phénomènes récents de recolonisation (impliquant des effets fondateurs) peuvent par exemple conduire à de faibles mesures de Fst (forte connectivité génétique observée), bien que les flux d'individus soient en réalité très faibles (faible connectivité démographique réalisée ; Sork *et al.* 1999; Marko & Hart 2011). Les mesures de dissimilarité entre individus reposent quant à elles sur une simple approche statistique (p. ex. mesure d'une proportion d'allèles en commun) et s'affranchissent généralement de tout modèle démographique théorique : dans ce contexte, faire le lien entre connectivité génétique et connectivité démographique est tout aussi problématique.

Il est donc indispensable de coupler données génétiques et données démographiques pour s'assurer de l'adéquation entre estimations de flux de gènes et de flux de dispersion et ainsi dégager une vision claire des différents processus affectant réellement les peuplements considérés (Mayer *et al.* 2009; Lowe & Allendorf 2010; Safner *et al.* 2011a; Lamy *et al.* 2012). Si l'idéal reste de pouvoir estimer les flux démographiques réels, ceci implique, encore une fois, la mise en place de procédures particulièrement contraignantes en terme de temps et de moyens (Murphy *et al.* 2010b). L'acquisition de données démographiques plus simples, mais également plus accessibles, peut être pertinente.

Dans le cadre d'un échantillonnage individu-centré, plusieurs types de données démographiques peuvent ainsi être récoltées sans pour autant mettre à mal les différents

avantages méthodologiques (notamment en terme de durée de prospection) qu'offrent cette approche. L'information la plus simple qu'il est possible de récolter lors d'un échantillonnage individu-centré concerne l'occurrence des individus. Elle s'est avérée particulièrement intéressante dans le cas de l'étude de la dispersion du triton alpestre en paysage naturel (Chapitre 4) : si la dispersion semblait plus importante en zone agricole, ce qui va à l'encontre de la littérature (p. ex. Joly *et al.* 2001; Mazerolle & Desrochers 2005; Stevens *et al.* 2006b; Janin *et al.* 2009), la probabilité d'occurrence y était également plus faible. Ceci peut s'expliquer par une plus grande propension à la dispersion des individus lorsque les sites de reproduction sont défavorables, et inversement par une plus grande propension à la philopatrie en zones prairiales, secteurs offrant des sites de reproduction de qualité et stables dans le temps (Semlitsch 2008; Kopecky *et al.* 2012). Il s'agit là d'une parfaite illustration de l'importance des données démographiques pour l'interprétation des données génétiques. Une seconde interprétation, s'affranchissant de l'utilisation des données d'occurrence, repose sur l'hypothèse que les secteurs de pentes pourraient constituer des zones de barrières aux flux de gènes ; il n'est toutefois pas possible de hiérarchiser ces deux hypothèses sans la mise en place d'études complémentaires, par exemple comportementales.

D'autres types de données, permettant de caractériser les processus démographiques en jeu sur la zone d'étude, peuvent être récoltées dans le cadre d'un échantillonnage individu-centré. La durée de prospection nécessaire à la détection ou la capture du petit nombre d'individus que l'on aura décidé d'échantillonner en chaque site peut par exemple servir d'estimation indirecte d'abondance, évitant ainsi une prospection systématique de tous les sites selon un protocole d'échantillonnage standardisé, fiable mais coûteux en temps (Fellers & Freel 1995). Enfin, l'obtention d'indicateurs de l'état physiologique des quelques individus récoltés (mesures de conditions physiques (poids, taille), dosages d'hormones de stress, etc.) peut permettre de se faire une idée de la qualité des sites prospectés ou du statut démographique des peuplements (Sztatecsny & Schabetsberger 2005; Janin *et al.* 2011).

Données d'occurrence, estimations d'abondance ou indicateurs d'état physiologique sont donc autant de données « démographiques » qu'il est aisé de récolter dans le cadre d'un échantillonnage individu-centré. Si elles peuvent permettre d'enrichir les conclusions issues d'analyses corrélatives en génétique du paysage, les combiner

directement aux données génétiques en une seule et même analyse statistique reste sans doute la procédure la plus prometteuse : à ce titre, l'utilité des modèles de gravité, intégrant à la fois des données inter- et intra-sites (Murphy *et al.* 2010a), mérite d'être plus amplement étudiée.

Concilier les avantages des approches individu-centrées et populationnelles

Si l'échantillonnage individu-centré apporte de nombreux avantages méthodologiques et inférentiels en génétique du paysage, mon propos n'est certainement pas de dénigrer l'utilité des approches basées sur un échantillonnage populationnel classique. De nombreux paramètres biologiques, dont la connaissance peut s'avérer indispensable en génétique de la conservation, ne peuvent être réellement appréciés que par le biais de l'étude de la variabilité génétique à l'échelle des populations : estimations de diversité génétique ou de taille efficace de populations, mesures de dépression de consanguinité ou de potentiel adaptatif ou encore détection d'évènements historiques tels qu'effets fondateurs ou *bottlenecks*, sont autant d'informations qu'il peut être indispensable de connaître pour la gestion et le maintien de la viabilité des peuplements (Beebee & Rowe 2001; Frankham *et al.* 2002; Broquet *et al.* 2010; Luquet *et al.* 2011). De la même manière, méthodes d'assignement ou analyses de lignées (Jones & Ardren 2003; Manel *et al.* 2005; Lowe & Allendorf 2010), basées sur un échantillonnage génétique populationnel, peuvent permettre une mesure directe des flux de dispersion entre populations, quand une approche individu-centrée n'en autorise qu'une estimation indirecte. Comme précisé en introduction, les approches bayesiennes et les approches basées sur la théorie de la coalescence, implémentées dans des logiciels comme MIGRATE (Beerli & Felsenstein 2001) ou BAYESASS (Wilson & Rannala 2003), permettent désormais d'obtenir des estimations fiables des flux de dispersion entre populations, tout en s'affranchissant de certaines contraintes liées à l'utilisation de modèles théoriques classiques. Si ces résultats peuvent être utilisés de manière *ad hoc*, par simple comparaison des patrons de dispersion avec les caractéristiques du paysage, ou directement en tant que variables explicatives dans des approches corrélatives, le principal inconvénient de ces méthodes reste que les estimations de flux de dispersion sont uniquement basées sur l'utilisation des données génétiques et ne reposent en rien

sur les caractéristiques du paysage. L'idéal serait sans doute de pouvoir disposer, à l'avenir, d'outils basés sur des approches bayesiennes ou de coalescence, intégrant dans leur processus d'estimation des flux de dispersion des informations sur la configuration du paysage.

Il n'en reste pas moins que les approches populationnelles entraînent une réduction de la couverture d'échantillonnage du paysage par rapport aux approches individu-centrées. Un plan d'échantillonnage idéal consisterait donc en une combinaison de ces deux approches individuelles et populationnelles (Broquet & Petit 2009; Chapitre 1) sous la forme d'un échantillonnage stratifié (Figure M7). Un grand nombre de prélèvements génétiques pourraient tout d'abord être obtenus par un échantillonnage individuel (aléatoire ou ciblé), afin d'assurer une couverture d'échantillonnage optimale de la zone d'étude. Quelques agrégats bien particuliers, choisis en fonction de leurs localisations respectives dans l'espace, pourraient ensuite être échantillonnés de manière plus ou moins exhaustive (Paetkau *et al.* 2004) afin de mettre en œuvre différents types d'analyses nécessitant un échantillonnage populationnel, et apportant des informations complémentaires quant à la perméabilité du paysage aux flux de gènes (Holderegger & Wagner 2008; Anderson *et al.* 2010).

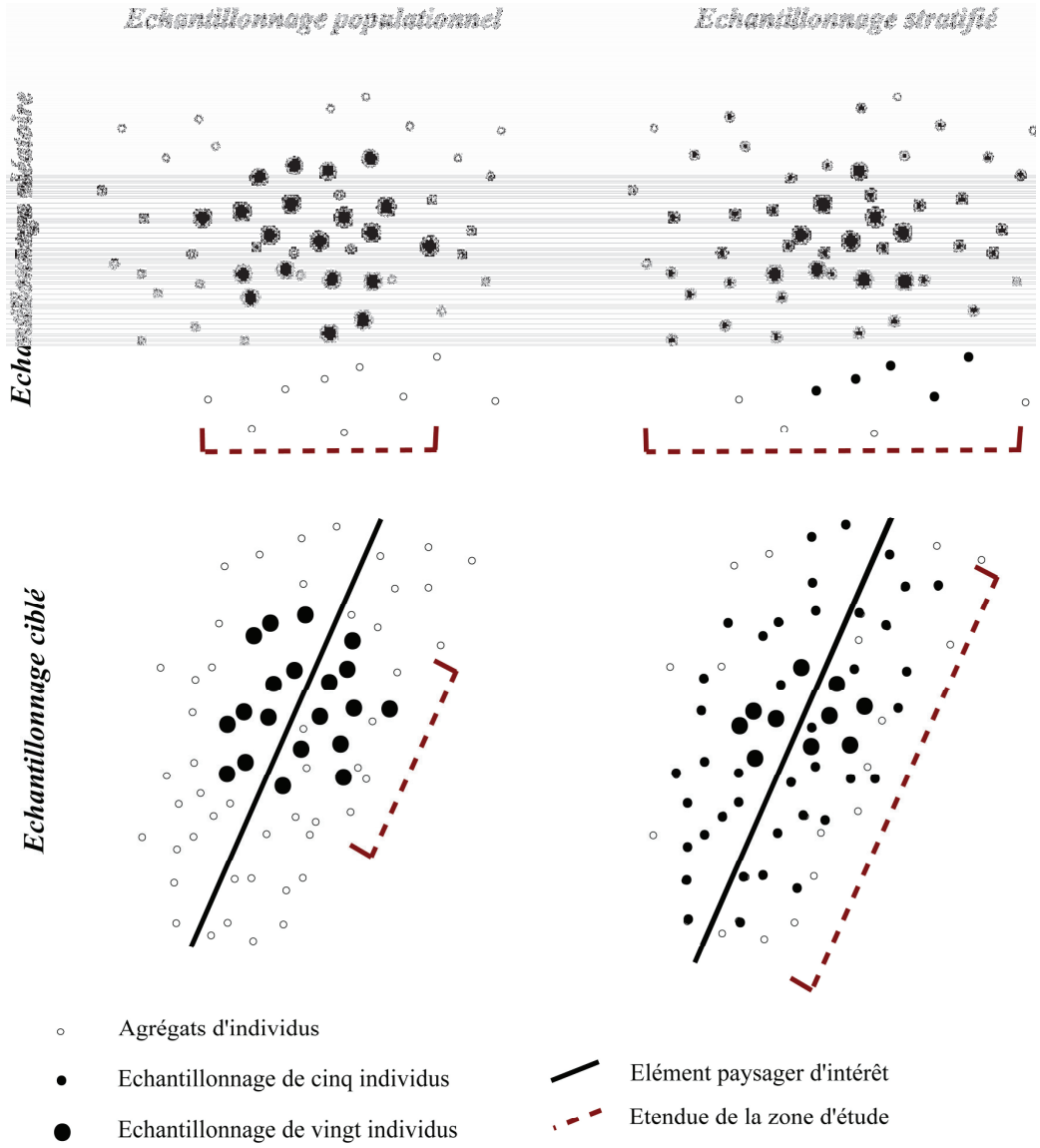


Figure M7 : Comparaison d'un échantillonnage populationnel classique et d'un échantillonnage stratifié, combinant approches individuelle et populationnelle. Dans cet exemple, 400 échantillons peuvent être analysés. L'échantillonnage populationnel porte sur 20 agrégats de 20 individus, tandis que l'échantillonnage stratifié porte sur 10 agrégats de 20 individus et 40 agrégats de 5 individus. L'échantillonnage stratifié permet une meilleure couverture d'échantillonnage de la zone d'étude (échantillonnage aléatoire) ou de l'élément paysager d'intérêt (échantillonnage ciblé) : augmentation du nombre d'agrégats échantillonnés et de la superficie de la zone d'étude, diminution de la distance entre points. Il autorise également l'obtention de données génétiques au niveau d'agrégat pour lesquels les densités d'individus sont trop faibles pour assurer l'obtention de fréquences alléliques non-biaisées (*ghost populations* ; Waples & Gaggiotti 2006), tout en permettant l'utilisation de méthodes issues de la génétiques des populations, toujours plus performantes.

L'approche individu-centrée … au-delà des questions de conservation ?

L'étude des flux de dispersion, ou, d'une manière plus générale, l'étude des caractéristiques démographiques d'un peuplement structuré par le biais d'une approche génétique, repose sur une analyse de la variabilité génétique au niveau de loci neutres, c'est-à-dire non soumis à sélection (Selkoe & Toonen 2006). Pour autant, l'identification de loci sous sélection peut permettre de comprendre les bases génétiques d'adaptations locales ou de processus de spéciation, des facteurs également très importants en biologie de la conservation (Manel *et al.* 2003; Segelbacher *et al.* 2010). L'identification de régions du génome soumises à sélection peut par exemple se faire par la comparaison de la diversité génétique d'un grand nombre de loci : certains présenteront des patrons de variabilité atypiques qui pourront être corrélés à des variables environnementales, et seront donc identifiés comme non neutres (e.g. Joost *et al.* 2007). Dans le cadre de telles études, il est nécessaire de mettre en place un échantillonnage assurant la collecte de matériel génétique dans une gamme de conditions environnementales (diversité de milieux ou de conditions climatiques susceptibles d'être rencontrés par les organismes) la plus large possible : ceci passe là encore par un échantillonnage réalisé à l'échelle de l'individu (Manel *et al.* 2012).

Quelque soit la question soulevée, la génétique du paysage repose donc sur une couverture d'échantillonnage optimale des conditions environnementales susceptibles d'influer sur la variabilité génétique observée (Storfer *et al.* 2007). Tout au long de ce mémoire, je me suis attaché à montrer que l'échantillonnage individu-centré était à même de répondre à cette exigence. Présentant par ailleurs un compromis intéressant entre efficacité de détection des structures génétiques et coût des analyses moléculaires, cette approche offre des avantages méthodologiques susceptibles, à l'avenir, de rendre l'outil génétique particulièrement attractif auprès des décideurs et des professionnels de l'environnement.

BIBLIOGRAPHIE

Abdelkrim J, Robertson BC, Stanton J-AL, Gemmell NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* **46**, 185-191.

Adriaensen F, Chardon JP, De Blust G*, et al.* (2003) The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning* **64**, 233-247.

Allag-Dhuisme F, Amsallem J, Barthod C*, et al.* (2010) Guide méthodologique identifiant les enjeux nationaux et transfrontaliers relatifs à la préservation et à la remise en bon état des continuités écologiques et comportant un volet relatif à l'élaboration des schémas régionaux de cohérence écologique: deuxième document en appui à la mise en œuvre de la Trame verte et bleue en France. *Proposition issue du comité opérationnel Trame verte et bleue. MEEDDM ed.*

Anderson CD (2010) Effects of Movement and Mating Patterns on Gene Flow among Overwintering Hibernacula of the Timber Rattlesnake (*Crotalus horridus*). *Copeia*, 54-61.

Anderson CD, Epperson BK, Fortin MJ*, et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology* **19**, 3565-3575.

Angelone S, Holderegger R (2009) Population genetics suggests effectiveness of habitat connectivity measures for the European tree frog in Switzerland. *Journal of Applied Ecology* **46**, 879-887.

Arens P, Van't Westende W, Bugter R, Smulders MJM, Vosman B (2000) Microsatellite markers for the European tree frog Hyla arborea. *Molecular Ecology* **9**, 1944-1946.

Arnaud JF (2003) Metapopulation genetic structure and migration pathways in the land snail Helix aspersa: influence of landscape heterogeneity. *Landscape Ecology* **18**, 333-346.

Austin JD, Gorman TA, Bishop D (2011) Assessing fine-scale genetic structure and relatedness in the micro-endemic Florida bog frog. *Conservation Genetics* **12**, 833-838.

Baguette M, Petit S, Queva F (2000) Population spatial structure and migration of three butterfly species within the same habitat network: consequences for conservation. *Journal of Applied Ecology* **37**, 100-108.

Baguette M, Schtickzelle N (2003) Local population dynamics are important to the conservation of metapopulations in highly fragmented landscapes. *Journal of Applied Ecology* **40**, 404-412.

Baguette M, Stevens VM (2003) Local populations and metapopulations are both natural and operational categories. *Oikos* **101**, 661-663.

Baguette M (2004) The classical metapopulation theory and the real, natural world: a critical appraisal. *Basic and Applied Ecology* **5**, 213-224.

Balkenhol N, Gugerli F, Cushman SA*, et al.* (2009a) Identifying future research needs in landscape genetics: where to from here? *Landscape Ecology* **24**, 455-463.

Balkenhol N, Waits LP (2009) Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology* **18**, 4151-4164.

Balkenhol N, Waits LP, Dezzani RJ (2009b) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* **32**, 818-830.

Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics* **5**, 251-261.

Beebee T, Rowe G (2001) Application of genetic bottleneck testing to the investigation of amphibian declines: a case study with natterjack toads. *Conservation Biology* **15**, 266-270.

Beebee TJC (2005) Conservation genetics of amphibians. *Heredity* **95**, 423-427.

Beebee TJC, Griffiths RA (2005) The amphibian decline crisis: A watershed for conservation biology? *Biological Conservation* **125**, 271-285.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4563-4568.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.03, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Bender DJ, Fahrig L (2005) Matrix structure obscures the relationship between interpatch movement and patch size and isolation. *Ecology* **86**, 1023-1033.

Bennett AF, Radford JQ, Haslem A (2006) Properties of land mosaics: Implications for nature conservation in agricultural environments. *Biological Conservation* **133**, 250-264.

Berger L, Speare R, Daszak P*, et al.* (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9031-9036.

Berset-Brandli L, Jaquiery J, Broquet T, Perrin N (2008a) Isolation and characterization of microsatellite loci for the European tree frog (Hyla arborea). *Molecular Ecology Resources* **8**, 1095-1097.

Berset-Brandli L, Jaquiery J, Broquet T, Ulrich Y, Perrin N (2008b) Extreme heterochiasmy and nascent sex chromosomes in European tree frogs. *Proceedings of the Royal Society B-Biological Sciences* **275**, 1577-1585.

Bifolchi A, Picard D, Lemaire C, Cormier JP, Pagano A (2010) Evidence of admixture between differentiated genetic pools at a regional scale in an invasive carnivore. *Conservation Genetics* **11**, 1-9.

Bjorklund M, Bergek S, Ranta E, Kaitala V (2010) The effect of local population dynamics on patterns of isolation by distance. *Ecological Informatics* **5**, 167-172.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology* **74**, 21-45.

Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests, Ghent University.

Boorman SA, Levitt PR (1973) Group selection on boundary of a stable population. *Theoretical Population Biology* **4**, 85-128.

Borcard D, Legendre P (2012) Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology* **93**, 1473-1481.

Boudjemadi K, Lecomte J, Clobert J (1999) Influence of connectivity on demography and dispersal in two contrasting habitats: an experimental approach. *Journal of Animal Ecology* **68**, 1207-1224.

Brede EG, Rowe G, Trojanowski J, Beebee TJC (2001) Polymerase chain reaction primers for microsatellite loci in the Common Toad *Bufo bufo*. *Molecular Ecology Notes* **1**, 308-310.

Brisson J, de Blois S, Lavoie C (2010) Roadside as Invasion Pathway for Common Reed (Phragmites australis). *Invasive Plant Science and Management* **3**, 506-514.

Brooker L, Brooker M, Cale P (1999) Animal dispersal in fragmented habitat : measuring habitat connectivity, corridor use, and dispersal mortality. In: *Conservation Ecology*.

Broquet T, Johnson CA, Petit E*, et al.* (2006a) Dispersal and genetic structure in the American marten, Martes americana. *Molecular Ecology* **15**, 1689-1697.

Broquet T, Ray N, Petit E, Fryxell JM, Burel F (2006b) Genetic isolation by distance and landscape connectivity in the American marten (Martes americana). *Landscape Ecology* **21**, 877-889.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2007) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics* **8**, 509-511.

Broquet T, Petit EJ (2009) Molecular Estimation of Dispersal for Ecology and Population Genetics. In: *Annual Review of Ecology Evolution and Systematics*, pp. 193-216. Annual Reviews, Palo Alto.

Broquet T, Angelone S, Jaquiery J*, et al.* (2010) Genetic Bottlenecks Driven by Population Disconnection. *Conservation Biology* **24**, 1596-1605.

Brown GP, Phillips BL, Webb JK, Shine R (2006) Toad on the road: Use of roads as dispersal corridors by cane toads (Bufo marinus) at an invasion front in tropical Australia. *Biological Conservation* **133**, 88-94.

Burnham KP, Anderson DR (2002) *Model selection and multimodel inference - A practical information-theoric approach*, 2nd Edition. Springer-Verlag, New York, New York, USA.

Carr LW, Fahrig L (2001) Effect of road traffic on two amphibian species of differing vagility. *Conservation Biology* **15**, 1071-1078.

Castellano S, Balletto E (2002) Is the partial Mantel test inadequate? *Evolution* **56**, 1871-1873.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics* **19**, 233-257.

Cercueil A, Francois O, Manel S (2007) The Genetical Bandwidth Mapping: A spatial and graphical representation of population genetic structure based on the Wombling method. *Theoretical Population Biology* **71**, 332-341.

Charbonnel N, Angers B, Rasatavonjizay R*, et al.* (2002a) The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. *Molecular Ecology* **11**, 2213-2228.

Charbonnel N, Angers B, Rasatavonjizay R, Bremond P, Jarne P (2002b) Evolutionary aspects of the metapopulation dynamics of *Biomphataria pfeifferi*, the intermediate host of *Schistosoma mansoni*. *Journal of Evolutionary Biology* **15**, 248-261.

Chardon JP, Adriaensen F, Matthysen E (2003) Incorporating landscape elements into a connectivity measure: a case study for the Speckled wood butterfly (Pararge aegeria L.). *Landscape Ecology* **18**, 561-573.

Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* **7**, 747-756.

Corlatti L, Hacklander K, Frey-Roos F (2009) Ability of Wildlife Overpasses to Provide Connectivity and Prevent Genetic Isolation. *Conservation Biology* **23**, 548-556.

Cosentino BJ, Schooley RL, Phillips CA (2011) Connectivity of agroecosystems: dispersal costs can vary among crops. *Landscape Ecology* **26**, 371-379.

Coulon A, Cosson JF, Angibault JM*, et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology* **13**, 2841-2850.

Coulon A, Guillot G, Cosson JF*, et al.* (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology* **15**, 1669-1679.

Crida A, Manel S (2007) WOMBSOFT: an R package that implements the Wombling method to identify genetic boundary. *Molecular Ecology Notes* **7**, 588-591.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation* **128**, 231-240.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist* **168**, 486-499.

Cushman SA, Landguth EL (2010a) Scale dependent inference in landscape genetics. *Landscape Ecology* **25**, 967-979.

Cushman SA, Landguth EL (2010b) Spurious correlations and inference in landscape genetics. *Molecular Ecology* **19**, 3592-3602.

Denoel M, Lehmann A (2006) Multi-scale effect of landscape processes and habitat quality on newt abundance: Implications for conservation. *Biological Conservation* **130**, 495-504.

Denoel M, Ficetola GF (2007) Landscape-level thresholds and newt conservation. *Ecological Applications* **17**, 302-309.

Denoel M, Ficetola GF (2008) Conservation of newt guilds in an agricultural landscape of Belgium: the importance of aquatic and terrestrial habitats. *Aquatic Conservation-Marine and Freshwater Ecosystems* **18**, 714-728.

Dodd CK, Barichivich WJ, Smith LL (2004) Effectiveness of a barrier wall and culverts in reducing wildlife mortality on a heavily traveled highway in Florida. *Biological Conservation* **118**, 619-631.

Donnelly P, Tavare S (1995) Coalescents and genealogical structure under neutrality. *Annual Review of Genetics* **29**, 401-421.

Dormann CF, McPherson JM, Araujo MB*, et al.* (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609-628.

Drescher K, Henderson J, McNamara K (2001) Farmland Prices Determinants

Driezen K, Adriaensen F, Rondinini C, Doncaster CP, Matthysen E (2007) Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (Erinaceus europaeus). *Ecological Modelling* **209**, 314-322.

Dunning JB, Danielson BJ, Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* **65**, 169-175.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571-2581.

Dyer RJ, Nason JD, Garrick RC (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology* **19**, 3746-3759.

Efron B, Tibshirani R (1993) *An Introduction to the Bootstrap* Chapman & Hall, New York, London.

Elmer KR, Davila JA, Lougheed SC (2007) Applying new inter-individual approaches to assess fine-scale population genetic diversity in a neotropical frog, Eleutherodactylus ockendeni. *Heredity* **99**, 506-515.

Emaresi G, Pellet J, Dubey S, Hirzel A, Fumagalli L (2011) Landscape genetics of the Alpine newt (Mesotriton alpestris) inferred from a strip-based approach. *Conservation Genetics* **12**, 41-50.

Epperson BK (2003a) *Geographical Genetics* Princeton University Press, Princeton and Oxford.

Epperson BK (2003b) *Geographical Genetics*, Princeton University Press.

Epperson BK, McRae BH, Scribner K*, et al.* (2010) Utility of computer simulations in landscape genetics. *Molecular Ecology* **19**, 3549-3564.

Epps CW, Palsboll PJ, Wehausen JD*, et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters* **8**, 1029-1038.

Fagan WF, Holmes EE (2006) Quantifying the extinction vortex. *Ecology Letters* **9**, 51-60.

Fahrig L, Merriam G (1994) Conservation of Fragmented Populations. *Conservation Biology* **8**, 50-59.

Fahrig L, Pedlar JH, Pope SE, Taylor PD, Wegner JF (1995) Effect of Road Traffic on Amphibian Density. *Biological Conservation* **73**, 177-182.

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

Fahrig L, Baudry J, Brotons L*, et al.* (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. *Ecology Letters* **14**, 101-112.

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**, 92-94.

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology* **16**, 1149-1166.

Fellers GM, Freel KL (1995) Technical Report NPS/WRUC/NRTR-95-01: A standardized protocol for surveying aquatic amphibians (ed. USDI), p. 122, Davis, California. .

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography* **16**, 265-280.

Forman RTT (2000) Estimate of the area affected ecologically by the road system in the United States. *Conservation Biology* **14**, 31-35.

Francois O, Ancelet S, Guillot G (2006) Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* **174**, 805-816.

Frank K, Wissel C (2002) A formula for the mean lifetime of metapopulations in heterogeneous landscapes. *American Naturalist* **159**, 530-552.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to conservation genetics.* Cambridge University Press, Cambridge.

Frankham R (2005) Genetics and extinction. *Biological Conservation* **126**, 131-140.

Funk WC, Blouin MS, Corn PS*, et al.* (2005a) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-496.

Funk WC, Donnelly MA, Lips KR (2005b) Alternative views of amphibian toe-clipping. *Nature* **433**, 193-193.

Garner TWJ, Schmidt BR, Hoeck P, Van Buskirk J (2003) Di- and tetranucleotide microsatellite markers for the Alpine newt (*Triturus alpestris*): characterization and cross-priming in five congeners. *Molecular Ecology Notes* **3**, 186-188.

Gauffre B, Estoup A, Bretagnolle V, Cosson JF (2008) Spatial genetic structure of a small rodent in a heterogeneous landscape. *Molecular Ecology* **17**, 4619-4629.

Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of extinction. In: *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soulé ME), pp. 19–34. Sinauer Associates, MA, Sunderland.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology* **19**, 3650-3663.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). University of Lausanne, Dorigny.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11**, 1103-1114.

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809-2815.

Grilo C, Bissonette JA, Santos-Reis M (2009) Spatial-temporal patterns in Mediterranean carnivore road casualties: Consequences for mitigation. *Biological Conservation* **142**, 301-313.

Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation-the correlated allele frequencies model revisited. *Bioinformatics* **24**, 2222-2228.

Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Molecular Ecology* **18**, 4734-4756.

Haila Y (2002) A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. *Ecological Applications* **12**, 321-334.

Hanski I, Gilpin M (1991) Metapopulation dynamics : brief-history and conceptual domain. *Biological Journal of the Linnean Society* **42**, 3-16.

Hanski I (1998) Metapopulation dynamics. *Nature* **396**, 41-49.

Hanski I (1999) *Metapopulation Ecology* Oxford University Press, Oxford.

Hanski I, Gaggiotti OE (2004) *Ecology, Genetics, and Evolution of Metapopulations.* Academic Press, San Diego, USA.

Harrison S (1991) Local extinction in a metapopulation context : an empirical evaluation. *Biological Journal of the Linnean Society* **42**, 73-88.

Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution* **16**, 629-636.

Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* **157**, 331-341.

Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* **58**, 199-207.

Holderegger R, Di Giulio M (2010) The genetic effects of roads: A review of empirical evidence. *Basic and Applied Ecology* **11**, 522-531.

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting Fst. *Nature Reviews Genetics* **10**, 639-650.

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.

Hyatt AD, Boyle DG, Olsen V*, et al.* (2007) Diagnostic assays and sampling protocols for the detection of Batrachochytrium dendrobatidis. *Diseases of Aquatic Organisms* **73**, 175-192.

Janin A, Lena JP, Ray N*, et al.* (2009) Assessing landscape connectivity with calibrated cost-distance modelling: predicting common toad distribution in a context of spreading agriculture. *Journal of Applied Ecology* **46**, 833-841.

Janin A, Lena JP, Joly P (2011) Beyond occurrence: Body condition and stress hormone as integrative indicators of habitat availability and fragmentation in the common toad. *Biological Conservation* **144**, 1008-1016.

Janin A, Léna JP, Deblois S, Joly P (2012) Use of stress-hormone levels and habitat selection to assess functional connectivity of a landscape for an amphibian. *Biological Conservation* **26**, 923–931.

Jaquiery J, Broquet T, Hirzel AH, Yearsley J, Perrin N (2011) Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology* **20**, 692-705.

Jaquiéry J, Broquet T, Hirzel AH, Yearsley J, Perrin N (2011) Inferring landscape effects on dispersal from genetic distances: how far can we go? *Molecular Ecology* **20**, 692-705.

Jehle R, Arntzen JW (2000) Post-breeding migrations of newts (Triturus cristatus and T. marmoratus) with contrasting ecological requirements. *Journal of Zoology* **251**, 297-306.

Jehle R, Burke T, Arntzen JW (2005) Delineating fine-scale genetic units in amphibians: Probing the primacy of ponds. *Conservation Genetics* **6**, 227-234.

Jodoin Y, Lavoie C, Villeneuve P*, et al.* (2008) Highways as corridors and habitats for the invasive common reed *Phragmites australis* in Quebec, Canada. *Journal of Applied Ecology* **45**, 459-466.

Joly P, Miaud C (1989) Fidelity to the breeding site in the alpine newt *Triturus alpestris*. *Behavioural Processes* **19**, 47-56.

Joly P, Miaud C (1993) How does a newt find its pond - the role of chemical cues in migrating newts (*Triturus alpestris*). *Ethology Ecology & Evolution* **5**, 447-455.

Joly P, Grolet O (1996) Colonization dynamics of new ponds, and the age structure of colonizing Alpine newts, *Triturus alpestris*. *Acta Oecologica* **17**, 599-608.

Joly P, Miaud C, Lehmann A, Grolet O (2001) Habitat matrix effects on pond occupancy in newts. *Conservation Biology* **15**, 239-248.

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

Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. *Molecular Ecology* **12**, 2511-2523.

Joost S, Bonin A, Bruford MW*, et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology* **16**, 3955-3969.

Kalinowski ST (2002) Evolutionary and statistical properties of three genetic distances. *Molecular Ecology* **11**, 1263-1273.

Kalinowski ST (2005) Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* **94**, 33-36.

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity* **101**, 423-436.

Kerth G, Melber M (2009) Species-specific barrier effects of a motorway on the habitat use of two threatened forest-living bat species. *Biological Conservation* **142**, 270-279.

Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **85**, 1049-1064.

Kimura M (1953) «Stepping stone» model of populations. *Annual Report National Institute of Genetics, Japan* **3**, 62-63.

Kindlmann P, Burel F (2008) Connectivity measures: a review. *Landscape Ecology* **23**, 879-890.

Kingman JFC (2000) Origins of the coalescent: 1974-1982. *Genetics* **156**, 1461-1463.

Kirchner F, Ferdy JB, Andalo C, Colas B, Moret J (2003) Role of corridors in plant dispersal: an example with the endangered Ranunculus nodiflorus. *Conservation Biology* **17**, 401-410.

Kopecky O, Vojar J, Denoel M (2012) Sex-specific effect of pool desiccation on the movement of Alpine Newts, *Mesotriton alpestris* (LAURENTI, 1768), among breeding sites (Caudata: Salamandridae). *Herpetozoa* **24**, 127-134.

Kovar R, Brabec M, Vita R, Bocek R (2009) Spring migration distances of some Central European amphibian species. *Amphibia-Reptilia* **30**, 367-378.

Kriger KM, Hines HB, Hyatt AD, Boyle DG, Hero JM (2006) Techniques for detecting chytridiomycosis in wild frogs: comparing histology with real-time Taqman PCR. *Diseases of Aquatic Organisms* **71**, 141-148.

Kriger KM, Pereoglou F, Hero JM (2007) Latitudinal variation in the prevalence and intensity of chytrid (Batrachochytrium dendrobatidis) infection in Eastern Australia. *Conservation Biology* **21**, 1280-1290.

Kuehn R, Hindenlang KE, Holzgang O*, et al.* (2007) Genetic effect of transportation infrastructure on roe deer Populations (Capreolus capreolus). *Journal of Heredity* **98**, 13-22.

Lamy T, Pointier JP, Jarne P, David P (2012) Testing metapopulation dynamics using genetic, demographic and ecological data. *Molecular Ecology* **21**, 1394-1410.

Landguth EL, Cushman SA (2010) CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources* **10**, 156-161.

Landguth EL, Cushman SA, Schwartz MK*, et al.* (2010) Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**, 4179-4191.

Landguth EL, Fedy BC, Oyler-McCance SJ*, et al.* (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources* **12**, 276-284.

Latter BDH (1973) Island model of population differentiation - General solution. *Genetics* **73**, 147-157.

Legendre L, Legendre P (1998) *Numerical Ecology* 2nd English ed. Elsevier., Amsterdam.

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources* **10**, 831-844.

Lesbarreres D, Lode T, Merila J (2004) What type of amphibian tunnel could reduce road kills? *Oryx* **38**, 220-223.

Lesbarreres D, Primmer CR, Lode T, Merila J (2006) The effects of 20 years of highway presence on the genetic structure of Rana dalmatina populations. *Ecoscience* **13**, 531-538.

Lesbarreres D, Fahrig L (2012) Measures to reduce population fragmentation by roads: what has worked and how do we know? *Trends in ecology & evolution (Personal edition)*.

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* **15**, 237–240.

Lewis OT, Thomas CD, Hill JK*, et al.* (1997) Three ways of assessing metapopulation structure in the butterfly Plebejus argus. *Ecological Entomology* **22**, 283-293.

Liebgold EB, Brodie ED, Cabe PR (2011) Female philopatry and male-biased dispersal in a direct-developing salamander, *Plethodon cinereus*. *Molecular Ecology* **20**, 249-257.

Lowe A, Harris S, Ashton P (2004) *Ecological genetics: Design, analysis and application.* Blackwell Publishing.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038-3051.

Luquet E, Lena JP, David P*, et al.* (2011) Consequences of genetic erosion on fitness and phenotypic plasticity in European tree frog populations (Hyla arborea). *Journal of Evolutionary Biology* **24**, 99-110.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189-197.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate techniques. *Trends in Ecology & Evolution* **20**, 136-142.

Manel S, Berthoud F, Bellemain E*, et al.* (2007) A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology* **16**, 2031-2043.

Manel S, Conord C, Despres L (2009) Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. *Bmc Evolutionary Biology* **9**.

Manel S, Joost S, Epperson BK*, et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology* **19**, 3760-3772.

Manel S, Albert CH, Yoccoz NG (2012) Sampling in Landscape Genomics. In: *Data Production and Analysis in Population Genomics: Methods and Protocols* (eds. Pompanon F, Bonin A). Springer Science+Business Media, New york.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**, 173-190.

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

Marsh DM, Trenham PC (2001) Metapopulation dynamics and amphibian conservation. *Conservation Biology* **15**, 40-49.

Mata C, Hervas I, Herranz J, Suarez F, Malo JE (2008) Are motorway wildlife passages worth building? Vertebrate use of road-crossing structures on a Spanish motorway. *Journal of Environmental Management* **88**, 407-415.

Mayer C, Schiegg K, Pasinelli G (2009) Patchy population structure in a short-distance migrant: evidence from genetic and demographic data. *Molecular Ecology* **18**, 2353-2364.

Mazerolle MJ, Desrochers A (2005) Landscape resistance to frog movements. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **83**, 455-464.

McCarthy MA, Parris KM (2004) Clarifying the effect of toe clipping on frogs with Bayesian statistics. *Journal of Applied Ecology* **41**, 780-786.

McGregor RL, Bender DJ, Fahrig L (2008) Do small mammals avoid roads because of the traffic? *Journal of Applied Ecology* **45**, 117-123.

McGuigan K, McDonald K, Parris K, Moritz C (1998) Mitochondrial DNA diversity and historical biogeography of a wet forest-restricted frog (Litoria pearsoniana) from mid-east Australia. *Molecular Ecology* **7**, 175-186.

Mech SG, Hallett JG (2001) Evaluating the effectiveness of corridors: a genetic approach. *Conservation Biology* **15**, 467-474.

Meunier FD, Verheyden C, Jouventin P (1999) Bird communities of highway verges: Influence of adjacent habitat and roadside management. *Acta Oecologica-International Journal of Ecology* **20**, 1-13.

Miaud C, Guyetant R, Faber H (2000) Age, size, and growth of the alpine newt, *Triturus alpestris* (Urodela : Salamandridae), at high altitude and a review of life-history trait variation throughout its range. *Herpetologica* **56**, 135-144.

Miller MP (2005) Alleles In Space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* **96**, 722-724.

Millions DG, Swanson BJ (2007) Impact of natural and artificial barriers to dispersal on the population structure of bobcats. *Journal of Wildlife Management* **71**, 96-102.

Monmonier M (1973) Maximum-difference barriers: An alternative numerical regionalization method. *Geographical Analysis* **5**, 245-261.

Morin PA, Luikart G, Wayne RK, Grp SNPW (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19**, 208-216.

Mousson L, Neve G, Baguette M (1999) Metapopulation structure and conservation of the cranberry fritillary Boloria aquilonaris is (lepidoptera, nymphalidae) in Belgium. *Biological Conservation* **87**, 285-293.

Murphy MA, Dezzani R, Pilliod DS, Storfer A (2010a) Landscape genetics of high mountain frog metapopulations. *Molecular Ecology* **19**, 3634-3649.

Murphy MA, Evans JS, Storfer A (2010b) Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology* **91**, 252-261.

Nathan R, Getz WM, Revilla E*, et al.* (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19052-19059.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Nurnberger B, Hofman S, Forg-Brey B*, et al.* (2003) A linkage map for the hybridising toads Bombina bombina and B-variegata (Anura : Discoglossidae). *Heredity* **91**, 136-142.

Olivieri I, Michalakis Y, Gouyon PH (1995) Metapopulation genetics and the evolution of dispersal. *American Naturalist* **146**, 202-228.

Olsson MPO, Widen P, Larkin JL (2008) Effectiveness of a highway overpass to promote landscape connectivity and movement of moose and roe deer in Sweden. *Landscape and Urban Planning* **85**, 133-139.

Pabijan M, Babik W, Rafinski J (2005) Conservation units in north-eastern populations of the Alpine newt (Triturus alpestris). *Conservation Genetics* **6**, 307-312.

Pabijan M, Babik W (2006) Genetic structure in northeastern populations of the Alpine newt (Triturus alpestris): evidence for post-Pleistocene differentiation. *Molecular Ecology* **15**, 2397-2407.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65.

Palo JU, Lesbarreres D, Schmeller DS, Primmer CR, Merila J (2004) Microsatellite marker data suggest sex-biased dispersal in the common frog *Rana temporaria*. *Molecular Ecology* **13**, 2865-2869.

Palomares F, Godoy JA, Lopez-Bao JV*, et al.* (2012) Possible Extinction Vortex for a Population of Iberian Lynx on the Verge of Extirpation. *Conservation Biology* **26**, 689-697.

Palsboll PJ, Berube M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends in Ecology & Evolution* **22**, 11-16.

Pannell JR, Charlesworth B (2000) Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society B-Biological Sciences* **355**, 1851-1864.

Parris KM, McCall SC, McCarthy MA*, et al.* (2010) Assessing ethical trade-offs in ecological field studies. *Journal of Applied Ecology* **47**, 227-234.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* **30**, 683-691.

Pellet J, Guisan A, Perrin N (2004) A concentric analysis of the impact of urbanization on the threatened European tree frog in an agricultural landscape. *Conservation Biology* **18**, 1599-1606.

Perret N, Pradel R, Miaud C, Grolet O, Joly P (2003) Transience, dispersal and survival rates in newt patchy populations. *Journal of Animal Ecology* **72**, 567-575.

Perry JC, Rowe L (2011) Rapid Microsatellite Development for Water Striders by Next-Generation Sequencing. *Journal of Heredity* **102**, 125-129.

Pidancier N, Miquel C, Miaud C (2003) Buccal swabs as a non-destructive tissue sampling method for DNA analysis in amphibians. *Herpetological Journal* **13**, 175-178.

Pinto N, Keitt TH (2009) Beyond the least-cost path: evaluating corridor redundancy using a graph-theoretic approach. *Landscape Ecology* **24**, 253-266.

Pinton F, Le Caro Y (2008) Protection de l'eau et remembrement en Puisaye-Forterre. *Programme Dialog. AgroPariTech and Université Rennes 2.* *www.afip.asso.fr/IMG/pdf/Monographie_Puisaye_Forterre.pdf*.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics* **6**, 847-859.

Poschadel JR, Moller D (2004) A versatile field method for tissue sampling on small reptiles and amphibians, applied to pond turtles, newts, frogs and toads. *Conservation Genetics* **5**, 865-867.

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

Prunier J, Kaufmann B, Grolet O*, et al.* (2012) Skin swabbing as a new efficient DNA sampling technique in amphibians, and 14 new microsatellite markers in the alpine newt (Ichthyosaura alpestris). *Molecular Ecology Resources* **12**, 524-531.

Prunier J, Kaufmann B, Fenet S*, et al.* (*submitted-a*) Improving statistical power in landscape genetic studies: Individual-based versus population-based sampling scheme.

Prunier J, Kaufmann B, Lena JP, Pompanon F, Joly P (*submitted-b*) Permeability of large transport infrastructures to gene flow in the alpine newt *Ichthyosaura alpestris*.

R Development Core Team (2011) R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Raufaste N, Rousset F (2001) Are partial Mantel tests adequate ? *Evolution* **55**, 1703-1705.

Ray N, Lehmann A, Joly P (2002) Modeling spatial distribution of amphibian populations: a GIS approach based on habitat matrix permeability. *Biodiversity and Conservation* **11**, 2143-2165.

Revilla E, Wiegand T, Palomares F, Ferreras P, Delibes M (2004) Effects of matrix heterogeneity on animal dispersal: From individual behavior to metapopulation-level parameters. *American Naturalist* **164**, E130-E153.

Rice WR (1989) Analysing tables of statistical tests. *Evolution* **43**, 223-225.

Ricketts TH (2001) The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist* **158**, 87-99.

Riley SPD, Pollinger JP, Sauvajot RM*, et al.* (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* **15**, 1733-1741.

Rittenhouse TAG, Semlitsch RD, Thompson FR (2009) Survival costs associated with wood frog breeding migrations: effects of timber harvest and drought. *Ecology* **90**, 1620-1630.

Roach JL, Stapp P, Van Horne B, Antolin MF (2001) Genetic structure of a metapopulation of black-tailed prairie dogs. *Journal of Mammalogy* **82**, 946-959.

Rosenberg DK, Noon BR, Megahan JW, Meslow EC (1998) Compensatory behavior of *Ensatina eschscholtzii* in biological corridors: a field experiment. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **76**, 117-133.

Rothermel BB, Semlitsch RD (2002) An experimental investigation of landscape resistance of forest versus old-field habitats to emigrating juvenile amphibians. *Conservation Biology* **16**, 1324-1332.

Rothermel BB (2004) Migratory success of juveniles: A potential constraint on connectivity for pond-breeding amphibians. *Ecological Applications* **14**, 1535-1546.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1-14.

Safner T, Miaud C, Gaggiotti O*, et al.* (2011a) Combining demography and genetic analysis to assess the population structure of an amphibian in a human-dominated landscape. *Conservation Genetics* **12**, 161-173.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011b) Comparison of Bayesian Clustering and Edge Detection Methods for Inferring Boundaries in Landscape Genetics. *International Journal of Molecular Sciences* **12**, 865-889.

Sala OE, Chapin FS, Armesto JJ*, et al.* (2000) Biodiversity - Global biodiversity scenarios for the year 2100. *Science* **287**, 1770-1774.

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103-113.

Schtickzelle N, Le Boulenge E, Baguette M (2002) Metapopulation dynamics of the bog fritillary butterfly: demographic processes in a patchy population. *Oikos* **97**, 349-360.

Schtickzelle N, Baguette M (2003) Behavioural responses to habitat patch boundaries restrict dispersal and generate emigration-patch area relationships in fragmented landscapes. *Journal of Animal Ecology* **72**, 533-545.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441-452.

Segelbacher G, Cushman SA, Epperson BK*, et al.* (2010) Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics* **11**, 375-385.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Selkoe KA, Watson JR, White C*, et al.* (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* **19**, 3708-3726.

Semlitsch RD (2008) Differentiating migration and dispersal processes for pond-breeding amphibians. *Journal of Wildlife Management* **72**, 260-267.

Seppa P, Laurila A (1999) Genetic structure of island populations of the anurans Rana temporaria and Bufo bufo. *Heredity* **82**, 309-317.

Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology* **12**, 253-262.

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

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. *Evolution* **47**, 264-279.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology* **14**, 67-73.

Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography* **28**, 110-128.

Smouse PE, Long JC, Sokal RR (1986) Multiple-regression and correlation extensions of the mantel test of matrix correspondence *Systematic Zoology* **35**, 627-632.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Snell C, Evans IH (2006) Discrimination of moor frog (Rana arvalis) and common frog (Rana temporaria) individuals using a RAPD technique. *Herpetological Journal* **16**, 363-369.

Sordello R, Comolet-Tirman J, De Massary J-C*, et al.* (2011) Trame verte et bleue : Critères nationaux de cohérence / Contribution à la définition du critère sur les espèces. . *Convention MNHN/MEDDTL fiche 3i, Rapport SPN 2011 - 21*.

Sork VL, Nason J, Campbell DR, Fernandez JF (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology & Evolution* **14**, 219-224.

Soto-Azat C, Clarke BT, Fisher MC, Walker SF, Cunningham AA (2009) Non-invasive sampling methods for the detection of Batrachochytrium dendrobatidis in archived amphibians. *Diseases of Aquatic Organisms* **84**, 163-166.

Spear SF, Peterson CR, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (Ambystoma tigrinum melanostictum). *Molecular Ecology* **14**, 2553-2564.

Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology* **19**, 3576-3591.

Stauffer D, Aharony A (1991) *Introduction to percolation theory.* 2nd edn. Taylor and Francis, London, UK.

Steinfartz S, Kusters D, Tautz D (2004) Isolation and characterization of polymorphic tetranucleotide microsatellite loci in the Fire salamander *Salamandra salamandra* (Amphibia: Caudata). *Molecular Ecology Notes* **4**, 626-628.

Stevens VM, Polus E, Wesselingh RA, Schtickzelle N, Baguette M (2004) Quantifying functional connectivity: experimental evidence for patch-specific resistance in the Natterjack toad (Bufo calamita). *Landscape Ecology* **19**, 829-842.

Stevens VM, Leboulenge E, Wesselingh RA, Baguette M (2006a) Quantifying functional connectivity: experimental assessment of boundary permeability for the natterjack toad (Bufo calamita). *Oecologia* **150**, 161-171.

Stevens VM, Verkenne C, Vandewoestijne S, Wesselingh RA, Baguette M (2006b) Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology* **15**, 2333-2344.

Storfer A, Murphy MA, Evans JS*, et al.* (2007) Putting the 'landscape' in landscape genetics. *Heredity* **98**, 128-142.

Storfer A, Eastman JM, Spear SF (2009) Modern Molecular Methods for Amphibian Conservation. *Bioscience* **59**, 559-571.

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

Stuart SN, Chanson JS, Cox NA*, et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* **306**, 1783-1786.

Swihart RK, Atwood TC, Goheen JR*, et al.* (2003) Patch occupancy of North American mammals: is patchiness in the eye of the beholder? *Journal of Biogeography* **30**, 1259-1279.

Sztatecsny M, Schabetsberger R (2005) Into thin air: vertical migration, body condition, and quality of terrestrial habitats of alpine common toads, *Bufo bufo*. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **83**, 788-796.

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. *Oikos* **68**, 571-573.

Tikka PM, Hogmander H, Koski PS (2001) Road and railway verges serve as dispersal corridors for grassland plants. *Landscape Ecology* **16**, 659-666.

Tischendorf L, Fahrig L (2000a) On the usage and measurement of landscape connectivity. *Oikos* **90**, 7-19.

Tischendorf L, Fahrig L (2000b) How should we measure landscape connectivity? *Landscape Ecology* **15**, 633-641.

Trenham PC, Koenig WD, Shaffer HB (2001) Spatially autocorrelated demography and interpond dispersal in the salamander Ambystoma californiense. *Ecology* **82**, 3519-3530.

Trombulak SC, Frissell CA (2000) Review of ecological effects of roads on terrestrial and aquatic communities. *Conservation Biology* **14**, 18-30.

Turner MG, Gardner RH, O'Neill RV (2001) Landscape Ecology in Theory and Practice (ed. Springer-Verlag).

Van Dyck H, Baguette M (2005) Dispersal behaviour in fragmented landscapes: Routine or special movements? *Basic and Applied Ecology* **6**, 535-545.

van Nouhuys S (2009) Metapopulation ecology. In: *Encyclopedia of Life Sciences* (ed. Sons JW), Chichester.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Vandewoestijne S, Baguette M (2004) Demographic versus genetic dispersal measures. *Population Ecology* **46**, 281-285.

Vaughan IP, Ormerod SJ (2005) Increasing the value of principal components analysis for simplifying ecological data: a case study with rivers and river birds. *Journal of Applied Ecology* **42**, 487-497.

Vilter A, Vilter V (1962) Migration de reproduction chez le Triton alpestre des Alpes vaudoises. *Comptes Rendus de l'Académie des Sciences* **156**, 2005-2007.

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.

Wagner A, Schabetsberger R, Sztatecsny M, Kaiser R (2011) Skeletochronology of phalanges underestimates the true age of long-lived Alpine newts (*Ichthyosaura alpestris*). *Herpetological Journal* **21**, 145-148.

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (Bufo exsul). *Molecular Ecology* **18**, 3847-3856.

Wang YH, Yang KC, Bridgman CL, Lin LK (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology* **23**, 989-1000.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**, 438-450.

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.

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

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. *Genetics* **146**, 427-441.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F-ST not equal 1/(4Nm+1). *Heredity* **82**, 117-125.

Wiens JA (1995) Habitat fragmentation - island v landscape perspectives on bird conservation. *Ibis* **137**, S97-S104.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

With KA (1997) The application of neutral landscape models in conservation biology. *Conservation Biology* **11**, 1069-1080.

Woltz HW, Gibbs JP, Ducey PK (2008) Road crossing structures for amphibians and reptiles: Informing design through behavioral analysis. *Biological Conservation* **141**, 2745-2750.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 0097-0159.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.

