Université de Lille

Evo-Eco-Paléo

Université de Lille - Sciences et Technologies

École Doctorale Sciences de la Matière du Rayonnement et de l'Environnement

Thèse de Doctorat pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LILLE SCIENCES ET TECHNOLOGIES

Discipline

BIOLOGIE ÉVOLUTIVE ET ÉCOLOGIE

DÉPARTEMENT DE BIOLOGIE

HISTOIRE ÉVOLUTIVE DE DEUX ESPÈCES D'AMPHIBIENS PIONNIÈRES, LE PÉLODYTE PONCTUÉ ET LE CRAPAUD CALAMITE, EN MILIEU FORTEMENT ANTHROPISÉ

Soutenue publiquement le 19 Décembre 2017 par

Leslie FAUCHER

Philippe JARNE Directeur de recherche, UMR-CNRS 5175, *Montpellier*	Rapporteur
Pierre JOLY Professeur, UMR-CNRS 5023 – LEHNA, *Université de Lyon*	Rapporteur
Virginie STEVENS Chargée de recherche, UMR-CNRS. 5321 'SETE', *Moulis*	Examinatrice
Pierre TABERLET Directeur de recherche, UMR UGA-USMB-CNRS 5553, *Grenoble*	Examinateur
Joël CUGUEN Professeur, UMR-CNRS 8198, *Université de Lille*	Examinateur
José GODIN Président du Groupe Ornithologique et Naturaliste du N-PdC	Invité
Jean-François ARNAUD Professeur, UMR-CNRS 8198, *Université de Lille*	Directeur de thèse

À mes grands-mères,
à mes grands-pères

© Laurent Lamacz

Photographie des terrils des fosses 11/19

Remerciements

Cette thèse aura avant tout été un projet riche en rencontres et en échanges portés par l'histoire de ces drôles de crapauds installés sur les terrils. De la définition des objectifs et des questions du projet à la rédaction des chapitres en passant par l'échantillonnage sur le terrain, le travail de biologie moléculaire en laboratoire et les analyses statistiques, de nombreuses personnes m'ont aidée à avancer et sans elles rien de tout ce qui est présenté ici n'aurait pu être fait. De manière générale, la bienveillance et l'enthousiasme de tous ceux qui ont entouré de près ou de loin ce travail de quatre années a été une grande source de motivation. Trouver les mots justes pour les remercier n'est pas une mince affaire. Ces quelques lignes me paraissent encore bien légères pour vous manifester toute la gratitude que j'ai envers vous.

****

*Pour commencer, mille mercis à **Jean-François**. Merci de m'avoir fait confiance pour travailler sur ce beau projet jusqu'à me donner l'opportunité de poursuivre sur une quatrième année. Merci pour ton implication tout au long de cette thèse. Merci car si la thèse a été un travail aussi plaisant et passionnant c'est parce que ton projet avait été élaboré consciencieusement, grâce aussi à ta pédagogie, ta patience et cette sérénité quotidienne qui m'auront permis d'avancer en m'épargnant pas mal de stress. Je ne doute pas de la chance immense que j'ai eu d'avoir un directeur de thèse aussi disponible, sincère et plein d'humanité.*

*Un grand merci au membre du jury qui se sont déplacés pour assister à ma soutenance et ont accepté d'évaluer mes travaux, **Philippe Jarne**, **Pierre Joly**, **Virginie Stevens**, **Pierre Taberlet**, et **Joël Cuguen**. Ça a été un réel plaisir d'avoir pu échanger avec vous à l'issue de ces quatre ans de travail. Je remercie aussi chaleureusement **José Godin**, qui n'a pas pu se joindre à nous à l'occasion de ma soutenance mais a eu la gentillesse de nous envoyer ses retours sur le manuscrit dans une période difficile pour lui.*

*Merci à **Aurélie Coulon**, **Julie Jaquiéry, Éric Petit**, **Stéphanie Rondel**, et **Cédric Vanappelghem** qui ont participé avec José Godin, à mon comité de thèse. En particulier, merci à **Cédric, Stéphanie, José, et Julie** qui étaient là dès les premières heures du projet. Merci pour vos conseils et le vif intérêt que vous avez porté à ce projet.*

*Un grand merci à Julie et à **Thomas Broquet** pour leur collaboration sur les différents chapitres de ma thèse. Merci pour votre implication à mes côtés et votre accueil en Bretagne. J'ai beaucoup appris en travaillant avec vous. J'en profite pour remercier aussi **Aude Ernoult** avec qui j'ai eu l'occasion, à Rennes, d'échanger sur les méthodes d'analyses de paysage.*

*Merci à ceux du **CEN : Alexandra Janczak**, **Loïc Coquel**, **Lucile Chastel**, du **CPIE Chaine des terrils : Bruno Derolez**, et du **GON : Robin Quevillart** qui ont échantillonné les mares du littoral et du bassin minier pendant plusieurs mois. Merci aussi à tous les bénévoles et étudiants du Nord, je pense notamment **Camille Creton** et **Celia Hurez** qui ont traqué les calamites et pélodytes à travers les terrils de nuit durant plusieurs semaines ; aux bénévoles de **Bretagne vivante**, des **Conservatoires des Espaces Naturels**, de la **LPO** et des **Universités**, en Bretagne, en Suisse, dans le Lot, la Bourgogne, la Franche-Comté, la Lorraine. Sans eux nous n'aurions pas récolté tous les précieux échantillons d'ADN de ces centaines de pélodytes et calamites. Et merci à **Claire Blaise** qui m'a aidée au pied levé en me fournissant des cartes de données historiques sur les populations de crapauds calamites.*

*Et puisque nous voilà sur le terrain, merci également à **Baptiste Faure** qui nous a appris à poser les PIT-TAGs, nous a prêté son matériel et conseillé pour le choix de sites de suivi de populations. Merci pour ton*

aide et ton enthousiasme sur cette étude. Je remercie par ailleurs ***Virginie Callipel*** *de nous avoir donné l'autorisation d'étudier la population de crapauds calamites de Monchecourt en 2014.*

Je n'oublie pas tous ceux les collègues de l'EEP et amis qui ont eu la gentillesse de m'accompagner à Monchecourt de nuit, souvent sous la pluie et jusqu'à ce que, tard dans la nuit, les températures frôlant les 0°c fassent taire les plus combatifs des crapauds calamites. Merci surtout à ***Florian Boizet*** *qui a participé avec courage et motivation, à trois mois d'échantillonnage et d'élevage de crapauds calamites. Je lui en suis très reconnaissante. Merci aussi à* ***Sylvain Fossier*** *qui a été en stage quelques journées avec nous.*

****

Je n'oublie pas non plus les enseignants et chercheurs qui ont contribué plus en amont à éveiller mon goût pour la science et la recherche. En particulier, un grand merci à ***Didier Jollivet*** *qui a été un soutien indispensable pour mon grand saut vers la thèse.*

****

Je tiens également à remercier ceux qui, sans forcément contribuer de manière directe à mon travail, ont partagé mon quotidien et ont contribué en grande partie à faire de cette expérience de thèse une si belle aventure.

Un grand merci à l'ensemble de ***l'équipe GEPV/EEP*** *qui m'a chaleureusement accueillie à Lille. Si la vie lilloise a été une très belle surprise, c'est en partie parce que l'équipe regorge de collègues motivés et dynamiques toujours disponibles pour donner un coup de main. La bonne humeur de nos pauses cafés animées de débats plus extravagants les uns que les autres me manquera beaucoup. En particulier, merci à* ***Sophie Gallina, Clément Mazoyer*** *et* ***Mathieu Genete*** *qui ont toujours su se rendre disponibles pour m'aider à trouver les ressources informatiques nécessaires aux analyses et m'assister pour que j'avance dans les méandres des codes informatiques. Merci à* ***Sylvain Billiard*** *pour les cours sur les modèles linéaires généralisés. Merci à* ***Anne Duputié*** *avec qui j'ai pu discuter de mon travail d'analyse du paysage. Merci à* ***Cécile Godé****. C'était formidable de travailler à la paillasse avec toi. Merci pour ta patience et ta pédagogie qui m'ont permis d'aborder avec confiance les manips de biomol. Merci à* ***Laura Hénocq*** *qui était avec nous pour poser les premières pierres à l'édifice durant son stage de M2 en analysant les premiers échantillons recueillis. Merci à* ***Jacky Grimonprez*** *qui a toujours été très disponible pour trouver des solutions à mes « accidents » informatiques. Merci également* ***Sandrine Belingheri*** *et* ***Michèle Delcourt*** *puis* ***Séverine Six*** *pour leur dynamisme et réactivité au secrétariat qui nous rendent les tâches administratives tellement plus faciles. Ayant eu l'occasion de faire du monitorat, je tiens à remercier l'équipe pédagogique qui a fait en sorte que les séances que j'encadrais se passent dans les meilleures conditions en fournissant des supports de TD et TP solides. Et un grand grand merci aux autres thésards, post-docs et ATER, collègues et amis,* ***Hélène, Claire, Diala, Marie, Roxane, Maxime, Samson, Monika, Vincent, Thibault, Thomas, Clotilde****,* ***Dima****,* ***Laura****,* ***Nicolas****,* ***Mathilde****,* ***Bénédicte****,* ***Julien****,* ***Natasha****,* ***Marina****,* ***Clémentine****,* ***Renato, Marie-Joe, Romuald, Alessandro*** *et à* ***Maryse, Laura, Estelle, Thomas, Amélie*** *qui sont arrivés plus récemment, avec qui j'ai été très heureuse de partager mon quotidien de thésarde pour un moment ou durant les quatre ans.*

Enfin une grosse pensée aux amis rencontrés à Lille ou restés plus au sud et à ma famille qui sont une source intarissable d'optimisme et de motivation.

****

Je remercie la ***région Nord-Pas de Calais*** *pour la bourse de thèse et le* ***CPER-Climibio*** *pour le poste d'ingénieur d'étude dans le cadre desquels j'ai réalisé ce travail.*

SOMMAIRE

PRÉAMBULE

La diversité génétique est le support de l'évolution de la vie, dans le temps et dans l'espace sous l'effet de forces micro-évolutives qui sont : la mutation, la migration, la dérive et la sélection naturelle. Or, depuis plusieurs siècles, de nouvelles pressions de sélection qui résultent des activités humaines ont émergés (Pereira *et al.* 2012). La surexploitation des ressources, la pollution, la destruction des habitats, les invasions biologiques, mais également la fragmentation des habitats, sont autant de pressions anthropogènes survenues rapidement et souvent avec de grandes intensités. Ceci entraine une érosion de la biodiversité d'une telle ampleur que l'on parle aujourd'hui de la sixième crise d'extinction massive (Wake & Vredenburg 2008; Allendorf *et al.* 2012). Les amphibiens sont le taxon le plus impacté : un tiers des espèces sont menacées d'extinction et plus généralement 70% des espèces sont décrites comme étant en déclin (Wake & Vredenburg 2008). Au-delà de la disparition des espèces par l'effet direct d'une surexploitation, du manque de ressources, ou des pandémies (maladies émergentes, invasions biologiques), les extinctions sont induites de façon plus insidieuse par le biais de la perte de diversité génétique intraspécifique. Ce phénomène résulte de la perte et de la fragmentation des habitats qui entrainent une réduction des tailles de populations et la diminution des échanges de migrants qui garantissent des flux de gènes entre elles. Cette perte de diversité génétique fragilise les populations face aux fluctuations environnementales et s'avère être une menace majeure pesant sur le maintien des populations (Allentoft & O'Brien 2010; Merilä & Hendry 2014). Des programmes de protection des écosystèmes et des populations ont été mis en place afin de préserver certaines aires de toute activité humaine, en créant par exemple des parcs nationaux, représentant actuellement environ 11% de la surface du globe. Mais ces mesures ne suffisent pas à écarter les menaces qui pèsent sur la plupart des écosystèmes (Rodrigues *et al.* 2004). Aujourd'hui il apparait essentiel de multiplier les projets visant à mieux comprendre l'impact des changements environnementaux pour mettre en place des programmes efficaces de conservation de la biodiversité au sein de zones d'activités humaines (Alberti 2015).

Dans cette thèse nous nous sommes intéressés au cas de populations d'amphibiens de deux espèces, le Pélodyte ponctué (*Pelodytes punctatus*) et le Crapaud calamite (*Bufo* [*Epidalea*] *calamita*), établies dans le nord de la France. Dans cette région ces deux espèces sont observées le long du littoral dans des aires naturelles ou semi-naturelles mais aussi au sein du bassin minier sur d'anciens sites industriels, les terrils. Ces habitats artificiels sont au cœur d'une zone de fortes activités agricoles et urbaines. Ce travail de thèse visait à étudier les niveaux de diversité génétique neutre des populations du littoral et du bassin houiller en prenant en compte les effets conjoints de processus biogéographiques de colonisation des terrils par le Crapaud calamite, de la connectivité du paysage et du régime d'appariement sur la structure génétique des populations.

INTRODUCTION

"What are the minimum conditions for the long-term persistence and adaptation of a species or a population in a given place ? This is one of the most difficult and challenging intellectual problems in conservation biology. Arguably, it is the quintessential issue in population biology, because it requires a prediction based on a synthesis of all the biotic and abiotic factors in the spatial-temporal continuum."

Michael E. Soulé (1987)

1. Diversité génétique neutre et notion de structure génétique

La diversité génétique présente au sein d'une espèce ne se répartit pas de façon homogène. Elle se structure dans le temps et l'espace au sein de populations géographiquement structurées. Pourtant les allèles ségrègent de façon aléatoire lors de la formation de gamètes destinées à transmettre l'information génétique à la génération suivante. Cet écart à une distribution aléatoire attendue sous l'équilibre de Hardy-Weinberg peut être imputé à quatre forces évolutives : la mutation, la migration, la dérive génétique et la sélection naturelle dans laquelle est inclue la sélection sexuelle (Wright 1931, Tableau 1).

Allèles : Variant d'un gène résultant d'une mutation de ce gène

Equilibre de Hardy-Weinberg : équilibre atteint dans une population idéale (*i.e.* de taille infinie, non soumise à sélection, sans mutation et où les individus s'accouplent au hasard (en panmixie) et dans laquellle les fréquences alléliques restent constantes au cours du temps.

Dérive génétique : résulte de tous les phénomènes aléatoires (stochasticité démographique et environnementale) entrainant des fluctuations fortuites de fréquences alléliques au sein des populations, indépendamment de toutes autres forces évolutives.

Tableau 1 : Effets des quatre forces évolutives modualnt la diversité génétique au sein des populations, inspiré de Bergstrom & Dugatkin (2012).

Forces évolutive	Effets sur la diversité génétique au sein de l'espèce	Variance génétique intra-population	Variance génétique inter-population
Mutation	Apport de nouveaux variants alléliques	Augmente	Augmente
Migration	Échanges d'allèles entre populations	Augmente	Diminue
Dérive	Fluctuations aléatoires des fréquences d'allèles au sein des populations	Diminue	Augmente
Sélection	Favorise certains allèles par rapport à d'autres	Diminue (sauf en cas de sélection balancée)	Augmente si les populations sont dans des conditions différentes, diminue si les populations sont dans des conditions identiques

Un des objectifs de la génétique des populations est d'identifier, à partir de l'étude des patrons spatiaux de structure génétique, la nature de la différentiation génétique observée entre populations et, si possible, d'identifier les facteurs responsables de cette structuration spatiale. La diversité génétique est garante du potentiel adaptatif des espèces. Des pertes drastiques de diversité génétique peuvent induire de la dépression de consanguinité et une diminution de la réponse immunitaire au sein des populations (Reed & Frankham 2003; Johansson *et al.* 2007; Siddle *et al.* 2007; Beebee *et al.* 2012; Marsh *et al.* 2017). La biologie de la conservation s'intéresse plus particulièrement aux facteurs qui conditionnent les niveaux de diversité génétique et influent sur le potentiel des populations à se maintenir dans le temps.

Dans le cadre de ces travaux de thèse, nous nous sommes focalisés sur les processus pouvant expliquer la structuration spatiale de la diversité génétique neutre sur une échelle régionale, à savoir : (i) les flux de gènes et fluctuations démographiques passées liées à l'historique de colonisation, (ii) les flux de gènes « contemporains » liés à l'influence de la matrice paysagère, et enfin (iii) les flux de gènes au sein des populations, liés au régime d'appariement (Figure 1).

Gène : unité de base transmise par hérédité à la descendance. En génétique des populations ce terme désigne les variants alléliques définissant les génotypes des individus

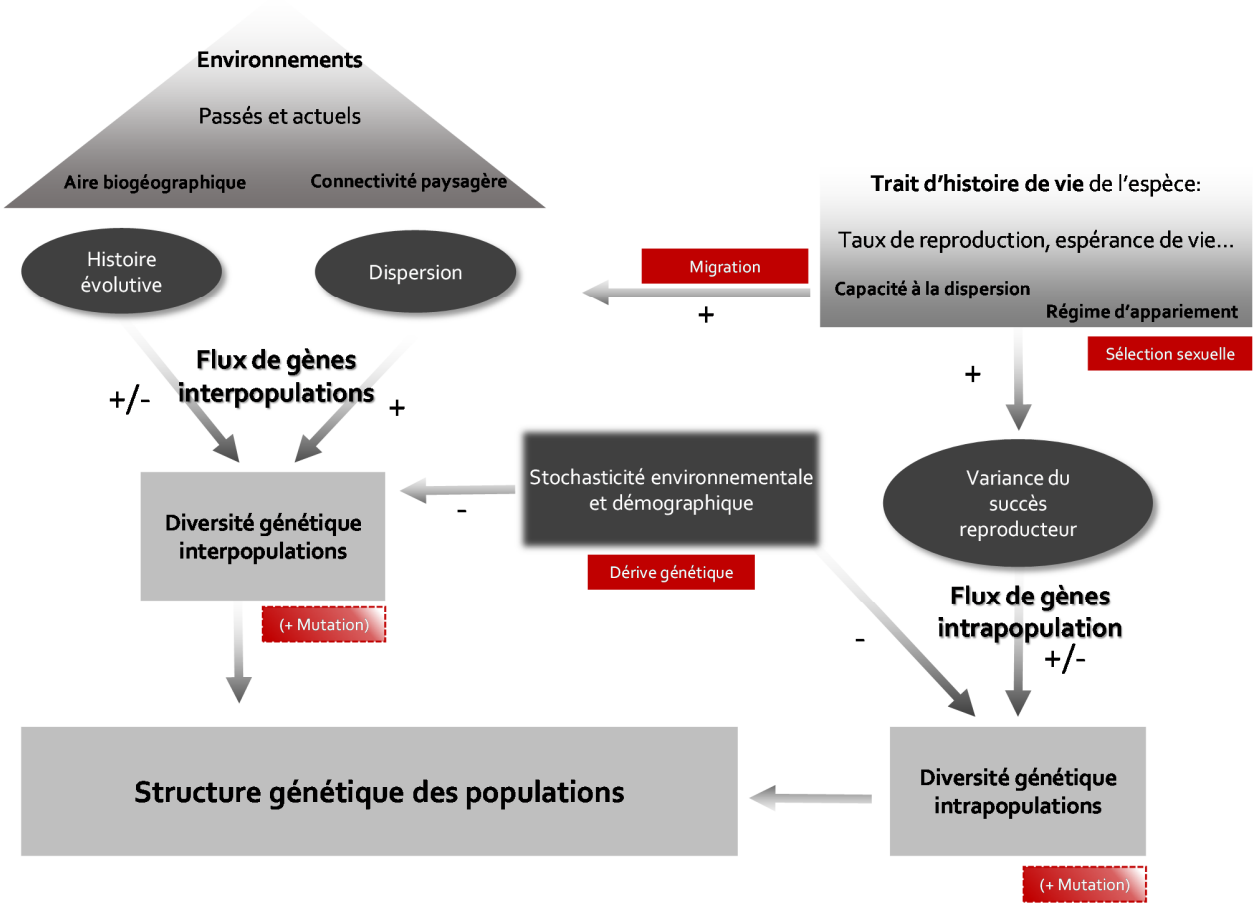


Figure 1: Schéma décrivant les paramètres conditionnant la structure génétique des populations ainsi que les forces évolutives (en rouge) ayant un effet positif ou négatif sur les niveaux de diversité génétique observés au sein et entre les populations.

a. Les flux géniques, définition

Le flux de gènes est un terme englobant tous les mécanismes induisant un déplacement des gènes d'une population vers une autre (Slatkin 1985). Le plus souvent, ce terme est employé comme synonyme de migration impliquant un évènement de dispersion d'une population à une autre, suivi d'un évènement de reproduction. Les flux de gènes peuvent toutefois désigner également les transferts de gènes qui s'effectuent d'une génération à l'autre au sein d'une même population. Les flux de gènes dépendent de facteurs endogènes liés aux traits d'histoire de vie de l'espèce, à ses capacités de dispersion ou son régime d'appariement… et de facteurs exogènes comme la prédation, la topographie ou le climat. Les flux de gènes varient donc d'une espèce à l'autre et fluctuent au cours du temps, façonnant ainsi les patrons géographiques de polymorphisme génétique.

J'emploie ici le terme **migration** pour désigner la force évolutive impliquant un transfert de gènes d'une population à l'autre suite à un évènement de **dispersion** suivi d'une reproduction.

b. Historique de colonisation et effet fondateur

Les fortes réductions d'effectifs liés à de la stochasticité environnementale ou démographique, de même que les évènements de fondation de nouvelles populations, induisent des pertes de diversité génétique par dérive génétique communément nommées « goulot d'étranglement » ou « effet fondateur ». L'histoire évolutive des populations conditionne en grande partie leur diversité génétique et leur niveau de différenciation génétique (Wade & McCauley 1988; McCauley *et al.* 1995; Le Corre & Kremer 1998). Lors de processus de colonisation, plus le nombre d'individus fondateurs est restreint, plus la perte de diversité génétique est élevée (Figure 2). Les populations issues d'évènements de colonisation distincts seront d'autant plus différenciées génétiquement entre elles que la proportion d'individus colonisateurs provenant de la même origine est élevée ou que l'effectif d'individus fondateurs est petit (Figure 2).

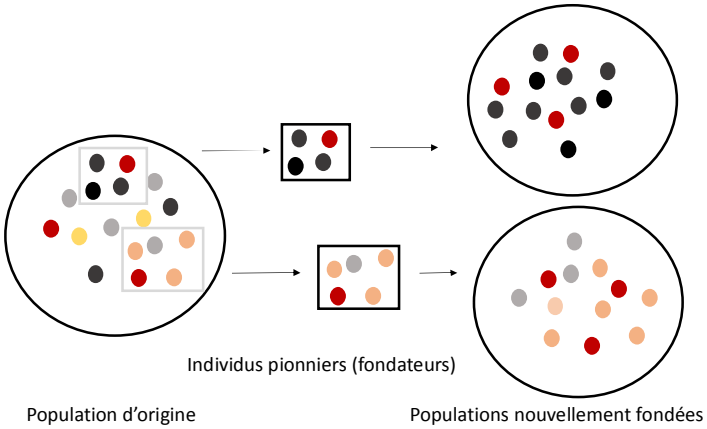


Figure 2 : Schéma de l'effet de goulot d'étranglement lié à la fondation d'une nouvelle population

La perte de diversité génétique est extrêmement dépendante de la taille efficace de la population. Sewall Wright (1969) définissait ainsi la taille efficace de la population comme « tout ce qui doit être substitué dans la formule (1/2*N*) pour décrire la perte effective d'hétérozygotie », 1/(2*N*) étant la probabilité que 2 gènes soient identiques par ascendance à la génération précédente dans une population idéale (Allendorf *et al.* 2012). Ce paramètre mesuré sur des populations naturelles prend en compte la perte de diversité génétique due, non seulement aux fluctuations démographiques des populations, mais également à la contribution inégale des individus à la génération suivante liée au régime d'appariement. Il est donc important de pouvoir estimer l'impact de l'historique des populations sur la structure génétique neutre des populations pour mieux évaluer les facteurs actuels entrainant une perte de diversité génétique.

c. Paysage et connectivité entre populations

Dans un environnement homogène, plus la distance géographique entre deux populations est grande, plus la différenciation génétique entre ces deux populations est élevée si la dispersion est spatialement contrainte (Wright 1943; Slatkin 1993). On parle alors d'isolement par la distance. Ceci s'explique par une diminution des flux de gènes entre les populations, diminution liée à une moins forte probabilité que les individus dispersent de l'une vers l'autre. Dans un environnement hétérogène, la probabilité de dispersion des individus d'une population à l'autre dépendra moins de la distance géographique Euclidienne qui sépare les populations que de l'impact des différents facteurs environnementaux sur les capacités et probabilités de dispersion des individus. Parmi ces facteurs environnementaux certains auront tendance à diminuer

Habitat : en écologie ce terme regroupe les conditions biotiques et abiotiques qui définissent un milieu et les ressources disponibles pour la survie et la reproduction des individus.

les probabilités de dispersion des individus entre populations structurées, on parlera alors d'effet barrière. Au contraire, d'autres facteurs augmenteront les probabilités de dispersion, on parlera d'effet corridor. La connectivité des populations dépend donc de l'ensemble des éléments du paysage qui relient les taches d'habitats entre elles et de leurs influences respectives sur la dispersion des individus pour une espèce donnée (Fischer & Lindenmayer 2007). Ainsi les discontinuités génétiques existant entre populations peuvent être mise en relation avec les éléments du paysage.

d. Régimes d'appariement et variance du succès reproducteur

Les niveaux de diversité génétique des populations sont conditionnés par les flux de gènes inter-populations mais également par les flux de gènes intra-population qui eux peuvent dépendre en partie du régime d'appariement. En effet, au sein d'une population, tous les individus ne contribuent pas de façon égale à la génération suivante. Plus le déséquilibre est élevé dans l'accès à la reproduction, plus la variance du succès reproducteur prenant en compte le nombre de jeunes produits augmente, ceci résultant en des flux de gènes biaisés d'une génération à l'autre. Cette variance du succès reproducteur est un facteur clé dans le processus de sélection sexuelle, même si elle n'implique pas forcément d'effet de la sélection sexuelle (Arnold & Wade 1984; Jones 2009). La sélection sexuelle résulte de « l'avantage que certains individus ont sur d'autres individus du même sexe et de la même espèce, en relation exclusive avec la reproduction » (Darwin 1871). Les caractères secondaires, morphologiques ou comportementaux qui amènent à un dimorphisme sexuel plus ou moins prononcé résultent de la sélection sexuelle pouvant agir de concert sur les deux sexes ou seulement l'un des deux. La sélection sexuelle peut ainsi impliquer une sélection intra-sexuelle par le biais de la compétition entre individus du même sexe ou une sélection inter-sexuelle par le biais d'un choix de partenaire (Møller & Danchin 2008).

Défini en 1977 par Emlen & Oring, le régime d'appariement regroupe l'ensemble des comportements liés aux interactions entre les partenaires de reproduction chez les espèces sexuées. Il se définit comme l'interaction de plusieurs paramètres : le nombre de partenaires, la manière d'accéder à la reproduction, la nature et la durée des liens entre les partenaires au cours d'une saison de reproduction, et l'importance des soins parentaux fournis par chacun des parents (Møller & Danchin 2008). Les régimes d'appariement se déterminent principalement par la capacité d'un sexe à monopoliser des partenaires du sexe opposé soit par association directe, soit en contrôlant l'accès aux ressources pour la reproduction. Quatre grandes catégories ont ainsi pu être définies sur la base

du nombre de partenaires sexuelles monopolisés sur une saison de reproduction : la monogamie, la polygynie, la polyandrie et la promiscuité/polygynandrie impliquant respectivement un seul partenaire par mâles et par femelles, plusieurs partenaires pour les mâles ou pour les femelles et plusieurs partenaires pour les mâles comme pour les femelles. Les régimes d'appariement diffèrent non seulement entre espèces, mais peuvent également fluctuer entre populations et même entre individus au sein d'une même population, et conditionnent en partie la diversité génétique observée au sein des populations (Frankham *et al.* 2002). Savoir comment les évènements de reproduction s'établissent et quelles incidences ils ont sur la taille efficace des populations est essentiel pour mieux évaluer la viabilité des populations.

2. *Écologie des amphibiens*

Les amphibiens comptent parmi les taxons animaux ayant les niveaux de différenciation génétique les plus élevés (Tableau 2, Allendorf *et al.* 2012). Ces estimations sont fortement dépendantes des marqueurs moléculaires utilisés et de l'échelle d'étude. Néanmoins cette observation suggère des flux de gènes généralement assez limités entre populations d'amphibiens induisant de forts effets de dérive génétique.

Les amphibiens regroupent les espèces de trois ordres différents : les anoures (grenouilles, crapauds et rainettes), les urodèles (tritons et salamandres) et les gymnophiones (cécilies) occupant une multitude d'habitats largement réparties sur l'ensemble du globe à l'exception de l'Antarctique (Wells 2010).

Tableau 2: Comparaison des niveaux de diversité génétique et des niveaux de différenciation génétique (F_{ST}) pour les différents taxons animaux (Ward *et al.* 1992) et plantes classés selon leur distribution géographique (Hamrick and Godt 1990, 1996). H_T=diversité génétique au sein de l'espèce, H_S= diversité génétique au sein des populations (tiré de Allendorf *et al.* 2012, p.50).

Taxons	H_t	Hs	F_{st}	Nb d'espèces
Amphibiens	0.136	0.094	0.315	33
Oiseaux	0.059	0.054	0.076	16
Poissons	0.067	0.054	0.135	79
Mammifères	0.078	0.054	0.242	57
Reptiles	0.124	0.09	0.258	22
Crustacées	0.088	0.063	0.169	19
Insectes	0.138	0.122	0.097	46
Mollusques	0.157	0.121	0.263	44
Plantes endémiques	0.096	0.063	0.248	100
Plantes régionales	0.15	0.118	0.216	180
Plantes ubiquiste	0.202	0.159	0.21	85

a. Tailles de population

Les populations d'amphibiens ont une dynamique démographique souvent complexe. Les stades larvaires sont inféodés au milieu aquatique tandis que les adultes peuvent rester totalement aquatiques, vivre en partie dans des points d'eau ou devenir totalement terrestres selon les espèces (Beebee 1996). La métamorphose signe donc le passage d'un habitat à un autre souvent très différent. Ainsi les populations d'amphibiens requièrent généralement plusieurs sites d'habitats entre lesquels les individus pourront migrer. Les populations peuvent présenter de fortes fluctuations dans le temps, avec plusieurs années de déclin suivies d'années de croissance brusque caractérisées par une explosion démographique (Beebee & Griffiths 2005). De nombreuses populations d'amphibiens présentent d'ailleurs de petites tailles efficaces, souvent bien inférieures à la taille de la population recensée par comptage des individus (Beebee 1996; Phillipsen *et al.* 2011; Wang & Shaffer 2017).

Ces faibles tailles efficaces peuvent s'expliquer par l'instabilité des sites de reproduction. La survie larvaire, par exemple, est très faible (0.2% à 6% chez les anoures, Beebee 1996). De fortes fluctuations peuvent notamment s'observer chez les espèces dites pionnières occupant des sites souvent éphémères, des flaques d'eau où des points d'eau en formation qui plus tard peuvent être amenés à se refermer entre autre avec la densification de la végétation. La survie des espèces est alors grandement tributaire d'évènement de dispersion et de colonisation qui pallient aux extinctions locales de populations (Stevens & Baguette 2008). Dans le cas le plus extrême, chaque population étudiée indépendamment ne présente pas une taille suffisante pour assurer sa pérennité. Mais si ces petites populations sont suffisamment bien connectées pour permettre des évènements de migration ou colonisation et recolonisation, et si les dynamiques locales sont désynchronisées, ces sous-populations vont former ce qu'on appelle une métapopulation qui perdurera dans le temps (Smith & Green 2005; Marsh & Trenham 2015). Ce système a été décrit chez de nombreux amphibiens à plusieurs reprises (Smith & Green 2005).

b. Capacité de dispersion

Les amphibiens sont connus pour avoir d'assez faibles capacités de dispersion, de l'ordre de quelques kilomètres. Du moins, c'est ce que la plupart des études rapportent. Sur 166 études, Smith & Green (2005) relèvent 70% de cas où la distance maximum de dispersion reste inférieure à un kilomètre. Ces données sont toutefois fortement dépendantes de l'échelle à laquelle les processus de dispersion sont étudiés et les auteurs suggèrent qu'il existe un biais lié aux faibles échelles spatiales

pour lesquelles les études sont réalisées. En effet, certains amphibiens semblent pouvoir disperser sur des distances de plus d'une dizaine de kilomètres (Smith & Green 2005). Il semblerait donc exister une grande variabilité inter et intraspécifique dans les capacités de dispersion des amphibiens (Tableau 3).

Tableau 3 : Quelques exemples extraits de la méta-analyse de Smith & Green (2005) pour illustrer la variabilité inter et intra-spécifique des distance de dispersion enregistrées chez diverses espèces d'amphibiens.

Espèces	Distance enregistrée (m)	Références
Bufo americanus	6437	Hamilton 1934
	547	Blair 1943
Bufo calamita	4411	Sinsch 1989
	400	Miaud *et al*. 2000
Bufo marinus	15100	Brechenridge & tester 1961
	35000	Kusano *et al* . 1995
Bufo boreas	200	Muths 2003
	6000	Bartlet 2000
Ambystoma coliforniense	670	Trenham *et al*. 2001
	129	Loredo *et al* 1997
Cryptobranchus allegani	990	Peterson 1987
	85	Wiggs 1977

Les contraintes à la dispersion des amphibiens peuvent être d'ordre physiologiques, comme la sensibilité à l'hygrométrie due à leur peau très perméable ou comportementales si l'espèce est philopatrique (Blaustein *et al.* 1994; Smith & Green 2005). Elles sont aussi fortement liées aux contraintes environnementales telles que la prédation, la topographie ou la couverture végétale. Certains éléments paysagers semblent notamment fortement conditionner le taux de dispersion et donc les niveaux de flux géniques entre populations (Cushman 2006). Pour ne citer que quelques exemples, certaines études ont montré l'effet des routes agissant comme des barrières sur la dispersion des amphibiens (Vos *et al.* 2001; Youngquist *et al.* 2017). D'autres ont mis en évidence l'importance des habitats ouverts, tels que les garrigues agissant comme corridors pour les populations méditerranéennes, ou les golfs facilitant les flux de gènes entre populations en zone urbanisée (Saarikivi *et al.* 2013; Gutiérrez-Rodríguez *et al.* 2017).

c. Régimes d'appariement

Chez les amphibiens comme chez la plupart des animaux la compétition entre mâles joue un rôle important pour l'accès à la reproduction (Wells 2010). Cette compétition résulte de la disparité entre mâles et femelles dans l'investissement énergétique pour la production de gamètes. Les mâles consacrant moins d'énergie à la production de gamètes, ils peuvent alors principalement s'investir dans lacompétition pour l'accès aux femelles qui constituent une ressource limitante (Trivers 1972). L'intensité de la sélection sexuelle varie donc selon le sexe considéré, la disponibilité des ressources, certains paramètres démographiques comme le sex-ratio, et le comportement des autres membres conspécifiques (Bergstrom & Dugatkin 2012).

Sex-ratio : ratio du nombre de mâles par femelle

Chez les amphibiens, et en particulier chez les anoures, le dimorphisme sexuel n'est pas fortement marqué d'un point de vue morphologique. Au contraitre, des comportements bien distincts peuvent s'observer : la parade nuptiale chez les urodèles, et le chant chez les anoures. Ces comportements suggèrent l'effet d'une sélection intra et inter-sexuelle impliquant une compétition entre mâles et potentiellement un choix de la part des femelles (Lehmann & Perrin 2003; Jaquiéry *et al.* 2010; Muralidhar *et al.* 2014). Ces mécanismes de sélection sexuelle peuvent significativement impacter la variance du succès reproducteur (Arnold & Wade 1984; Howard *et al.* 1997; Jones & Ratterman 2009). Le régime d'appariement doit donc être étudié pour mesurer l'effet du fonctionnement intra-populationnel sur la structure génétique de populations d'amphibiens.

3. Importance de la connectivité paysagère en biologie de la conservation

Dans le contexte actuel de fortes perturbations de l'environnement par les activités humaines, la perte des habitats et leur fragmentation s'avèrent être une cause majeure de la perte de biodiversité (Fischer & Lindenmayer 2007). En tant que telle, la fragmentation n'induit pas forcément de processus d'extinction (Fahrig 2017). La fragmentation du paysage a toujours existé puisque les écosystèmes présentent tous une diversité de facteurs environnementaux qui modulent les taux de dispersion au sein des espèces. Une fragmentation de l'habitat peut notamment mener à des adaptations locales modulant la valeur sélective des individus localisés dans des environnements hétérogènes soumis

Valeur sélective : elle décrit la capacité d'un génotype à se propager à travers le temps *via* un meilleur taux de survie et un meilleur succès reproducteur. C'est une mesure de la force de sélection

à des pressions de sélections différentes (Kawecki & Ebert 2004). Toutefois, la diminution de la taille d'habitats favorables et la diminution de leur disponibilité qui opère conjointement avec l'apparition d'éléments fragmentants, entrainent des effets négatifs sur la biodiversité. Les habitats de plus petites tailles ont de faibles capacités de charge et les populations sont donc restreintes à de petits effectifs en plus d'être isolées géographiquement les unes des autres (Figure 3).

Tache d'habitat homogène | Tache d'habitats fragmentés | Tache d'habitats fragmentés avec perte de certains habitats

Figure 3 : Effets respectifs de la fragmentation des habitats et de la fragmentation couplée à la perte d'habitats sur la structure et la composition d'une tache de paysage. A chaque couleur différente est attribué un habitat spécifique, la couleur grise symbolise des habitats non exploitables par l'espèce (inspiré de Fahrig 2017).

Les amphibiens dépendent, pour la grande majorité, de milieux humides ce qui peut considérablement limiter leur dispersion. Les milieux humides sont en effet assez rares, ils représenteraient entre 4 à 6% des zones émergées de la planète et on estime qu'au cours du dernier siècle, 64% des zones humides ont disparus (Hu *et al.* 2017). Dans ce contexte de régression des milieux humides, l'environnement peut rapidement isoler les populations les unes des autres.

Pour les populations caractérisées par un fonctionnement en métapopulations, comme c'est le cas de beaucoup d'amphibiens, ces modifications de l'habitat peuvent avoir des effets très rapides en limitant les échanges de gènes entre populations, empêchant les processus de (re)colonisation. Chez d'autres taxons comme les mammifères, des études ont montré que l'isolement géographique pouvaient conduire à un effondrement rapide des populations en quelques générations seulement (Krebs *et al.* 1969). De manière plus générale, des effectifs réduits de populations induisent une vulnérabilité plus fortes des populations face aux effets de la dérive génétique entrainant une érosion de la diversité génétique et une disparition d'allèles rares (Johansson *et al.* 2007). Si ces populations sont isolées géographiquement, elles perdent le bénéfice d'apport de nouveaux allèles et peuvent rapidement subir des effets de dépression de consanguinité, une perte de leur réponse immunitaire et une diminution de leur capacité d'adaptation. Les conséquences de cette perte de valeur sélective sont multiples. Ceci peut, par exemple, entrainer des diminutions du taux de survie et du taux de croissance des populations conduisant alors à une diminution de leurs effectifs. Ce processus cyclique prenant

en compte cette rétroaction positive des facteurs démographiques et des facteurs génétiques est ainsi décrit sous le nom de vortex d'extinction (Gilpin & Soulé 1986, Figure 4).

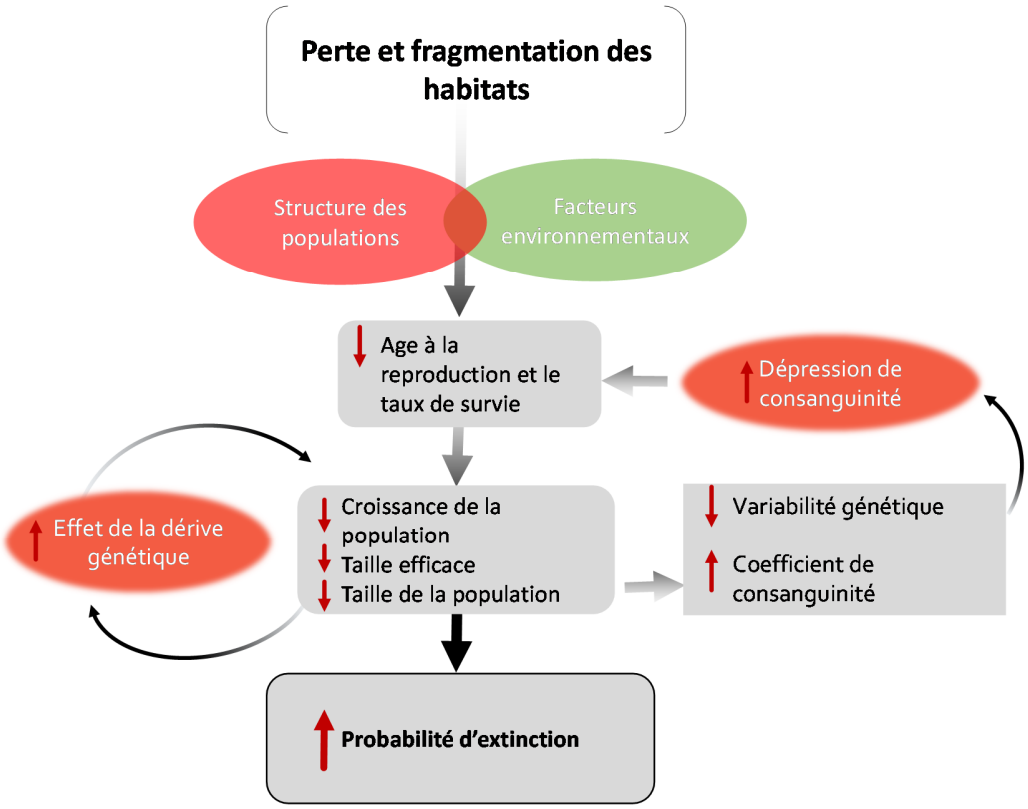


Figure 4 : Schéma simplifié du vortex d'extinction où, initié par la perte et la fragmentation des habitats, les facteurs démographiques et génétiques interagissent rétroactivement (d'après Soulé & Mills 1998). La structure des populations inclue la structure d'âge, le sex-ratio, les taux intrinsèques de natalité et de mortalité, les interactions comportementales, la distribution des populations et leurs réponses physiologiques densité-dépendantes. Les facteurs environnementaux comprennent les facteurs biotiques et abiotiques interagissant avec la population dans l'habitat donné.

Un nombre croissant d'études s'intéressent ainsi aux flux de gènes inter-populations chez les amphibiens et à l'impact de la fragmentation du paysage (*p.e.* Vos *et al.* 2001; Spear *et al.* 2005; Stevens *et al.* 2006; Arens *et al.* 2007; Wang 2009; Noël & Lapointe 2010; Coster *et al.* 2015; Youngquist *et al.* 2017; Gutiérrez-Rodríguez *et al.* 2017). Selon l'espèce et selon le type d'habitat, la structure génétique spatiale peut-être restreinte à moins de un kilomètre ou bien s'étendre à plus d'une dizaines de kilomètres (Figure 5). En Europe de l'Ouest il a été montré que les populations d'amphibiens sont caractérisées par un déclin général de leur niveau de diversité génétique liée à l'anthropisation croissante des milieux (Cushman 2006; Miraldo *et al.* 2016). Ceci souligne l'importance de prendre en compte l'interaction entre paysage et dispersion pour évaluer les capacités des populations à se maintenir au sein des zones de fortes activités humaines.

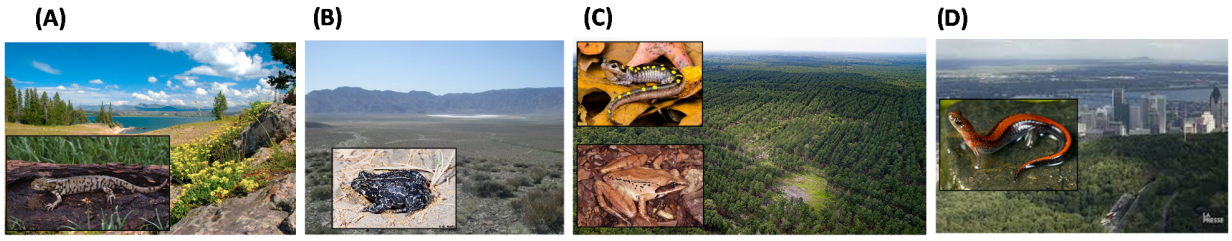


Figure 5 : Quelques exemples d'estimation de l'échelle spatiale à laquelle opèrent les flux de gènes chez des espèces d'amphibiens en aires naturelles (A et B) ou anthropisées (C et D), obtenus par une approche de génétique des populations. (A) Spear *et al*, 2005, (B) Wang 2009, (C) Coster *et al*, 2015, (D) Noël & Lapointe 2010.

4. Cadre de l'étude et objectifs généraux de la thèse

a. Particularités des friches minières de la plaine du Nord de la France

La plaine du Nord-Pas de Calais est une vaste plaine agricole aujourd'hui très urbanisée. C'est la région avec l'un des plus fort taux de surfaces urbanisées et agricoles et, en contrepartie, le taux le plus faible de surfaces boisées en France métropolitaine. Par ailleurs les zones humides ne représentent pas plus de 0.5% de la surface occupée. Depuis l'ouverture de la première exploitation minière en 1757 à la fermeture du dernier puit en 1990 le bassin minier s'est doté de nouveaux habitats très singuliers, les terrils (Figure 6).

Figure 6 : Photographie des terrils de la région de Liévin dans les années 70.

Issus des résidus miniers ces terrils constituent aujourd'hui un agrégat de grès et de schiste carbonifère qui petit à petit ont pu se végétaliser. Ils représentent en tout 2500 hectares. Certains sont exploités pour l'utilisation du schiste, d'autres sont devenus des bases de loisir ou encore des zones

naturelles protégées. Les terrils et le bassin minier du Nord-Pas de Calais ont, depuis 2012, obtenus le statut de patrimoine mondial de l'UNESCO. Leur sol très minéral et leur albédo élevé, la température du sol étant en moyenne 5°C plus élevé sur les terrils qu'ailleurs dans la région, en font des milieux atypiques dans le nord de la France. De plus, ne faisant l'objet d'aucune exploitation agricole (sauf de rares exceptions), ils sont relativement préservés des pollutions liées aux intrants couramment rencontrés en zones agricoles. Ces sites hébergent une partie de la flore et de la faune régionales originaires d'espaces aux conditions édapho-climatiques proches, comme celles que l'on rencontre sur les dunes et les coteaux calcaires (Lemoine 2012). Parmi ces espèces, le Pélodyte ponctué et le Crapaud calamite y ont trouvé un habitat pionnier adapté. Les nappes phréatiques, dont le niveau a pu remonter en surface après l'abandon des galeries minières, et le sol rendu imperméable par la compaction du schiste dans certaines zones, ont entrainé la mise en place de points d'eaux permanents ou plus éphémères. Enfin, les blocs de grès fournissent une zone de refuge idéale pour ces deux espèces d'amphibiens. Néanmoins, le paysage est dominé au voisinage par les terrains agricoles (~72%) et des zones urbaines (~16%). La matrice paysagère est donc profondément fragmentée par des éléments anthropogènes et interpelle sur le devenir des populations aujourd'hui établies sur les sites miniers et qui pourraient se retrouver isolées de toutes autres populations tant les barrières à la dispersion semblent nombreuses.

A l'ouest, séparé du bassin minier par une zone moins urbaine, le littoral présente de vastes aires dunaires avec un paysage plus continu et moins perturbé par les activités humaines. Une grande partie de la côte est classée en espaces protégés gérés par des organismes de protection de la nature. On y relève toutefois une régression du nombre de sites de pontes de l'ordre de 3.5% dû à une reconversion des terrains pâturés en des terres de cultures arables (Arntzen *et al*. 2017).

b. Le Crapaud calamite et le Pélodyte ponctué

Le Pélodyte ponctué (*Pelodytes punctatus*, Daudin, 1803) et le Crapaud calamite (*Bufo calamita,* Laurenti, 1768), devenu récemment *Epidalea calamita*) sont deux anoures affectionnant les sites pionniers (Figure 7 A et B). On les trouve dans les habitats littoraux où ils ont pour sites de reproduction des pannes dunaires, dans des environnements prairiaux, et enfin dans des sites plus anthropisées tels que les carrières ou les terrils où ils peuvent pondre dans des ornières ou flaques temporaires (Beebee & Denton 1996; Beja & Alcazar 2003; Stevens *et al.* 2003; Salvidio *et al.* 2004; Flavenot *et al.* 2015). Les points d'eau sont généralement ensoleillés et peu profonds, même si ces

espèces s'accommodent parfois de mares assez profondes notamment en ce qui concerne le Pélodyte ponctué (Figure 7 C à E). Les faibles niveaux d'eau et l'ensoleillement favorisent un développement assez rapide des têtards grâce au réchauffement de l'eau et à une limitation de la compétition interspécifique.

Le Crapaud calamite à une aire de répartition couvrant une grande partie de l'Europe, depuis le Portugal jusqu'au sud de la Suède (Beebee *et al.* 2012). Selon la base de données SiRF (Système d'information Régionale du la Faune), la présence du crapaud calamite a été signalée pour la première fois en 1829 sur le littoral du Pas-de-Calais. Des données indiquent par ailleurs sa présence en 1899 au sein du bassin minier. Aujourd'hui on recense 42 sites miniers pour lesquels des populations de Crapauds calamites y sont établies. Les tailles de populations sont parfois très importantes allant jusqu'à plusieurs centaines de mâles chanteurs estimés au sein d'une population.

Le Pélodytes ponctué est, dans le Nord de la France, situé en limite d'aire de répartition géographique. Son aire de distribution se résume à la France, au Portugal, à une partie de l'Espagne, et à une zone restreinte au Nord-Ouest de l'Italie. Dans le nord de la France la répartition du Pélodyte ponctué reste beaucoup plus éparse que celle du Crapaud calamite et les tailles de populations sont généralement petites, de l'ordre d'une 10^{aine} d'individus. A l'heure actuelle, sont recensés 25 sites occupés dans le bassin minier et le long du littoral (sources, SiRF). Ces deux espèces d'amphibiens sont sur la liste rouge de l'IUCN des amphibiens menacés de France métropolitaine. Ils sont sous statut de protection nationale, l'habitat du Crapaud calamite étant également sous statut de protection.

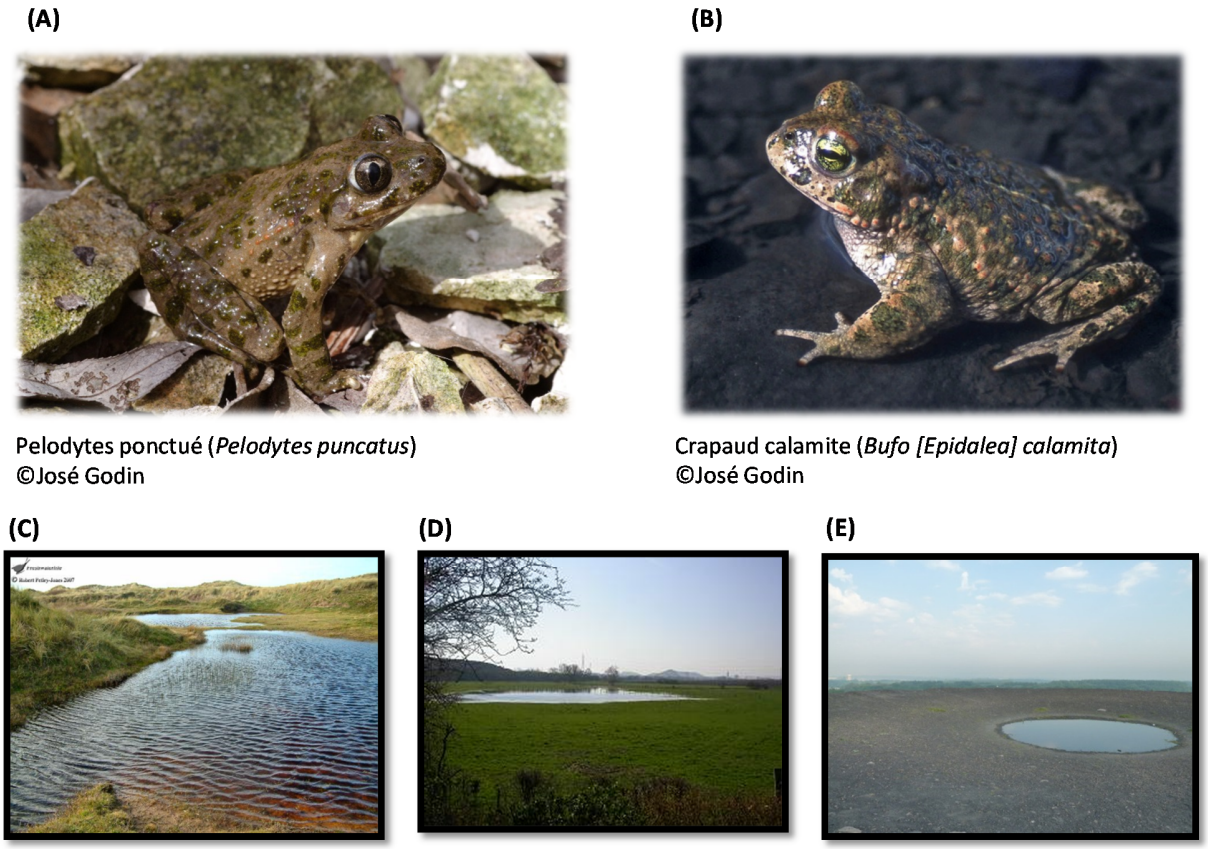


Figure 7 : Photographies d'un Pélodyte ponctué (A) et d'un Crapaud calamite (B) adultes, et de trois sites de reproduction typiques : (C) panne dunaire, (D) mare prairiale, (E) point d'eau temporaire sur un terril.

c. Architecture générale de ce document de thèse

Les populations de Pélodyte ponctué et de Crapaud calamite du bassin minier du nord de la France occupent des sites artificiels mis en place au cours de ces trois derniers siècles. Pour mieux appréhender les différents facteurs conditionnant les niveaux de diversité génétique au sein des populations et leur structure génétique spatiale, il a été nécessaire d'intégrer l'effet de différents processus historiques et plus contemporains.

Dans un premier temps, de nouveaux marqueurs nucléaires et mitochondriaux ont été développés afin d'avoir une puissance statistique suffisante pour décrire une histoire évolutive récente et déterminer comment s'opèrent les évènements de reproduction chez le Crapaud calamite. Ce travail fait l'objet du Chapitre I intitulé : *« Développement de nouveaux marqueurs nucléaires et mitochondriaux pour l'étude de la structure génétique des populations de Bufo [Epidalea] calamita »*

Les populations d'amphibiens du bassin minier sont établies dans des sites datant de moins de trois siècles, ce qui implique qu'il y ait eu des évènements de fondation récents. Ces processus sont connus pour fortement impacter les niveaux de diversité génétique et la structuration génétique des populations (McCauley *et al.* 1995). La structuration génétique des populations de Crapaud calamite a d'abord été étudiée sur une grande échelle géographique, en intégrant des populations échantillonnées ailleurs en France, en Suisse et en Suède, afin de pouvoir déterminer les scénarios biogéographiques les plus probables ayant conduits à l'établissement des populations de Crapaud calamite au sein du bassin houiller. Cette étude fait l'objet du chapitre II intitulé : « *Histoire évolutive du Crapaud calamite dans le nord de la France* ».

La réduction des flux de gènes entre populations est une des causes majeures d'extinction des populations. Le bassin minier est une zone fortement anthropisée ayant la particularité de fournir de nouveaux habitats propices à des espèces pionnières telles que le Pélodyte ponctué et le Crapaud calamite. Toutefois, ces habitats sont répartis de façon irrégulière au sein d'un paysage fortement fragmenté par des zones d'activités agricoles et des zones urbaines. Dans cette troisième partie nous avons donc entrepris d'évaluer les niveaux de connectivité entre populations en essayant de déterminer les variables paysagères pouvant impacter positivement ou négativement les flux géniques entre populations. Ce travail fait l'objet du chapitre III intitulé : « *Impact du paysage et des tailles efficaces sur les patrons de flux géniques chez le Pélodyte ponctué et le Crapaud calamite établis dans des zones fortement anthropisées* ».

Enfin le dernier volet d'étude de ces travaux de thèse concerne le fonctionnement intra-population pour un site de reproduction du Crapaud calamite. Le Crapaud calamite est une espèce polygyne fonctionnant en lek où une forte compétition entre mâles semble s'exercer (Arak 1988; Tejedo 1992). La variance du succès reproducteur mâle semble soumise à une dynamique complexe et très dépendante de la densité de population (Tejedo 1988; Stevens *et al.* 2003). Par ailleurs, les femelles de cette espèce peuvent apparaitre plusieurs fois au sein du lek et certaines observations laissent penser qu'il puisse y avoir un choix des femelles (Arak 1988). Nous avons donc étudié les processus de sélection sexuelle sous-jacents à ce régime d'appariement en mesurant (i) la variance du nombre d'accouplement ainsi que (ii) la variance du succès reproducteur pour chacun des deux sexes. Une étude des traits des adultes ainsi qu'un suivi de leur présence au cours d'une saison entière de

reproduction a été réalisé. Enfin, un suivi de traits liés à la survie de la progéniture a également été entrepris sur un échantillon d'œufs prélevé sur chacune des ponte produites dans la saison. Ce travail fait l'objet du chapitre IV intitulé « *Analyse du succès reproducteur en population naturelle* ».

Depuis le processus biogéographique de colonisation des terrils par le Crapaud calamite jusqu'au régime d'appariement rencontré au sein d'une population, en passant par l'analyse multi-espèces de la connectivité des populations, les approches de génétique des populations employées dans ce travail de thèse ont visé à décrire les effets de différents processus micro-évolutifs sur les niveaux de diversité génétique observés dans ces habitats singuliers que sont les terrils (Figure 8).

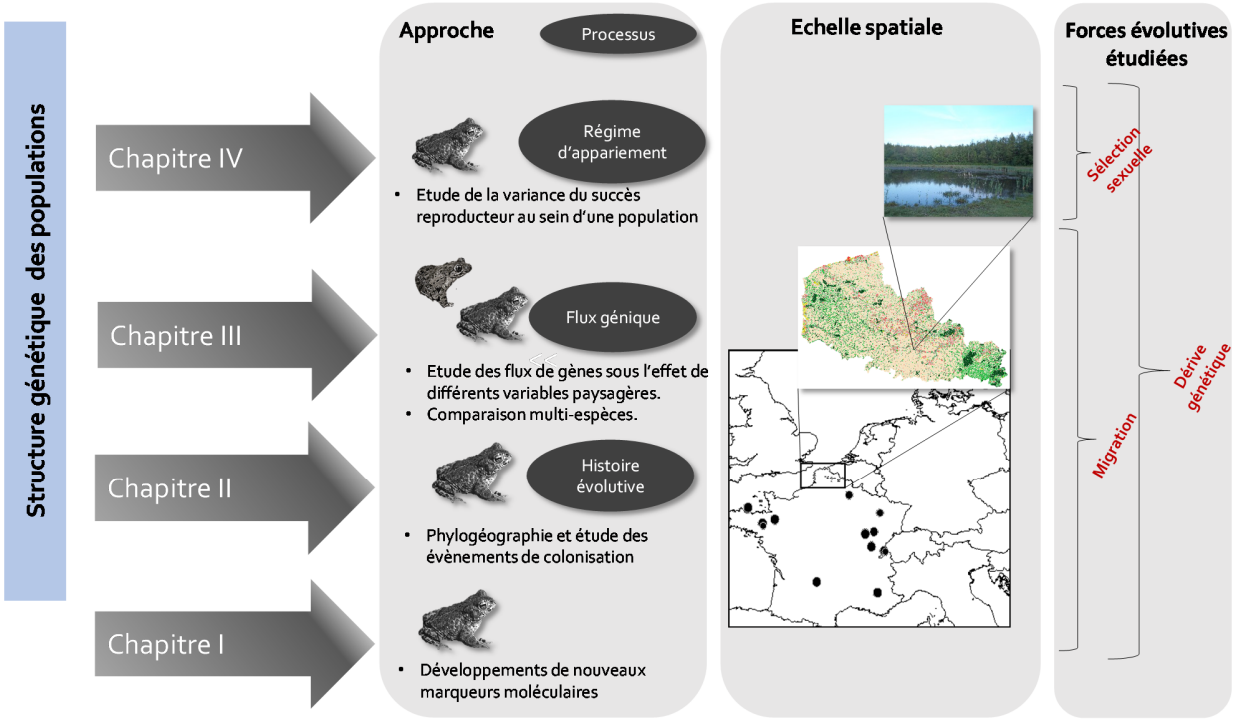


Figure 8 : Organigramme présentant les différents processus micro-évolutifs abordés dans ce travail de thèse à partir de l'étude de la structure génétique neutre des populations, ainsi que les échelles spatiales et forces évolutives qui y sous-tendent.

Références

Alberti M (2015) Eco-evolutionary dynamics in an urbanizing planet. *Trends in Ecology and Evolution*, **30**, 114–126.

Allendorf FW, Luikart GH, Aitken SN (2012) *Conservation and the Genetics of Populations*. Wiley-Blackwell, Oxford (UK).

Allentoft ME, O'Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, **2**, 47–71.

Arak A (1988) Female mate selection in the natterjack toad : active choice or passive attraction ? *Behavioral Ecology and Sociobiology*, **22**, 317–327.

Arens P, Sluis T Van Der, Van't Westende WPC *et al.* (2007) Genetic population differentiation and connectivity among fragmented Moor frog (*Rana arvalis*) populations in the Netherlands. *Landscape Ecology*, **22**, 1489–1500.

Arnold SJ, Wade MJ (1984) On the measurement of natural and sexual selection: theory. *Evolution*, **38**, 709–719.

Arntzen JW, Abrahams C, Meilink WRM, Iosif R, Zuiderwijk A (2017) Amphibian decline, pond loss and reduced population connectivity under agricultural intensification over a 38 year period. *Biodiversity and Conservation*, **26**, 1411–1430.

Beebee TJC (1996) *Ecology and conservation of amphibians*. Chapman & Hall, London.

Beebee TJC, Cabido C, Eggert C *et al.* (2012) 40 years of natterjack toad conservation in Europe. *FrogLog*, **101**, 40–44.

Beebee TJC, Denton JS (1996) *The natterjack toad conservation handbook (Species recovery program)*. English Nature, Northminster House, Peterborough PE1 1UA.

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

Beja P, Alcazar R (2003) Conservation of Mediterranean temporary ponds under agricultural intensification: an evaluation using amphibians. *Biological Conservation*, **114**, 317–326.

Bergstrom CT, Dugatkin LA (2012) *Evolution* (B Twitchell, Ed,). W. W. Norton & Company, Inc, New-York & Londres.

Blaustein AR, Wake DB, Sousa WP (1994) Amphibian declines : judging stability, persistence, and susceptibility of populations to local and global Extinctions. *Conservation Biology*, **8**, 60–71.

Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, **11**, 495–512.

Coster SS, Babbitt KJ, Cooper A, Kovach AI (2015) Limited influence of local and landscape factors on finescale gene flow in two pond-breeding amphibians. *Molecular Ecology*, **24**, 742–758.

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

Darwin C (1871) *The descent of man, and selection in relation to sex*. Princeton, N.J. : Princeton University Press, London.

Fahrig L (2017) Ecological responses to habitat fragmentation per se. *Annual Review of Ecology, Evolution, and Systematics*, **48**, 1-23.

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

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

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

Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of species extinction. In: *Conservation Biology* (ed Soulé ME), pp. 19–34. Sunderland, Massachusets.

Gutiérrez-Rodríguez J, Gonçalves J, Civantos E, Martínez-Solano I (2017) Comparative landscape genetics of pond-breeding amphibians in Mediterranean temporal wetlands: The positive role of structural heterogeneity in promoting gene flow. *Molecular Ecology*, **26**, 5407-5420.

Hamrick J. L., and M. J. Godt 1990 Allozyme diversity in plant species. Pp 43-63 in Brown, A. H. D., M. T. Clegg, A.

L. Kahler, and B. S. Weir, eds. Plant Population Genetics, Breeding, and Genetic Resources. Sinauer, Sunderland, MA.

Hamrick J. L., and M. J. Godt 1996 Conservation genetics of endemic plants. Pp 281-304 in Avise, J. C. and J. L. Hamrick, eds. Conservation genetics: Case Histories from Nature. Chapman and Hall, New York.

Howard RD, Moorman R, Whiteman H (1997) Differential effects of mate competition and mate choice on eastern tiger salamanders. *Animal behaviour*, **53**, 1345–56.

Hu S, Niu Z, Chen Y, Li L, Zhang H (2017) Global wetlands: potential distribution, wetland loss, and status. *Science of the Total Environment*, **586**, 319–327.

Jaquiéry J, Broquet T, Aguilar C, Evanno G, Perrin N (2010) Good genes drive female choice for mating partners in the lek-breeding European treefrog. *Evolution*, **64**, 108–115.

Johansson M, Primmer CR, Merilä J (2007) Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (*Rana temporaria*). *Molecular Ecology*, **16**, 2693–2700.

Jones AG (2009) On the opportunity for sexual selection, the bateman gradient and the maximum intensity of sexual selection. *Evolution*, **63**, 1673–1684.

Jones AG, Ratterman NL (2009) Mate choice and sexual selection: what have we learned since Darwin? *Pnas*, **106**.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Krebs CJ, Keller BL, Tamarin RH (1969) Microtus population biology: demographic changes in fluctuating populations of *M. ochrogaster* and *M. pennsylvanicus* in southern Indiana. *Ecology*, **50**, 587–607.

Lehmann L, Perrin N (2003) Inbreeding avoidance through kin recognition : choosy females boost male dispersal. *The American Naturalist*, **162**, 638–652.

Lemoine G (2012) Interêt, gestion, protection et valorisation des terrils miniers dans le Nord–Pas-de-Calais. *Annales de la Société Géologique du Nord*, **19**, 1–8.

Marsh DM, Trenham PC (2015) Metapopulation dynamics and amphibian conservation. *Conservation Biology*, **15**, 40–49.

Marsh JN, Vega-Trejo R, Jennions MD, Head ML (2017) Why does inbreeding reduce male paternity? Effects on sexually selected traits. *Evolution*, 71, 2728–2737.

McCauley DE, Raveill JAY, Antonovicst J (1995) Local founding events as determinants of genetic structure in a plant metapopulation. *Heredity*, **75**, 630–636.

Merilä J, Hendry AP (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, **7**, 1–14.

Miraldo A, Li S, Borregaard MK *et al.* (2016) An Anthropocene map of genetic diversity. *Science*, **353**, 1532–1535.

Møller AP, Danchin E (2008) Behavioural ecology and conservation. In: *Behavioural Ecology* (eds Danchin E, Giraldeau L-A, Cézilly F), pp. 647–664. Oxford University Press, Oxford, UK.

Muralidhar P, de Sá FP, Haddad CFB, Zamudio KR (2014) Kin-bias, breeding site selection and female fitness in a cannibalistic neotropical frog. *Molecular Ecology*, **23**, 453–463.

Noël S, Lapointe FJ (2010) Urban conservation genetics: study of a terrestrial salamander in the city. *Biological Conservation*, **143**, 2823–2831.

Pereira HM, Navarro LM, Martins IS (2012) Global biodiversity change: the bad, the good, and the unknown. *Annual Review of Environment and Resources*, **37**, 25–50.

Phillipsen IC, Funk WC, Hoffman EA, Monsen KJ, Blouin MS (2011) Comparative analyses of effective population size within and among species: ranid frogs as a case study. *Evolution*, **65**, 2927–2945.

Reed DH, Frankham R (2003) Correlation between Ffitness and genetic diversity. *Conservation Biology*, **17**, 230–237.

Rodrigues ASL, Andelman SJ, Bakarr MI *et al.* (2004) Effectiveness of the global protected area network in representing species diversity. *Nature*, **428**, 640–643.

Saarikivi J, Knopp T, Granroth A, Merilä J (2013) The role of golf courses in maintaining genetic connectivity between common frog (*Rana temporaria*) populations in an urban setting. *Conservation Genetics*, **14**, 1057–1064.

Salvidio S, Lamagni L, Bombi P, Bologna MA (2004) Distribution, ecology and conservation of the parsley frog

(*Pelodytes punctatus*) in Italy (Amphibia, Pelodytidae). *Italian Journal of Zoology*, **71**, 73–81.

Siddle H V., Kreiss A, Eldridge MDB *et al.* (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. *Proceedings of the National Academy of Sciences of the USA*, **104**, 16221–16226.

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 non-equilibrium populations. *Evolution*, **47**, 264–279.

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

Soulé M. E. (1987). Introduction. Pp. 1-10 in Soulé, M. E., ed. Viable Populations for Conservation. Sinauer, Sunderland, MA.

Soulé ME, Mills LS (1998) No need to isolate genetics. *Science*, **282**, 1658–1659.

Spear SF, Peterson CR, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology*, **14**, 2553–2564.

Stevens VM, Baguette M (2008) Importance of habitat quality and landscape connectivity for the persistence of endangered natterjack toads. *Conservation Biology*, **22**, 1194–204.

Stevens VM, Verkenne C, Vandewoestijne S, Wesselingh RA, Baguette M (2006) Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology*, **15**, 2333–2344.

Stevens VM, Wesselingh RA, Baguette M (2003) Demographic processe in a small, isolated population of Natterjack Toads (*Bufo calamita*) in southern Belgium. *Herpetological Journal*, **12**, 59–67.

Tejedo M (1988) Fighting for females in the toad *Bufo calamita* is affected by the operational sex ratio. *Animal Behaviour*, **36**, 1765–1769.

Tejedo M (1992) Large male mating advantage in natterjack toads, *Bufo calamita*: sexual selection or energetic constraints? *Animal Behaviour*, **44**, 557–569.

Trivers R (1972) Parental investment and sexual selection. *Sexual Selection and the Descent of Man*, **13**, 53–95.

Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJ (2001) Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity*, **86**, 598–608.

Wade MJ, McCauley DE (1988) Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution*, **42**, 995–1005.

Wake DB, Vredenburg VT (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11466–11473.

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Molecular Ecology*, **18**, 3847–56.

Wang IJ, Shaffer HB (2017) Population genetic and field-ecological analyses return similar estimates of dispersal over space and time in an endangered amphibian. *Evolutionary Applications*, **10**, 630–639.

Ward, R. D., D. O. F. Skibinsk, and M. Woodwark. 1992. Protein heterozity, protein structure, and taxonomic differentiation. *Evolutionary Biology* **26**, 73-160.

Wells KD (2010) *The ecology and behavior of amphibians*. University of Chicago Press, Chicago, USA.

Wright S (1931) Evolution in Mendelian Populations. *Genetics*, **16**, 97–159.

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

Youngquist MB, Inoue K, Berg DJ, Boone MD (2017) Effects of land use on population presence and genetic structure of an amphibian in an agricultural landscape. *Landscape Ecology*, **32**,147-162.

CHAPITRE I

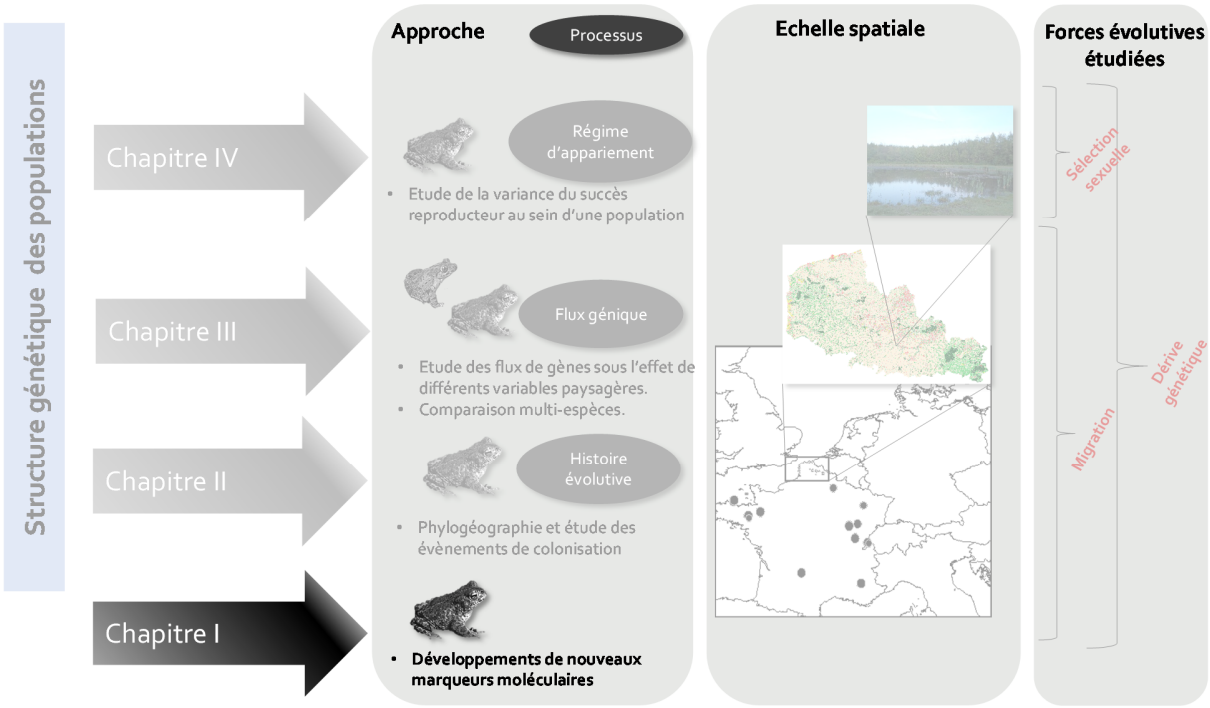
Développement de nouveaux marqueurs nucléaires et mitochondriaux pour l'étude de la structure génétique des populations de *Bufo* [*Epidalea*] *calamita.*

Les marqueurs moléculaires sont aujourd'hui utilisés en routine en génétique et en écologie évolutive. Ils constituent un outil puissant pour inférer les processus micro évolutifs classiquement étudiés en génétique des populations : les taux de migration, la taille efficace des populations, les goulots d'étranglement ou encore les niveaux d'apparentement génétique entre individus (Slatkin 1985; McCartney-Melstad & Shaffer 2015; Ellegren & Galtier 2016). Dans cet article, nous présentons des marqueurs moléculaires développés dans le cadre de cette thèse afin d'établir précisément les génotypes multilocus nucléaires et les lignées maternelles mitochondriale chez le Crapaud calamite : 22 nouveaux marqueurs nucléaires microsatellites et 18 marqueurs KASPar de polymorphisme mononucléotidique (SNP, Single Nucleotide Polymorphism) localisés dans des portions de régions du 16S et de la D-loop du génome mitochondrial.

Les séquences microsatellites sont des répétitions en tandem de motifs de 2 à 6 nucléotides et sont largement répandues dans le génome de la plupart des taxons. Leur taux de mutation est souvent élevé, en moyenne $5x10^{-4}$ par locus et par génération, induisant un fort niveau de diversité génétique qui peut être exploité pour étudier des processus évolutifs ayant lieu sur des petites échelles de temps (Balloux & Lugon-Moulin 2002; Selkoe & Toonen 2006). Trois études avaient précédemment publié des marqueurs microsatellites nucléaires pour le Crapaud calamite. Le polymorphisme de ces loci était assez faible et leur amplification présente un fort taux d'échec. C'est pourquoi de nouveaux marqueurs microsatellites ont été développés pour ce travail de thèse afin d'améliorer la puissance statistique (i) des estimations de flux de gènes et les inférences sur les évènements démographiques passés (Chap II, III) et (ii) des tests d'assignements de parenté et d'analyse de consanguinité individuelle(Chap IV) (Landguth *et al.* 2012).

Les marqueurs mitrochondiaux de type SNPs, quant à eux, permettent de déterminer le polymorphisme des individus sur une mutation ponctuelle. Le taux de mutation des SNPs varie selon l'espèce et la région du génome considérée. Beaucoup de SNPs n'ont pas d'implication fonctionnelle et permettent d'estimer les mêmes paramètres que les marqueurs microsatellites. Les SNPs ont l'avantage de pouvoir être traité très rapidement, notamment avec la méthode KASPar, ce qui permet une acquisition rapide de données génotypiques. Le génome mitochondrial présente un taux de mutation inférieur au génome nucléaire et a une hérédité exclusivement maternelle. En développant des SNPs pour explorer le polymorphisme porté par le génome mitochondrial au sein des populations, nous avons pu décrire une structure génétique des populations liée à l'existence de plusieurs lignées maternelles et résultant d'évènements d'ordre biogéographique (Chap II).

Ce chapitre de thèse a fait l'objet d'une publication (Faucher *et al.* 2016).



Références

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155–65.

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

Faucher L, Godé C, Arnaud J-F (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo* (*Epidalea*) *calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). *Journal of Heredity*, **107**, 660–665.

Landguth EL, Fedy BC, Cance SJOC, Garey AL (2012) Effects of sample size , number of markers , and allelic richness on the detection of spatial genetic pattern. , 276–284.

McCartney-Melstad E, Shaffer HB (2015) Amphibian molecular ecology and how it has informed conservation. *Molecular Ecology*, **24**, 5084–5109.

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

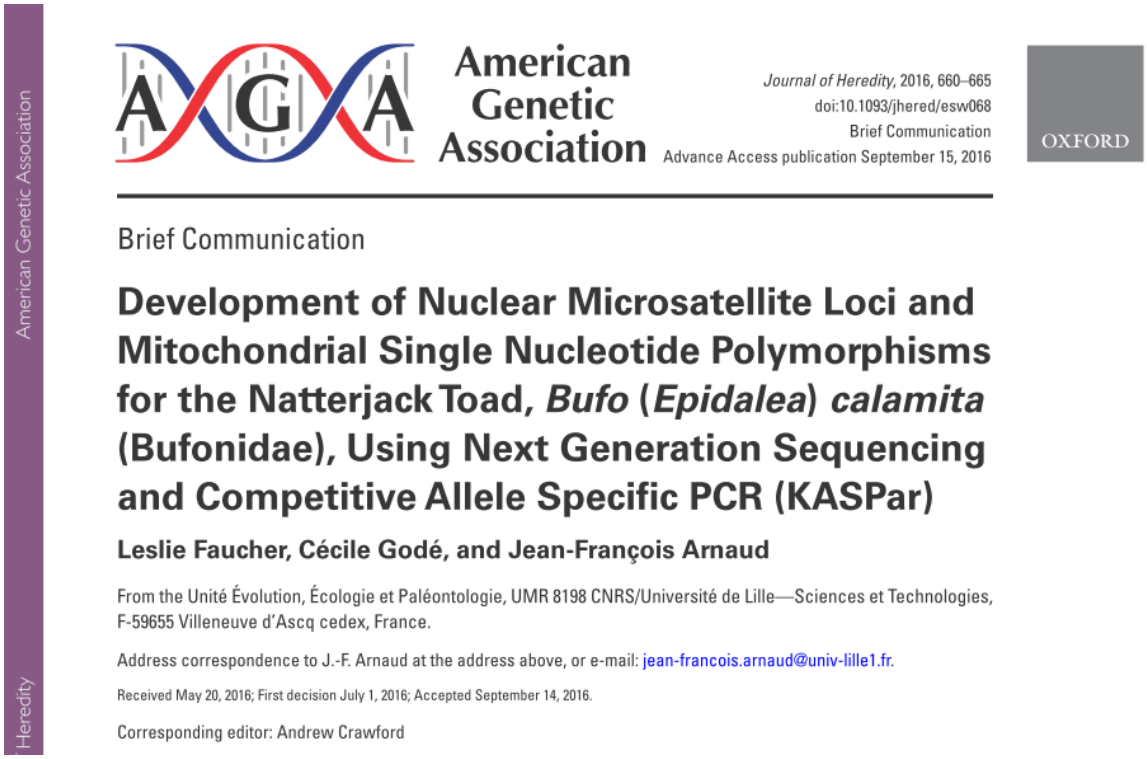
Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology, Evolution, and Systematics*, **16**, 393–430.

DEVELOPMENT OF NUCLEAR MICROSATELLITE LOCI AND MITOCHONDRIAL SINGLE NUCLEOTIDE POLYMORPHISMS FOR THE NATTERJACK TOAD, *BUFO* (*EPIDALEA*) *CALAMITA* (BUFONIDAE), USING NEXT GENERATION SEQUENCING AND COMPETITIVE ALLELE SPECIFIC PCR (KASPAR).

Leslie Faucher, Cécile Godé and Jean-François Arnaud*

Unité Évolution, Écologie et Paléontologie, UMR 8198 CNRS / Université de Lille – Sciences et Technologies, F-59655 Villeneuve d'Ascq cedex, France
* Corresponding author
E-mail: jean-francois.arnaud@univ-lille1.fr

Subject areas: Population structure and phylogeography; conservation genetics and biodiversity
Keywords: Amphibians, nuclear microsatellites, mtDNA SNPs, Kompetitive allele-specific PCR, conservation genetics, genetic diversity, parentage, *Bufo calamita*



Abstract

Amphibians are undergoing a major decline worldwide and the steady increase in the number of threatened species in this particular taxa highlights the need for conservation genetics studies using high-quality molecular markers. The natterjack toad, *Bufo* (*Epidalea) calamita*, is a vulnerable pioneering species confined to specialized habitats in Western Europe. To provide efficient and cost-effective genetic resources for conservation biologists, we developed and characterized 22 new nuclear microsatellite markers using next-generation sequencing. We also used sequence data acquired from Sanger sequencing to develop the first mitochondrial markers for KASPar assay genotyping. Genetic polymorphism was then analyzed for 95 toads sampled from five populations in France. For polymorphic microsatellite loci, number of alleles and expected heterozygosity ranged from 2 to 14 and from 0.035 to 0.720, respectively. No significant departures from panmixia were observed (mean multilocus F_{IS}=-0.015) and population differentiation was substantial (mean multilocus F_{ST}=0.222, $P < 0.001$). From a set of 18 mitochondrial SNPs located in the 16S and D-loop region, we further developed a fast and cost-effective SNP genotyping method based on competitive allele-specific PCR amplification (KASPar). The combination of allelic states for these mtDNA SNP markers yielded 10 different haplotypes, ranging from 2 to 5 within populations. Populations were highly differentiated ($G_{ST} = 0.407$, $P < 0.001$). These new genetic resources will facilitate future parentage, population genetics and phylogeographical studies and will be useful for both evolutionary and conservation concerns, especially for the set-up of management strategies and the definition of distinct evolutionary significant units.

Introduction

Amphibians have dramatically declined worldwide over the past 30 years (Beebee and Griffiths 2005). One of the major processes currently acknowledged to be involved in the increase of the number of threatened amphibian species is landscape fragmentation and associated habitat loss. The ensuing genetic erosion due to decreasing gene flow among populations and reduced effective population sizes lowers the evolutionary potential of populations (Allentoft and O'Brien 2010). Hence, conservation genetics studies of amphibians are essential for establishing conservation priorities and for the delineation of management units with adequate evolutionary capabilities (Beebee 2005; Igawa et al. 2015; MacCartney-Melstad et al. 2015).

The natterjack toad, *Bufo (Epidalea) calamita* (Laurenti, 1768) (see Dubois & Bour 2010 for recent changes in taxonomical status), has a wide geographical distribution range, extending from continental southern Europe (Spain and Portugal) to north-western Europe (Denmark and southern Sweden) and including Ireland and United Kingdom (Rowe et al. 2006). This emblematic pioneering species is confined to temporary ponds in highly specialized habitats as marshes, coastal sand dunes or heathland habitats (Beebee and Denton 1996). In spite of its large geographical distribution, *B. calamita* is a highly protected species classified as rare and vulnerable, and suffering from the loss of suitable breeding ponds and favorable connected habitats, especially in the northern parts of its range (Allentoft et al. 2009). Consequently, this species now often inhabits human-made areas as substitution habitats, such as quarries, and is subject to numerous local conservation actions devoted to protect and reinforce isolated and declining populations (Beebee and Denton 1996; Godin 2002; Flavenot et al. 2015).

Nuclear microsatellite loci and mtDNA markers are powerful tools to get insights into patterns of dispersal events among populations, to delineate maternal lineages and to identify evolutionary significant units for the set-up of subsequent conservation management strategies (Slatkin 1995; Frankham et al. 2010; Allendorf et al. 2013; Ritchie et al. 2016; Wirtz et al. 2016). The natterjack toad was among the first European amphibian for which microsatellite markers were isolated for population genetics and phylogeographical inferences (see Rowe et al. 1997, 1998, 2000, 2006; Rogell et al. 2005). To gain further insights into fine-scaled population structure and individual siring success in the natterjack toad, this study aims at (i) isolating and screening an additional set of nuclear microsatellite markers of high resolution using high-throughput sequencing techniques and at (ii) developing fast and cost-effective mitochondrial DNA (mtDNA) markers using the KBiosciences competitive allele-specific PCR

amplification (KASPar^{TM}) of target SNPs. The KASPar method is simple and cost-effective as compared to other SNP genotyping assays: it allows bi-allelic scoring of SNPs through competitive binding of two allele-specific forward primers (Cuenca et al. 2013; Semagn et al. 2014). We successfully isolated and characterized 22 nuclear microsatellite markers and 18 mitochondrial SNP markers suitable to shed light on the intensity of gene flow among subdivided populations and to delineate maternal lineages of conservation relevance for remnant populations.

Material and methods

DNA extraction

A non-destructive sampling method was used. Plain sterile 15SC Copan (Brescia Italia) swabs were used to collect buccal cells for *B. calamita* individuals which were released after sampling. Swabs were dried prior DNA extraction. Total genomic DNA was isolated from these buccal swabs using the Macherey Nagel (Düren, Germany) NucleoSpin® 96 trace kits following the manufacturer instructions.

Isolation of nuclear microsatellite loci

Total genomic DNA was sent to the GENOSCREEN genomic platform, Lille, France (www.genoscreen.fr). A random pool of eight toads was chosen for Shotgun sequencing of a 55.2*ng* genomic library. A high-throughput method for isolating high-quality microsatellite markers for non-model organisms was then used as described in Malausa et al. (2011). By coupling next-generation sequencing and multiplex microsatellite enrichment, 1μg of genomic DNA was used for the development of microsatellite libraries through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries, as in Favre-Bac et al. (2014) and Poux et al. (2015). GS-FLX libraries were realized according to manufacturer's protocols and sequenced on a GS-FLX Titanium PicoTiterPlate (Roche Diagnostics, Mannheim, Germany). Of 26136 randomly fragmented sequences, 47 loci with the longest repeat sequences were initially tested for further development and routine genotyping on eight individuals of *B. calamita*.

Nuclear microsatellite loci genotyping

PCR reactions were performed in optimized cost-effective multiplexes that successfully amplified different sets of microsatellite loci in a single PCR reaction (see Table 1). Forward primers for the selected loci were labelled with PET™, NED™, 6-FAM™ or HEX™ fluorescent dye (Applied biosystem, Foster City, California, USA). We used 10μl volume

containing 1μl (5-20ng) of genomic DNA, 1X multiplex PCR master mix (QIAGEN Hilden, Germany), and 0.1-0.3μM of labelled forward and reverse primer. The PCR cycling program had an initial denaturation of 95°C for 15 min; 35 cycles of 94°C for 30s, annealing temperature of 55°C for 1min 30s, and 72°C for 1min; and a final extension at 60°C for 30 min. PCR was conducted on a Eppendorf Mastercycler pro 384 (Applied Biosystem, Foster City, California, USA). 1.5μl of PCR product were pooled with 0.25μl of GeneScan 500 LIZ size standard (Applied Biosystems) and 9.75μl of deionized formamide (Applied Biosystems). The PCR amplicons were subsequently electrophoresed and sized using an ABI PRISM 3130 Sequencer (Applied Biosystems) and the software GeneMapper version 5 (Applied Biosystems), respectively. Individuals with failed amplifications or dubious genotypes were checked with a second round of genotyping.

Mitochondrial DNA acquisition from Sanger sequencing

In addition to these nuclear microsatellite loci, we further identified mtDNA single nucleotide polymorphisms (SNPs) from mitochondrial sequence data acquired from Sanger sequencing. These mtDNA SNPs were located in the 16S and D-loop region (see GenBank accessions in Table S1). Based on the entire mitochondrial genome of a related species (*Bufo japonica*, GenBank accession number: AB303363.1) described in Igawa et al. (2008), two mtDNA regions were selected for Sanger sequencing by adapting universal primers. The D-loop, situated in the control region, is known to be the most variable region along the circular mitochondrial genome (Vanbrabant et al., 2009). The 16S region covers conserved coding and hypervariable non-coding regions. The set of slightly modified universal primers were as follows: 16sar_F (CGCCTGTTTACCAAAAACAT') and 16sbr_R (CCGGTCTGAACTCA-GATCACGT), LX12SN1_F (TACACACCGCCCGTCA) and LX16S1R_R (GACCTGG-ATTTCTCCGGTCTGAACTC), LX16S1_F (GGTTTA-CGACCTCGATGTTGGATCA) and Met3850H_R (GGTATGGGCCCAAAAGCTT), adding up to 3592bp for the 16S region; Control BH_R (GTCCATTGGAGATTAAGATCTACCA) and ControlWrev-L_F (GACA-TACTATGTATAATCGAGCAT), CytA-L_F (GAAT-CGGGGGTCAACCAGTAGAAGA-CCC) and ControlK-H_R (AATGGT-CAAAATGGCTGAGATTG) adding up to 1752bp for the D-loop region. PCR amplifications were performed in a Eppendorf Mastercycler pro 384 using 25μl of mix comprising 1X of Buffer 10X (Dream *Taq* Buffer, ThermoFisher Scientific, Waltham, MA USA), 2.5mM of MgCl2, 2μl (2.5mM) of dNTP, 0.2mM of BSA, 0.2μM) of each primer, 0.652U of Dream *Taq* polymerase (ThermoFisher Scientific) and 5μl (5-20ng) of template DNA. PCR conditions were as follow: 95°C for 5min; 40 cycles of 94°C for 30s,

annealing temperature of 55°C for 1min, and 72°C for 1min 30s; and a final extension step at 72°C for 10min. Raw data of sequences were read, verified and aligned using CODONCODE Aligner v. 5.1.5 (CODONCODE Corporation, Centerville, MA, USA) and BIOEDIT v. 7.2.5 (Hall 1999). Polymorphic SNPs were confirmed by a second round of PCR amplification with independent sequencing. A total of 94 individuals from different French localities were used for this first step of direct mitochondrial sequencing.

MtDNA SNPs genotyping using the KASPar methodology

We describe here an accessible and cost-effective SNP genotyping method based on KBioscience's competitive allele-specific PCR amplification of target sequences and endpoint fluorescence genotyping (KASPar™) using a LightCycler 480 (Roche Diagnostics, Mannheim, Germany). From the initial direct sequencing, 18 polymorphic SNPs were identified and used for competitive allele-specific PCR amplification, see Table S1. SNP primers design for KASPar genotyping were defined by LGC group (http://www.lgcgroup.com/genomics) from each SNP locus flanking sequence. Two allele-specific oligonucleotides and one common oligonucleotide were defined for each locus (Table S1). The KASPar system applies two fluorescence resonance energy transfer (FRET) cassettes, where fluorometric dye is conjugated to the primer but quenched via resonance energy transfer (reviewed in Semagn et al. 2014). KASPar genotyping is based on competitive allele-specific PCR and enables bi-allelic scoring of single nucleotide polymorphisms (e.g. Cuenca et al. 2013; Martin et al. 2016). The SNP-specific KASP Assay mix (designed by LGC group, www.lgcgroup.com) and the universal KASP Master Mix were added to DNA samples. A thermal cycling reaction was then carried out, followed by an end-point fluorescent read. Allelic discrimination was finally completed through competitive annealing of two allele-specific forward primers, each having a unique tail sequence corresponding to a distinct labelled FRET cassette in the Master Mix, one labelled with FAM™ dye and the other with HEX™ dye. Sequences flanking each SNPs and each specific and common primers that were used are detailed in Table S1.

PCR amplification reactions were performed using either the Eppendorf Mastercycler pro 384 (Applied Biosystem, Foster City, California, USA) or the LightCycler 480 (Roche Diagnostics, Mannheim, Germany), with $4\mu l$ (5-20ng) of genomic DNA in $8\mu l$ reaction volume containing 1X of universal KASP Master mix (2X) and $0.11\mu l$ of the SNP-specific KASP Assay mix. This SNP-specific KASP Assay mix was composed of the two forward primers for allele X and Y, labelled in FAM^{TM} and HEX^{TM} respectively, and the common reverse primer, see Table S1. Cycling conditions for PCR amplifications were the following for all mtDNA SNPs

but one: 15 min at 94°C, 10 cycles of 20s at 94°C and 60s at 65°C-57°C with a drop of 0.8°C per cycle, 40 cycles of 20s at 94°C and 60s at 57°C, ending by 12°C for 10 min. Cycling conditions for the remaining mtDNA SNP (Bc_ControlR_9_T/C) were as follow: 15 min at 94°C, 10 cycles of 20s at 94°C and 60s at 61°C-55°C with a drop of 0.6°C per cycle, 40 cycles of 20s at 94°C and 60s at 55°C, and finally 12°C for 10 min. Fluorescence was detected using a LightCycler® 480 (Roche Diagnostics, Mannheim, Germany) and automatic allele calls were checked using the LightCycler 480 SW1.5.0 SP3 version 1.5.0.39 software (Roche Diagnostics GmbH, Mannheim, Germany). Lastly, as mitochondrial genome is inherited as a single linkage unit, individual mitochondrial haplotypes were defined as the combination of allelic states for the 18 detected mtDNA polymorphic SNP markers.

Table 1 Locus name, primer sequence (5'-3'), repeat motif from the original sequence, allelic size range (bp), multiplex number, fluorescent dye used and the GenBank accession number for 22 newly isolated microsatellite loci in the natterjack toad *Bufo* (*Epidalea) calamita*. Also presented are the mean measures of genetic diversity and genetic differentiation estimated over five populations (totalizing $N = 95$ toads): the total number of allele (A_n), the mean observed (H_o) and expected (H_e) heterozygosity, the mean intra-population fixation index (F_{IS}) and the mean population genetic differentiation (F_{ST} and R_{ST}). *NS*: non-significant; ***: $P < 0.001$. – : not calculable.

Locus name	Primer sequences (5'-3')	Repeat motif	Allelic size range	Multiplex number	Dye	GenBank Accession no.	A_n	H_o	H_e	F_{IS}	F_{ST}	R_{ST}
BC05	F: CATTGATATGGCTGCCAACTT R: CATGGGGATCAATGGCTACT	$(GT)_{13}$	106-116	1	NED	KX237573	4	0.548	0.519	-0.049^{NS}	0.275***	0.391***
BC08	F: CTCTTGTGCAAGATCTCTGGG R: TACTGACTGCTGCCCTCTCC	$(TAGA)_{11}$	241-279	1	HEX	KX237574	14	0.642	0.654	0.005^{NS}	0.137***	0.193***
BC11	F: AGCCTTCTTTGCATCACTGC R: TAGCGGGAAGAGATGTACGC	$(GATA)_{11}$	128-158	1	HEX	KX237575	8	0.581	0.594	0.035^{NS}	0.319***	0.320***
BC19	F: CCAAGGAAGAAACTGTGGCA R: AACATACATACACTCACACCCACA	$(TG)_{10}$	170-174	1	FAM	KX237576	3	0.502	0.451	-0.112^{NS}	0.365***	0.457***
BC22	F: TGCAGATTGCCAGCAGTTTA R: CACTTCCTCAAGGTGGTGCT	$(GATA)_{9}$	314-339	1	FAM	KX237577	7	0.696	0.72	0.074^{NS}	0.137***	0.243***
BC29	F: GTTGCGACTGGGGAAATAAC R: GCTTCACAAGACATGCAGGA	$(ATCT)_{9}$	172-188	1	NED	KX237578	5	0.55	0.602	0.032^{NS}	0.198***	0.177***
BC38	F: TACAGTTAAGGACCCCGTCG R: GGCCACTGTCCGTGGTTAC	$(CA)_{8}$	251	1	PET	KX237579	1	-	-	-	-	-
BC46	F: TGAATAGACAGACATTTGTCCAAGA R: TTCTACCGGTCAACCTATCCA	$(AGAT)_{8}$	111-131	1	FAM	KX237580	6	0.667	0.667	0.016^{NS}	0.155***	0.265***
BC01	F: TCCATAATCAGGCGCTCATA R: TCTATTCTCTTAAACCGGAGAGG	$(TAC)_{16}$	85-127	2	FAM	KX237581	9	0.564	0.519	-0.082^{NS}	0.177***	0.210***
BC02	F: TTGCTTGAGAAAAGTCCAACA R: ACTTGCCAACTCTCCCAGAA	$(GATA)_{14}$	191-218	2	FAM	KX237585	8	0.603	0.68	0.081^{NS}	0.166***	0.496***
BC04	F: TGCTCCTGACAATTAACTTTGG R: ATCTGTGTCAGGGCATCTCC	$(CA)_{13}$	132-143	2	NED	KX237584	6	0.63	0.597	-0.082^{NS}	0.178***	0.068***
BC24	F: ACGGTTTTCTGAAGCAATGG R: GCATGTGCAGAAGACTTCAAA	$(AC)_{9}$	206-213	2	NED	KX237586	4	0.61	0.623	0.034^{NS}	0.193***	0.271***
BC25	F: CAGTTGTTTTCCGAGGTGGT R: AAGGAAGCTGAATTTTGGTTGA	$(CT)_{9}$	201-207	2	PET	KX237587	3	0.187	0.178	-0.058^{NS}	0.450***	0.368***
BC28	F: ACTTTGGCAAGGGAACCAG R: TTGTCAGTTAACCAGCGTGC	$(GA)_{9}$	175-183	2	HEX	KX237588	3	0.487	0.432	-0.172^{NS}	0.242***	0.155***
BC34	F: TGCATAGCCTTGTGAAGCTG R: GTGACACCATGTCCCAGATG	$(AC)_{9}$	108-112	2	HEX	KX237582	3	0.381	0.361	-0.033^{NS}	0.424***	0.593***
BC35	F: GGGTATGGTGGCTGGTAATG R: TCACAAGTAGCCACAGTAAAGGA	$(TC)_{9}$	108-110	2	PET	KX237583	2	0.035	0.035	-0.021^{NS}	0.013NS	0.013NS
BC15	F: TGCTCCTCAAGTGTTGTTGG R: TGGGACGACAGGAACGTACT	$(AG)_{11}$	89-99	3	FAM	KX237589	4	0.222	0.18	-0.258^{NS}	0.204***	0.151***
BC18	F: CCTTAATGGCCCAAGCCTAT R: AGACAGGGATGGATAGATGGA	$(ATCT)_{10}$	175-191	3	FAM	KX237590	5	0.682	0.659	0.003^{NS}	0.135***	0.094***
BC37	F: TCACCTGTACCCCTCTGGG R: CCATCCATGACACAGACCAG	$(ATCT)_{9}$	87-116	3	HEX	KX237591	8	0.693	0.711	-0.015^{NS}	0.130***	0.243***
BC39	F: TCTGTCCTTCTGTCCAATCTG R: GCACCTTTGTTCAGGATGGT	$(TCTA)_{8}$	167-195	3	HEX	KX237592	8	0.69	0.64	-0.049^{NS}	0.142***	0.310***
BC09	F: GGTGGTGGCACATTTCTTTT R: GTAGTTTGCCAGCAATGCCT	$(TAGA)_{11}$	237-273	3	NED	KX237593	7	0.714	0.67	-0.058^{NS}	0.222***	0.314***
BC45	F: CCCTTGCAGCCAAAATAAAA R: TAACAGGAAACGGATTTGGG	$(TAGA)_{8}$	118-156	3	PET	KX237594	9	0.583	0.592	0.047^{NS}	0.338***	0.300***
Mean over all loci							5.77	0.509	0.504	-0.015^{NS}	0.222***	0.273***

Population sampling and statistical analyses

We tested the polymorphism of isolated suitable microsatellite markers and we validated the 18 SNP assays using KASPar genotyping chemistry on a diversity panel of 95 individuals coming from five geographically distinct populations located in France (Figure S1). Names, sample sizes and coordinates (WGS84) of sampled populations are the following: "North coastline" (*N*=20; Latitude: 1.5832, Longitude: 50.4869), "North mining area" (*N*=30, Lat.: 2.7967, Long.: 50.4038), "Picardie" (*N*=13; Lat.: 3.9334, Long.: 49.5534), "Lorraine" (*N*=10, Lat.: 6.1050, Long.: 48.5509) and "Brittany" (*N*=22; Lat.: -1.9920, Long.: 48.0160). Classic parameters of genetic diversity and genetic differentiation across the five surveyed populations were estimated for each loci using FSTAT version 2.9.3 (Goudet 1995).

Results and Discussion

Of the 47 nuclear microsatellite loci initially tested, 22 had easily readable PCR amplification products with no stutter peaks or dubious electrophenograms. 21 markers out of the 22 tested were polymorphic. In order to reduce genotyping costs and efficiency, primer pairs were successfully combined into three multiplexes, each including six to height markers (Table 1). Out of 210 comparisons and after Bonferroni corrections, no linkage disequilibrium was observed for any pairs of loci. Moderate to high polymorphism was observed with clear sizing PCR products. Excluding the locus BC38 which was monomorphic, the observed number of alleles varied from 2 to 14 across loci (mean = 5.77), for a total of 127 alleles on the studied samples (Table 1). The mean observed (*Ho*) and mean expected heterozygosity (*He*) values ranged from 0.035 to 0.714 and from 0.035 to 0.720, respectively (mean = 0.509 and 0.504). Mean F_{IS} estimates across populations and across loci (intra-population fixation index, measuring departures from panmixia) were all conform to Hardy-Weinberg expectations (*i.e.* not statistically significantly different from zero), with F_{IS} estimates varying from -0.258 to 0.074 for a mean multilocus value of -0.015. Except for locus BC35, most loci detected population genetic structuring as F_{ST} estimates (inter-population fixation index, measuring the extent of population genetic differentiation) were highly significant (all at $P < 0.001$) and ranged from 0.130 to 0.450 with a mean multilocus estimate of 0.222. Estimates of R_{ST}, a similar statistic based on models which take into account the microsatellite evolution, especially the variance in allele length (see Slatkin 1995), were also all significant (ranging from 0.094 to 0.593, with a mean estimate of 0.273) making these microsatellite loci markers excellent tools to detect phylogeographical signals.

KASPar genotyping data conformed to previous Sanger sequencing data and, out of the 18 mtDNA SNP primers tested, only two were monomorphic (Bc_ControlR_6_T/C and Bc_ControlR_19_C/A) across the five surveyed populations. The combination of allelic states for polymorphic mtDNA SNPs yielded a total of ten different haplotypes. The within-population number of mtDNA haplotypes ranged from 2 to 5 (mean = 3), toad populations located in human-modified habitats ("North mining area" and "Picardie" populations) being the most diverse (Figure S1). MtDNA genetic differentiation among populations was highly significant, with a G_{ST} estimate (estimator designed for mitochondrial variation and strongly related to F_{ST} index, see Pons and Petit 1996) of 0.407 ($P < 0.001$). For each mtDNA haplotype, the Table S2 and the Figure S1 give (i) their detailed allelic combinations, (ii) their geographical distributions over the surveyed populations, and (iii) their genealogical relationships based on a median-joining network. Finally, the KASPar technology proved to be very fast and cost-effective as compared to classical microsatellite genotyping and direct Sanger sequencing. Indeed, the genotyping cost per individual was over four fold and 14 fold cheaper for the KASPar SNP genotyping (runtime of 2 min.) than for microsatellite genotyping (runtime of 45 min.) and direct Sanger sequencing (runtime of 2h30), respectively.

Overall, the newly developed nuclear microsatellite markers show high levels of polymorphism and can be applied for low-cost high throughput genotyping. The same holds for mtDNA SNP markers that exhibited different maternal lineages within populations. Therefore, they will be markers of choice for fine and large-scale population genetics studies designed to infer inbreeding levels, effective population sizes and patterns of gene flow among populations, and to trace back phylogeographical patterns. In addition, they are proving useful for ongoing parentage analyses assessing male and female reproductive success, which is of immediate value for monitoring and conservation actions.

Acknowledgments

We wish to thank Julie Jaquiery, Stéphanie Rondel, Alexandra Janczak and Loic Coquel for invaluable help in sampling populations, Céline Poux for technical advices on the choice of mtDNA sequences and Isabelle De Cauwer for critical reading on an early version of this manuscript. We also express our gratitude to two anonymous Reviewers and the Associate Editor whose suggestions greatly improved the manuscript. This work was funded by the "Région Nord–Pas-de-Calais" (AMPHIDIV project).

Data Availability

Nuclear DNA sequences used to design the 22 microsatellite primers have been assigned to the following GenBank numbers: KX237573- KX237594. Mitochondrial DNA sequences used to define the 18 polymorphic SNPs have been assigned to the following GenBank number: KX237595- KX237630. Microsatellite and mtDNA genotypes are deposited in the Dryad repository: doi: 10.1093/jhered/esw068.

References

Allendorf FW, Luikart G, Aitken SN. 2013. Conservation and the genetics of populations. John Wiley & Sons, Ltd, Chichester, West Sussex, UK.

Allentoft ME, O'Brien J. 2010. Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*. 2: 47-71.

Allentoft ME, Siegismund HR, Briggs L, Andersen LW. 2009. Microsatellite analysis of the natterjack toad (*Bufo calamita*) in Denmark: populations are islands in a fragmented landscape. *Cons. Genet.* 10:15-28.

Beebee TJC, Denton J. 1996. The Natterjack Toad, Conservation Handbook. Communications and Grants Team, English Nature, Northminster House, Peterborough, UK.

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? *Biol. Cons.* 125:271-285.

Cuenca J, Aleza P, Navarro L, Ollitrault O. 2013. Assignment of SNP allelic configuration in polyploids using competitive allele-specific PCR: application to citrus triploid progeny. *Ann. Bot. (London)*. 111:731-742.

Dubois A., Bour R. 2010. The nomenclatural status of the nomina of amphibians and reptiles created by Garsault (1764), with a parsimonious solution to an old nomenclatural problem regarding the genus *Bufo* (Amphibia, Anura), comments on the taxonomy of this genus, and comments on some nomina created by Laurenti (1768). *Zootaxa*. 2447:1-52.

Igawa T, Nozawa M, Nagaoka M, Komaki S, Oumi S, Fujii T, Sumida M. 2015. Microsatellite marker development by multiplex ion Torrent PGM sequencing: a case study of the endangered *Odorrana narina* complex of frogs. *J. Hered.* 106:131-137.

Favre-Bac L, Godé C, Arnaud J-F. 2014. Characterization of polymorphic microsatellite markers for the fine-leaved water-Dropwort *Oenanthe aquatica* and the Gypsywort *Lycopus europaeus*, two farmland remnant wetland species. *Cons. Genet. Res.* 6:995-998.

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

Frankham R, Ballou JD, Briscoe DA. 2010. Introduction to conservation genetics, 2nd edition. Cambridge University Press, Cambridge, UK.

Godin J. 2002. Degré de rareté, évolution de la distribution et particularités de l'herpétofaune de la Région Nord-Pas de Calais. *Bull. Soc. Herp. Fr.* 104:16-35.

Goudet J. 1995. FSTAT (Version 1.2). A computer program to calculate *F*-Statistics. *J. Hered.* 86:485-486.

Hall TA. 1999. BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95-98.

Igawa T, Kurabayashi A, Usuki C, Fujii T, Sumida M. 2008. Complete mitochondrial genomes of three neobatrachian anurans : a case study of divergence time estimation using different data and calibration settings. *Gene.* 407:116-129.

Malausa T, Gilles A, Meglécz E, Blanquart H, Duthoy S, Costedoat C, Dubut V, Pech N, Castagnone-Sereno P, Délye C, Feau N, Frey P, Gauthier P, Guillemaud T, Hazard L, Le Corre V, Lung-Escarmant B, Malé P-J, Ferreira S, Martin J-F. 2011. High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Mol. Ecol. Res.* 11:638-644.

Martin H, Touzet P, Van Rossum F, Delalande D, Arnaud J-F. 2016. Phylogeographic pattern of range expansion provides evidence for cryptic species lineages in *Silene nutans* in Western Europe. *Heredity.* 116:286-294.

McCartney-Melstad E, Shaffer HB. 2015. Amphibian molecular ecology and how it has informed conservation. *Mol. Ecol.* 24:5084-5109.

Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics.* 144:1237-1245.

Poux C, Gothland M, Holl A-C, Spilmont N, Arnaud J-F. 2015. Highly polymorphic microsatellite markers in two species, the invasive shore crabs *Hemigrapsus sanguineus* and *Hemigrapsus takanoi* (Decapoda, Varunidae). *Cons. Genet. Res.* 7:569-572.

Ritchie H, Jamieson AJ, Piertney SB. 2016. Isolation and characterization of microsatellite DNA markers in the deep-sea amphipod *Paralicella tenuipes* by Illumina Miseq sequencing. *J. Hered.* 107: 367-371.

Rogell B, Gyllenstrand N, Höglund J. 2005. Six polymorphic microsatellite loci in the natterjack toad, *Bufo calamita*. *Mol. Ecol. Notes.* 5:639-640.

Rowe G, Beebee TJC, Burke T. 1997. PCR primers for polymorphic microsatellite loci in the anuran amphibian *Bufo calamita*. *Mol. Ecol.* 6:401-402.

Rowe G, Beebee TJC, Burke T. 1998. Phylogeography of the natterjack toad *Bufo calamita* in Britain: genetic differentiation of native and translocated populations. *Mol. Ecol.* 7:751-760.

Rowe G, Beebee TJC, Burke T. 2000. A further four polymorphic microsatellite loci in the natterjack toad *Bufo calamita*. *Cons. Genet.* 1:371-372.

Rowe G, Harris DJ, Beebee TJC. 2006. Lusitania revisited: a phylogeographic analysis of the natterjack toad *Bufo calamita* across its entire biogeographical range. *Mol. Phyl. Evol.* 39:335-346.

Semagn K, Babu R, Hearne S, Olsen M. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol. Breeding.* 33:1-14.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics.* 139:457-462.

Vanbrabant J, Goddeeris B, Martens K, Schön I. 2009 Development of primers for amplification and further research on the D-loop control region for the common spadefoot toad, *Pelobates fuscus* (Amphibia). *Cons. Genet. Res.* 1: 495-499.

Wirtz S, Böhm C, Fritz J, Hankeln T, Hochkirch A. 2016. Isolation of microsatellite loci by next-generation sequencing of the critically endangered northern bald ibis, *Geronticus eremita*. *J. Hered.* 107: 363-366.

Supplementary files

Table S1: Characteristics of 18 mitochondrial SNP markers defined by Sanger sequencing and used for KASPar genotyping in the natterjack toad, *Bufo (Epidalea) calamita*. a: SNP locus name; b: GenBank accession numbers for the genomic mitochondrial DNA fragment obtained by Sanger sequencing; c: Allele X and Allele Y forward primers, respectively labelled with FAM or HEX fluorescent dyes; d: common reverse primer. Within genomic regions surrounding the SNPs, the latter are indicated in bold within brackets.

SNP marker[a]	GenBank accession numbers[b]	Genomic sequence surrounding the SNP	SNP-specific forward primers[c]	Common reverse primers[d]	Allele X (FAM)	Allele Y (HEX)
Bc_16S_1_G/A	KX237595 KX237596	GATCAATTAAACCCTTTTAATAAACCATTTTTATATTCTAGTAA**[G/A]**GGAGATTGAAAATTTTATTGGAGYCATAACGACAGTACCGCA	Allele X(FAM): CCTTTTAATAAACCATTTTTATATTCTAGTAAG Allele Y(HEX): CCTTTTAATAAACCATTTTTATATTCTAGTAAA	CGGTACTGTCGTTATGRCTCCAATAAA	G	A
Bc_16S_3_C/T	KX237597 KX237598	CCCCTTGTACCTTTTGCATCATGGTTTAACTAGT**[C/T]**TAATCAAGCRAAACGTATTTAAGTTTGACTTCCCGAAA	Allele X(FAM): ACCTTTTGCATCATGGTTTAACTAGTC Allele Y(HEX): GTACCTTTTGCATCATGGTTTAACTAGTT	GGGAAGTCAAACTTAAATACGTTTYGCTT	C	T
Bc_16S_6_G/A	KX237599 KX237600	CTTCTCTTAAATGCGAGTGTACATCAGAAAGGACAAACC**[G/A]**CTGATATTTAACACCTCTGAGCTTAAAGTAATAACTTCTCA	Allele X(FAM): TAAGCTCAGAGGTGTTAAATATCAGC Allele Y(HEX): CTTTAAGCTCAGAGGTGTTAAATATCAGT	AAATGCGAGTGTACATCAGAAAGGACAAA	G	A
Bc_16S_7_T/C	KX237601 KX237602	AATAACTTCTCAAGAAAATTTTACTTGCACCAATGT**[T/C]**AACCTAACACAAGAGCATTCTAAGAAAGATTAAAA	Allele X(FAM): CTCAAGAAAATTTTACTTGCACCAATGTT Allele Y(HEX): CTCAAGAAAATTTTACTTGCACCAATGTC	ATCTTTCTTAGAATGCTCTTGTGTTAGGTT	T	C
Bc_16S_9_A/G	KX237603 KX237604	TGTCTCCTTTTTCTAATCAGTGAAACTAATCTCCCC**[A/G]**TGAAGAAGCGGGGATAAAAATATAAGACGAGAAGACC	Allele X(FAM): TCTAATCAGTGAAACTAATCTCCCCA Allele Y(HEX): CTAATCAGTGAAACTAATCTCCCCG	CTCGTCTTATATTTTTATCCCCGCTTCTT	A	G
Bc_16S_11_T/-	KX237605 KX237606	TAAACAAAATAGCACTTGCCACATAAC(AT/--)ACGATTTCCTGAA**[-/T]**TTTTTACTTCACCAAGCAATATGA	Allele X(FAM): TATAGTCATATTGCTTGGTGAAGTAAAAAT Allele Y(HEX): CTTATAGTCATATTGCTTGGTGAAGTAAAAAA	GAGCTTTAAACAAAATAGCACTTGCCACAT	-	T
Bc_ControlR_2_T/C	KX237607 KX237608	CCTTATATTTCAACAATAATAAATCTTCACATGCYAA**[T/C]**GAAGATTTACAAYACGWATGGTGTAGTACATATTAT	Allele X(FAM): ATTTCAACAATAATAAATCTTCACATGCYAAT Allele Y(HEX): TCAACAATAATAAATCTTCACATGCYAAC	TACATGCATAATATGTACTACACCATWCGT	T	C
Bc_ControlR_5_T/C	KX237609 KX237610	ACAAYACGWATGGTGTAGTACATATTATGCATGTATA**[T/C]**AGACATACTATGTATAATAGAGCATTCATCTTATTTCCCCATG	Allele X(FAM): GATGAATGCTCTATTATACATAGTATGTCTA Allele Y(HEX): ATGAATGCTCTATTATACATAGTATGTCTG	CGWATGGTGTAGTACATATTATGCATGTAT	T	C
Bc_ControlR_6_T/C	KX237611 KX237612	AGCATTCATCTTATTTCCCCATGCATATCATTACCCATAAT**[T/C]**ATGAAAAATTCACATAATATrCCAA(--/AA)AAATTTTAA	Allele X(FAM): CCCCATGCATATCATTACCCATAATT Allele Y(HEX): CCCCATGCATATCATTACCCATAATC	TAATAATGTTTATTAAACAGGGGTGGGKAA	T	C
Bc_ControlR_7_T/C	KX237613 KX237614	CTTATATTTCAACAATAATAAATCTTCACATGC**[C/T]**AATGAAGATTTACAAYACGTATGGTGTAGTACATATTAT	Allele X(FAM): TATATTTCAACAATAATAAATCTTCACATGCC Allele Y(HEX): ATATTTCAACAATAATAAATCTTCACATGCT	ACTACACCATACGTRTTGTAAATCTTCATT	C	T
Bc_ControlR_9_T/C	KX237615 KX237616	TATGTATAATAGAGCATTCATCTTATTTCCCCATGCATATCA**[T/C]**ATyAACAATATATYTTTAAyTTYACATATTAAGACTCACATAA	Allele X(FAM): CATTCATCTTATTTCCCCATGCATATCAT Allele Y(HEX): CATCTTATTTCCCCATGCATATCAC	CGTGAAACGTTTRTTGATGTGATATAAT	T	C
Bc_ControlR_13_C/T	KX237617 KX237618	ATAAAATTATTATCAAATTTAArATTATATCACATCAA**[C/T]**AAACGTTTCACGAAGTTGGRATGATCTTAATGAATAAAGACT	Allele X(FAM): ATCATYCCAACTTCGTGAAACGTTTG Allele Y(HEX): GATCATYCCAACTTCGTGAAACGTTTA	TCTTATTTCCCCATGCATATCAYAT	C	T
Bc_ControlR_14_G/A	KX237619 KX237620	ATTATATCACATCAAyAAACGTTTCACGAAGTTGG**[G/A]**ATGATCTTAATGAATAAAGACTRTTACTCGT	Allele X(FAM): ATCAAYAAACGTTTCACGAAGTTGGG Allele Y(HEX): CATCAAYAAACGTTTCACGAAGTTGGA	CTCGATTGACGAGTGAAYACGAGTAA	G	A
Bc_ControlR_18_G/A	KX237621 KX237622	TCTCCTTGAATTRAGAGCACAAATACCCACATCAAA(-/A)TT**[A/G]**AAACTCAAGTATTTTCCTTATMTAATGAAT	Allele X(FAM): CATTCATTAKATAAGGAAAATACTTGAGTTTT Allele Y(HEX): CATTCATTAKATAAGGAAAATACTTGAGTTTC	TGAATTRAGAGCACAAATACCCACATCAAA	A	G
Bc_ControlR_19_C/A	KX237623 KX237624	AAACTCAAGTATTTTCCTTAT**[A/C]**TAATGAATGYTTAATAGACATACAACTTATATATTGCACAATCTCARTCGTT	Allele X(FAM): ATATAAGTTGTATGTCTATTAARCATTCATTAT Allele Y(HEX): ATATAAGTTGTATGTCTATTAARCATTCATTAG	TGAATTRAGAGCACAAATACCCACATCAAA	A	C
Bc_ControlR_20_C/T	KX237625 KX237626	CTCAAGTATTTTCCTTATMTAATGAATG**[C/T]**TTAATAGACATACAACTTATATATTGCACAATCTCARTCGTTCAGA	Allele X(FAM): ATTGTGCAATATATAAGTTGTATGTCTATTAAG Allele Y(HEX): GTGCAATATATAAGTTGTATGTCTATTAAA	AAACTCAAGTATTTTCCTTATMTAATGAAT	C	T
Bc_ControlR_24_C/T	KX237627 KX237628	GATTAATTAACTAATTATTAATCAACAA**[C/T]**AATGArATAACAAACTTAyTCCAATTAAyTCAAGAATATTAAATTT	Allele X(FAM): CCWTGATTAATTAACTAATTATTAATCAACAAC Allele Y(HEX): CCWTGATTAATTAACTAATTATTAATCAACAAT	CGTATTGACTRGGATTATTAAGGCTCAATT	C	T
Bc_ControlR_26_T/A	KX237629 KX237630	CCyAGTCAATACGAATAATACTTATAATTTAAC**[A/T]**TACATAATAACCATCCATGGTCTATTTAATAAA	Allele X(FAM): TATTAAATAGACCATGGATGGTTATTATGTAT Allele Y(HEX): ATTAAATAGACCATGGATGGTTATTATGTAA	GAGCCTTAATAATCCYAGTCAATACGAATA	T	A

Table S2: Detailed allelic state combinations corresponding to the ten haplotypes (Hap) defined from 18 mitochondrial SNP markers in five populations of the natterjack toad, *Bufo* (*Epidalea) calamita*.

mtDNA haplotype name	Allelic state SNP combinations
Hap 1	ATATGGATTTTGCTAACT
Hap 2	ATGTGGATTTTGCTAACT
Hap 3	ATATGGGTTTTGCTTACT
Hap 4	ATATGGATTTTGTTAACT
Hap 5	ATACGGATTTTGTTAACT
Hap 6	ACATGGATCTTACTTACT
Hap 7	ATATGGATTTTGCTTACT
Hap 8	ACATGGATTTTGCTTACT
Hap 9	ATATGGATTTTGTTTACT
Hap 10	ATATGGATTTCGCTTACT

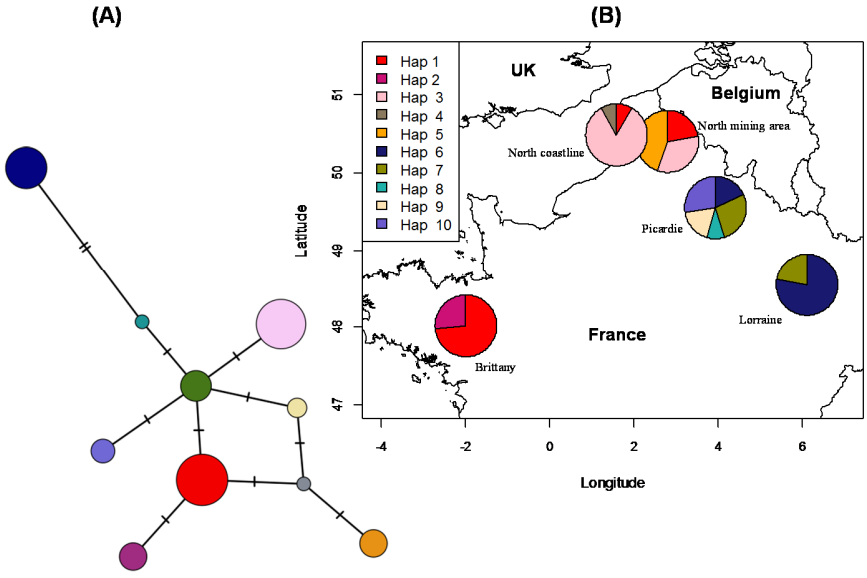


Figure S1 (Supplementary Material) : (A) Median joining network showing the genealogical relationships among 10 mtDNA haplotypes observed in five populations of the natterjack toad *Bufo* (*Epidalea) calamita*. Network design was carried out using the software POPART (http://popart.otago.ac.nz). (B) Geographical distribution of the 10 mtDNA haplotypes over five populations of *B. calamita* sampled in Western Europe (France). MtDNA haplotypes (Hap) are based on the combination of allelic states of 18 SNPs into a single haplotype.

CHAPITRE II

Histoire évolutive du Crapaud calamite dans le nord de la France

« Vers la droite, le terri barrait la vue, colossal comme une barricade de géants, déjà couvert d'herbe dans sa partie ancienne, consumé à l'autre bout par un feu intérieur qui brûlait depuis un an, avec une fumée épaisse, en laissant à la surface, au milieu du gris blafard des schistes et des grès, de longues traînées de rouille sanglante. Puis, les champs se déroulaient, des champs sans fin de blé et de betteraves, nus à cette époque de l'année, des marais aux végétations dures, coupés de quelques saules rabougris, des prairies lointaines, que séparaient des files maigres de peupliers. Très loin, de petites taches blanches indiquaient des villes, Marchiennes au nord, Montsou au midi; tandis que la forêt de Vandame, à l'est, bordait l'horizon de la ligne violâtre de ses arbres dépouillés. »

Emile Zola, *Germinal* (1885)

© CPIE Chaine des terrils

Photographie d'un crapaud calamite aux abords d'un terril

Initialement localisé au sein d'habitats littoraux et prairiaux, le Crapaud Calamite a progressivement colonisé des habitats industriels créés puis laissés à l'abandon, tels que les terrils mis en place au cours de ces trois derniers siècles dans le bassin houillier du nord de la France (Godin 2002; Lemoine 2012). Dans quelle mesure ces nouveaux sites peuvent-il être considérés comme des refuges de biodiversité durable ? Le processus de colonisation peut s'accompagner de perte drastique de diversité génétique au sein de populations nouvellement fondées (Micheletti & Storfer 2015). Parallèlement, la forte fragmentation du milieu peut dégrader la connectivité entre les populations et entrainer également une érosion de la diversité génétique au sein des populations. Pour mieux appréhender le statut et de la dynamique de ces populations fondatrices, nous avons, dans cette première partie, entrepris d'étudier leur histoire évolutive en intégrant des données sur la structure génétique de ces populations ainsi que sur celles vivants dans les habitats semi-naturels, localisés essentiellement le long du littoral et en bordure du bassin minier.

Les objectifs de ce chapitre de thèse étaient de déterminer si les populations localisées dans le bassin minier s'étaient formées à partir d'individus fondateurs en provenance d'habitats originels littoraux ou prairiaux, ou au contraire si ces populations avaient été fondées suite à des introductions accidentelles d'origine anthropique, dont les sources restaient à déterminer. En termes de priorisation de gestion conservatoire des populations de Crapaud calamite dans le Nord - Pas-de-Calais, il était aussi primordial d'avoir des informations sur les niveaux de diversité génétique présents au sein de ces populations et sur les flux de gènes opérant entre populations.

Pour répondre à ces questions, près d'un millier d'individus ont été échantillonnés dans plus de 50 sites du nord de la France, et 250 individus ont été échantillonnés dans 15 sites ailleurs en France, en Suisse ou en Suède. Les individus ont été capturés à la main de nuit durant la période de reproduction et les prélèvements noninvasifs d'ADN ont été effectués à l'aide d'écouvillons buccaux permettant de prélever des cellules épithéliales (Figure II 1).

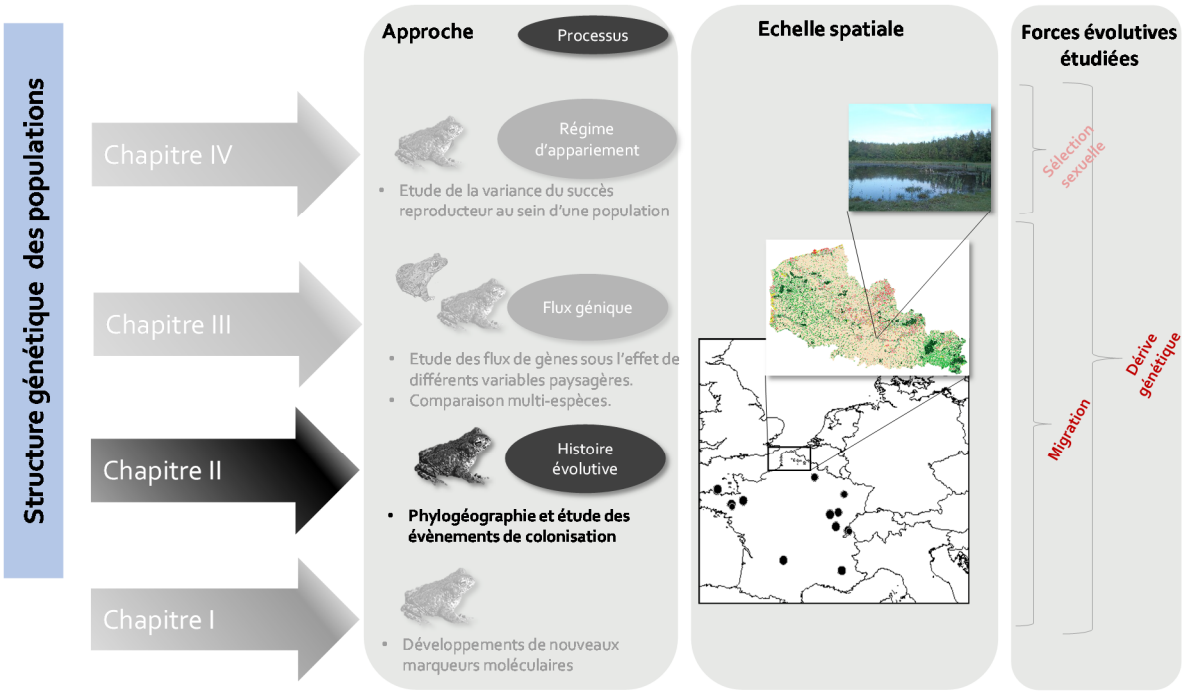
Figure II 1 : Photographie d'un prélèvement de cellules épithéliales réalisé à l'aide d'un écouvillon sur un Crapaud calamite mâle. © JF Arnaud

Tous les individus ont été génotypés sur 37 marqueurs microsatellites. De plus, 15 marqueurs de polymorphisme de type SNP mitochondriaux ont été caractérisés sur un sous-échantillon de 855 individus.

Diverses approches statistiques nous ont permis de décrire la structure génétique nucléaire et mitochondriale de l'ensemble des 68 populations échantillonnées puis, à plus fine échelle sur les 53 populations localisées dans le nord de la France. Une approche d'approximation bayésienne nous a permis d'évaluer le scénario de colonisation le plus probable. Sur la région du nord de la France, les niveaux d'apparentement génétique entre individus, et les niveaux de consanguinité, ont été comparés entre habitats semi-naturels localisés le long du littoral et habitats post-industriel localisés dans le bassin minier. L'existence d'éventuels goulots d'étranglements induits par des réductions rapides des effectifs au sein des populations a également été étudiée. Prenant en compte ces paramètres intra-population, nous avons également analysé la structure génétique spatiale dans les différentes zones occupées par les populations de Crapauds calamites afin d'évaluer l'impact potentiel des différentes matrices paysagères sur les flux de gènes.

Il en ressort que les populations de Crapaud calamite observées dans le bassin minier résultent à la fois d'une remontée des individus depuis le dernier épisode glaciaire ayant eu lieu il y a environ 15 000 ans, mais aussi des activités humaines récentes qui ont probablement joué un rôle important dans la dispersion et l'introduction accidentelle d'individus dans plusieurs sites du bassin minier. Aucune trace de consanguinité n'est observée au sein de ces zones industrielles, une très forte diversité génétique y est mesurée, et plusieurs lignées maternelles d'origines géographiques différentes ont été découvertes. Ceci illustre ainsi un « hot-spot » de diversité génétique au sein de ces habitats pourtant fortement impactés par l'homme. Ces résultats encourageants ne permettent toutefois pas d'affirmer qu'en tant que telles les populations pourront se maintenir dans le temps. En effet, ces populations diverses génétiquement ont une histoire évolutive très récente, et nous montrons par ailleurs un isolement génétique très fort, autrement dît des flux génique restreints, ceci pour des populations qui sont pourtant spatialement très proches (<5kms). Ces ruptures de flux géniques observées sur des échelles spatiales très réduites pourraient être imputables aux infrastructures humaines dans cette région qui est fortement urbanisée. Ainsi, même si ces populations récemment implantées au sein du bassin minier ne semblent pas sujettes à une érosion génétique ou de faibles tailles efficaces, l'enjeu sur le long terme est de faciliter leurs connexions au travers de la préservation d'habitats propices au cycle de vie et à la dispersion d'espèces d'amphibiens.

Ce chapitre a fait l'objet d'une publication (Faucher *et al.* 2017). Le travail portant sur les analyses de paysage et visant à déterminer les éléments impactant les flux de gènes fait l'objet du chapitre III qui suivra.



Références

Faucher L, Hénocq L, Vanappelghem C *et al.* (2017) When new human-modified habitats favour the expansion of an amphibian pioneer species: evolutionary history of the natterjack toad (*Bufo calamita*) in a coal basin. *Molecular Ecology*, **26**, 4434–4451.

Godin J (2002) Degré de rareté, évolution de la distribution et particularités de l'herpétofaune de la région Nord - Pas de Calais. *Bulletin de la Société Herpétologique de France*, **104**, 16–35.

Lemoine G (2012) Interêt, gestion, protection et valorisation des terrils miniers dans le Nord–Pas-de-Calais. *Annales de la Société Géologique du Nord*, **19**, 1–8.

Micheletti SJ, Storfer A (2015) A test of the central-marginal hypothesis using population genetics and ecological niche modelling in an endemic salamander (*Ambystoma barbouri*). *Molecular Ecology*, **24**, 967–979.

WHEN NEW HUMAN-MODIFIED HABITATS FAVOR THE EXPANSION OF AN AMPHIBIAN PIONEER SPECIES: EVOLUTIONARY HISTORY OF THE NATTERJACK TOAD (*BUFO CALAMITA*) IN A COAL BASIN

Leslie Faucher[1], Laura Hénocq[1], Cédric Vanappelghem[1,2], Stéphanie Rondel[3], Robin Quevillart[4], Sophie Gallina[1], Cécile Godé[1], Julie Jaquiéry[5] & Jean-François Arnaud[1*]

[1] Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France
[2] Conservatoire d'espaces naturels du Nord et du Pas de Calais, Lillers, France
[3] Centre Permanent d'Initiatives pour l'Environnement - Chaîne des Terrils, Loos-en-Gohelle, France
[4] Groupe ornithologique et naturaliste du Nord – Pas-de-Calais, Lille, France
[5] Université de Rennes 1, UMR CNRS 6553 – ECOBIO, 35042 Rennes Cedex, France

* Corresponding author
E-mail: jean-francois.arnaud@univ-lille1.fr

Keywords: man-made habitats; amphibians; nuclear and mitochondrial diversity; spoil heaps; colonization process; genetic structure

Received: 1 February 2017 | Revised: 17 May 2017 | Accepted: 5 June 2017
DOI: 10.1111/mec.14229

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

When new human-modified habitats favour the expansion of an amphibian pioneer species: Evolutionary history of the natterjack toad (*Bufo calamita*) in a coal basin

Leslie Faucher[1] | Laura Hénocq[1] | Cédric Vanappelghem[1,2] | Stéphanie Rondel[3] | Robin Quevillart[4] | Sophie Gallina[1] | Cécile Godé[1] | Julie Jaquiéry[5] | Jean-François Arnaud[1]

[1]Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, Lille, France
[2]Conservatoire d'espaces naturels du Nord et du Pas de Calais, Lillers, France
[3]Centre Permanent d'Initiatives pour l'Environnement - Chaine des Terrils, Loos-en-Gohelle, France
[4]Groupe ornithologique et naturaliste du Nord – Pas-de-Calais, Lille, France
[5]UMR CNRS 6553 – ECOBIO, Université de Rennes 1, Rennes Cedex, France

Correspondence
Jean-François Arnaud, Univ. Lille, CNRS, UMR 8198 – Evo- Eco- Paleo, Lille, France
Email: jean-francois.arnaud@univ-lille1.fr

Funding information
Région Nord-Pas De Calais, Grant/Award Number: AMPHIDIV project; European Regional Development Fund, Grant/Award Number: CLIMIBIO

Abstract

Human activities affect microevolutionary dynamics by inducing environmental changes. In particular, land cover conversion and loss of native habitats decrease genetic diversity and jeopardize the adaptive ability of populations. Nonetheless, new anthropogenic habitats can also promote the successful establishment of emblematic pioneer species. We investigated this issue by examining the population genetic features and evolutionary history of the natterjack toad (*Bufo [Epidalea] calamita*) in northern France, where populations can be found in native coastal habitats and coalfield habitats shaped by European industrial history, along with an additional set of European populations located outside this focal area. We predicted contrasting patterns of genetic structure, with newly settled coalfield populations departing from migration-drift equilibrium. As expected, coalfield populations showed a mosaic of genetically divergent populations with short-range patterns of gene flow, and native coastal populations indicated an equilibrium state with an isolation-by-distance pattern suggestive of post-glacial range expansion. However, coalfield populations exhibited (i) high levels of genetic diversity, (ii) no evidence of local inbreeding or reduced effective population size, and (iii) multiple maternal mitochondrial lineages, a genetic footprint depicting independent colonization events. Furthermore, Approximate Bayesian Computations suggested several evolutionary trajectories from ancient isolation in glacial refugia during the Pleistocene, with biogeographical signatures of recent expansion probably confounded by human-mediated mixing of different lineages. From an evolutionary and conservation perspective, this study highlights the ecological value of industrial areas, provided that ongoing regional gene flow is ensured within the existing lineage boundaries.

Introduction

Anthropogenic land-use changes are by far the most important contemporary cause of native habitat loss and landscape fragmentation (Fischer & Lindenmayer 2007). The loss of habitats and isolation of populations can decrease levels of gene flow among populations and reduce the effective population size (N_e), two key components of population genetic variability (Baguette *et al.* 2013; Ellegren & Galtier 2016). This is a major threat to a wide array of taxa, especially amphibians that are the vertebrate lineage with the highest number of species facing extinction (Beebee & Griffiths 2005; Wake & Vredenburg 2008; Allentoft & O'Brien 2010; Rivera-Ortíz *et al.* 2015). The vulnerability of amphibians facing habitat loss is attributed to numerous factors including relatively limited dispersal capabilities, high breeding site fidelity, a complex life cycle requiring different types of habitats (aquatic habitats for breeding period and/or larvae stages and terrestrial habitats otherwise) often combined with narrow habitat tolerance, and generally small N_e with high demographic fluctuations (Beebee & Griffiths 2005; McCartney-Melstad & Shaffer 2015; Lourenço *et al.* 2017).

Anthropogenic habitats can differ from native habitats with respect to their colonization history and may have different evolutionary effects (Alberti 2015). Urban environments can host trapped and isolated populations. In this context, recent studies investigated the strength of genetic drift induced by reduced functional connectivity across an urban matrix (*e.g.* Munshi-South *et al.* 2013; Beninde *et al.* 2016; Furman *et al.* 2016; Lourenço *et al.* 2017). However, it appears that new anthropogenic habitats can also promote the establishment of pioneer species in highly disturbed areas (Amici *et al.* 2015; Flavenot *et al.* 2015). This is particularly true for post-industrial environments such as former mining areas that may act as biodiversity refugia (Brock *et al.* 2007). Nonetheless, studies dealing with population genetic features in such *a posteriori*-colonized environments are very scarce and, to the best of our knowledge, only focused on a few plant species (Brock *et al.* 2007; Esfeld *et al.* 2008; Prinz *et al.* 2009; Pauwels *et al.* 2012). Our work aims at bridging this gap by assessing the current genetic patterns and evolutionary history of an amphibian species in natural and reclaimed coal mine environments.

In northern France, three centuries of coal extraction has led to high habitat fragmentation and the formation of spoil heaps, *i.e.* artificial hills constituted by mineral materials extracted with the coal. Spoil heaps (some of which cover 90 ha and exceed 140 m in height) have thin and nutrient-poor soil. Since the cessation of mining activities in the late 1980s, spoil heaps have

constituted new continental xeric primary ecosystems suitable for pioneer species, as well as a specialized flora adapted to anthropogenic metal stress (Godin 2002; Lemoine 2012; Pauwels *et al.* 2012). Presented as a negative symbol for a long time, the spoil heaps of northern France are the latest addition to UNESCO's world heritage list and landowners, agencies and volunteer associations all work to protect this cultural and natural heritage and to improve habitats for biodiversity and wildlife monitoring (Lemoine 2012).

In this study, we investigated the population genetic structure of the natterjack toad (*Bufo* [*Epidalea*] *calamita,* Laurenti, 1768) that established in this new human-modified habitat, and examined the related historical process of colonization. The natterjack toad is a protected pioneer anuran species characterized by a wide geographical distribution ranging from central and western Europe and extending northward to Sweden (Rowe *et al.* 2006). This species generally establishes in nutrient-poor environments with open and unshaded light sandy soils, such as coastal dunes, lowland heaths, semi-deserts, and meadows or marshes (Beebee & Denton 1996; Denton *et al.* 1997; Wilkinson & Griffiths 2013). Species occurrence declines in the northern parts of its range where populations are more fragmented and isolated (Allentoft *et al.* 2009). Based on historical records dating back to 1878 and current mapping within the coalfield areas of northern France, the natterjack toad was historically present in inland semi-natural habitats in the direct vicinity of mining areas. Subsequently, this species successfully settled on the slopes and in the surroundings of abandoned spoil heaps that constitute suitable breeding and resting habitats with temporary ponds (Godin 2002). Today, most historical semi-natural populations previously located in the neighborhood of the spoil heaps became extinct or undetectable (see Fig. S1, Godin 2002; Lemoine 2012). Outside the former coalfields, there are numerous populations established along the coastline area in coastal sand dunes and saltmarshes known to be their “native habitats”.

We conducted an extensive population genetic structure analysis based on both nuclear and mitochondrial data. We performed an exhaustive sampling of natterjack toad populations located in the anthropogenic habitats of the mining area and the natural coastal areas of northern France. Our main objectives were (i) to determine whether coastal natural and inland coalfield populations differ in terms of patterns of genetic structure and (ii) to assess the colonization and evolutionary history of inland coalfield populations. Owing to founder effects during the establishment of coalfield populations and to high habitat fragmentation impeding gene flow, our expectations were twofold. First, we expected reduced migration and increased drift for mining area populations

compared to native coastal populations, with higher levels of genetic differentiation and lower levels of genetic diversity. Next, we hypothesized a common ancestry between coastal natural populations and mining area populations, the latter being thought to be founded by natural colonization events and/or accidental translocations of individuals during former mining activities from previous semi-natural inland habitats hosting the natterjack toad (Godin 2002). Lastly, we discussed the evolutionary implications and validity of our findings for the long-term management of pioneer species in anthropogenic areas.

Materials and methods

Study area

Northern France is a broad open plain with, in the west, a shoreline composed of cliffs and wide beaches sheltering wildlife protection areas and, in the east, inland areas mainly composed of anthropogenic habitats (Fig. 1A and Fig. S1). The coalfield area of northern France covers 1200 km² from east to west (Fig. 1A). This industrial area illustrates the significant human impact with landscapes shaped over three centuries by coal extraction, which was virtually a mono-industry in this region from the 1700s to the 1900s. Coal activities led to the formation of more than 300 spoil heaps connected by railways along with extensive urban development. Today, the mining area consists of a very fragmented landscape dominated by urban habitats and conventional farming and crossed by several chains of spoil heaps constituting new continental open habitats. Some spoil heaps are still exploited for schist extraction, others have been abandoned or are now managed for leisure or wildlife protection. Natural early successional habitats such as coastal sand dunes are likely to be the primary "native" sites for *B. calamita*. This species has now successfully colonized numerous spoil heaps, potentially from former semi-natural habitats such as marshes or grassland in the neighborhood of coalfield areas (Fig. S1). Our sampling thus focused on the newly colonized mining area and the neighboring coastal area of northern France likely to host native populations. Additional populations located in different areas of the biogeographical range were also collected to calibrate the genetic divergence of coastline and mining area populations of northern France and to shed light on the settlement history of the natterjack toad in this peculiar area (Fig. 1A).

Sampling procedure

Overall, a total of 68 populations (1209 adult individuals) were sampled. We collected DNA from 959 *B. calamita* individuals in 53 different populations in the focal study area in northern France, with a mean (±SD) of 18 ±12 individuals sampled per population: 32 populations (MA-1 to MA-32) were located within the mining area and 21 populations (C-1 to C-21) were located along the coastal area (Fig. 1A, Table 1). To date, no *B. calamita* populations have been recorded between the coastal and coalfield areas, and we consider population sampling to be exhaustive based on reported locations from 20 years ago (Godin 2002, Fig. S1). To assess broad genetic structure patterns and test the origin of northern France populations, 250 additional individuals were collected from 15 populations in other parts of the *B. calamita* geographical distribution with a mean (±SD) of 17 ± 6 individuals per population; 12 of these populations were collected in other regions of France, 2 in Switzerland and 1 in Sweden (see Fig. 1A and Table 1). Individuals were caught by night during the breeding period (from March to June). Each locality was sampled once in 2013 or 2014. A non-invasive sampling method was used (Broquet *et al.* 2006). Plain sterile 15SC Copan (Brescia Italia) swabs were used to take samples of buccal cells for each captured individual and were stored at -20°C or dried prior to DNA extraction in the laboratory. Individuals were immediately released after buccal swabbing.

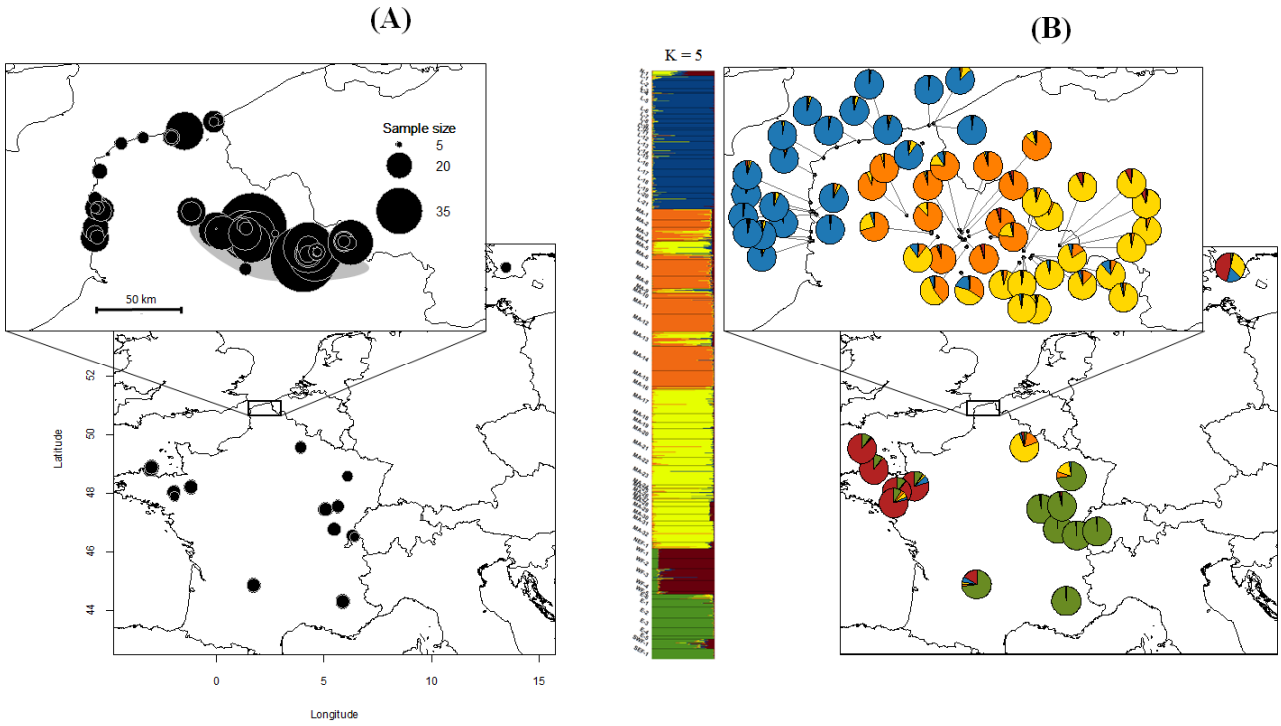


Figure 1: (A) Location of the 68 populations of *Bufo [Epidalea] calamita* sampled in Western Europe. Inset map focuses on fine-scale geographical distribution of 53 populations located in the focal study area in northern France, with circle sizes being proportional to the number of sampled individuals. Shaded area indicates the location of the coalfield area in northern France. (B) Results of Bayesian clustering assignment probability of individuals (barplot on the left) and of populations (maps on the right) giving their membership in *K*= 5 genetic clusters. Each individual is represented by a thin line partitioned into *K* colored segments displaying the individual estimated membership fractions in *K* clusters. Each population is represented by a diagram partitioned into *K* colors displaying the mean population estimated membership fractions in *K* clusters.

Table 1: Estimates of nuclear and mitochondrial genetic variation in 68 populations of *Bufo [Epidalea] calamita*. Sample size (N), total number of alleles (A_T), the mean number of alleles (A_n), the observed heterozygosity (H_O), expected heterozygosity (H_E), the nuclear allelic richness (A_r) and the mean intra-population fixation index (F_{IS}) according to Weir & Cockerham (1984). Significance of F_{IS} was tested using a *U*-test with 10,000 iterations for each population over all loci and corrected using a Bonferroni adjustment technique (Rice 1989). Also reported are the number of genotyped individuals for mitochondrial SNP polymorphism (N_mito), the number of different mitochondrial haplotypes (M_T) and the haplotype richness (M_r), and the one-sample effective population size (N_e) estimates using linkage disequilibrium method with jackknife 95% CIs. ABC Analyses: specific populations used in ABC analyses. –: Not applicable (insufficient sampling size). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

ID	Area	Latitude (WGS84)	Longitude (WGS84)	N	A_T	A_n	H_O	H_E	A_r	F_{IS}	N_mito	M_T	M_r	N_e	95% CIs	ABC analyses
C-01	North Coastline	2.5398	51.0821	9	118	3.371	0.471	0.523	3.19	0.104 *	9	3	2.96	–	–	X
C-02	North Coastline	2.504	51.0728	17	129	3.686	0.525	0.504	3.12	-0.043	15	2	2	53	30.7 - 149	
C-03	North Coastline	2.5044	51.0686	8	107	3.057	0.468	0.467	2.97	-0.001	8	2	2	–	–	
C-04	North Coastline	2.4501	50.976	2	63	–	–	–	3.29	–	2	2	–	–	–	
C-05	North Coastline	2.2778	51.0199	29	133	3.8	0.473	0.521	2.9	0.094 ***	10	1	1	46.9	33.3 - 73.8	
C-06	North Coastline	2.1699	50.9883	13	134	3.829	0.51	0.504	3.31	-0.013	13	2	2	–	–	X
C-07	North Coastline	2.1744	50.9877	10	120	3.429	0.534	0.501	3.16	-0.071 *	10	2	2	-	–	
C-08	North Coastline	1.954	50.983	9	122	3.486	0.509	0.507	3.07	-0.005	9	2	2	–	–	
C-09	North Coastline	1.7849	50.9519	10	117	3.343	0.489	0.497	3.14	0.017	10	2	2	–	–	X
C-10	North Coastline	1.6803	50.8945	4	83	–	–	–	3.24	–	4	1	-	–	–	
C-11	North Coastline	1.6292	50.7968	12	119	3.4	0.486	0.514	3.12	0.046	11	1	1	–	–	X
C-12	North Coastline	1.5788	50.649	11	111	3.171	0.423	0.443	2.91	0.047	10	3	2.92	–	–	
C-13	North Coastline	1.5876	50.5991	18	129	3.686	0.465	0.457	3.03	-0.02	13	2	2	12.5	9.8 - 16.2	
C-14	North Coastline	1.5987	50.5984	10	125	3.571	0.483	0.486	3.27	0.008	7	1	1	-	–	X
C-15	North Coastline	1.5785	50.5968	8	121	3.457	0.443	0.471	3.33	0.065	7	1	1	-	–	X
C-16	North Coastline	1.6237	50.5829	21	109	3.114	0.45	0.412	2.64	-0.093 **	14	1	1	8.6	6.5 - 11.5	
C-17	North Coastline	1.6244	50.5823	16	102	2.914	0.409	0.387	2.5	-0.057	12	1	1	28	15.9 - 73.4	
C-18	North Coastline	1.5832	50.4869	20	145	4.143	0.499	0.485	3.3	-0.029	12	3	2.67	198.7	73.8 - ∞	X
C-19	North Coastline	1.5758	50.456	13	138	3.943	0.479	0.486	3.35	0.015	12	3	2.84	–	–	X
C-20	North Coastline	1.5889	50.4398	11	124	3.543	0.463	0.455	3.32	-0.021	11	3	2.88	–	–	
C-21	North Coastline	1.5758	50.435	22	157	4.486	0.516	0.498	3.15	-0.037	14	3	2.94	316.1	81.3 - ∞	X
MA-01	Northern Coalfield	2.3279	50.5772	22	104	2.971	0.439	0.455	2.71	0.035	15	2	1.93	13.1	10 - 17.5	
MA-02	Northern Coalfield	2.3291	50.5788	15	102	2.914	0.493	0.474	3.25	-0.039	10	4	3.84	6.9	4.9 - 9.4	
MA-03	Northern Coalfield	2.5198	50.4786	27	127	3.629	0.544	0.528	2.93	-0.031 *	25	4	3.66	20.2	16.2 - 25.8	
MA-04	Northern Coalfield	2.5218	50.4858	2	67	–	–	–	3.23	–	2	2	–	–	–	
MA-05	Northern Coalfield	2.5675	50.4591	26	107	3.057	0.47	0.455	3.11	-0.032	15	2	2	22.3	16.2 - 32.9	
MA-06	Northern Coalfield	2.738	50.5018	11	130	3.714	0.506	0.516	2.66	0.019	15	2	2	–	–	X
MA-07	Northern Coalfield	2.74	50.4615	34	140	4	0.518	0.534	4	0.034	11	4	3.87	36.7	28.9 - 48.4	X
MA-08	Northern Coalfield	2.7469	50.4576	25	140	4	0.533	0.533	4	-0.002	22	5	4.25	18.2	15.1-22.4	X
MA-09	Northern Coalfield	2.7534	50.2618	10	122	3.486	0.513	0.512	3.14	-0.002	16	3	3	–	–	
MA-10	Northern Coalfield	2.7574	50.4856	10	126	3.6	0.531	0.521	3.6	-0.021	10	5	4.84	–	–	X
MA-11	Northern Coalfield	2.7656	50.4461	33	129	3.686	0.486	0.485	3.69	0	14	2	1.76	18.3	15.1-22.3	
MA-12	Northern Coalfield	2.7797	50.4461	36	125	3.571	0.507	0.49	3.57	-0.035	14	2	1.76	57.9	40.7-92.9	
MA-13	Northern Coalfield	2.7967	50.4038	30	130	3.714	0.523	0.501	3.71	-0.044	15	4	3.87	50.9	36.5-79.3	
MA-14	Northern Coalfield	2.8022	50.492	52	134	3.829	0.504	0.5	3.39	-0.008	11	3	2.87	96.1	64.1-172.4	X
MA-15	Northern Coalfield	2.811	50.4625	32	111	3.171	0.464	0.444	3.11	-0.048	15	5	4.38	49.4	33.5-83.7	
MA-16	Northern Coalfield	2.9823	50.4583	6	95	–	–	–	2.94	–	6	2	–	–	–	
MA-17	Northern Coalfield	3.2041	50.3168	50	135	3.857	0.494	0.484	3.21	-0.026	15	4	3.72	29	24.5-34.7	
MA-18	Northern Coalfield	3.2332	50.3449	18	136	3.886	0.544	0.552	3.06	0.015	15	3	2.93	50.2	33.1-94.3	
MA-19	Northern Coalfield	3.2371	50.3447	13	128	3.657	0.497	0.533	3.38	0.069*	11	4	3.76	–	–	X
MA-20	Northern Coalfield	3.2381	50.3453	21	142	4.057	0.554	0.553	3.35	-0.002	14	3	2.75	16.6	13.2-21.4	
MA-21	Northern Coalfield	3.2409	50.3827	17	127	3.629	0.506	0.508	3.15	-0.031	15	3	2.72	75.8	39.1-474.4	X
MA-22	Northern Coalfield	3.2457	50.3751	38	142	4.057	0.545	0.529	3.27	0.003	13	5	4.53	51.2	38.6-72.3	
MA-23	Northern Coalfield	3.2899	50.3425	40	118	3.371	0.561	0.504	2.9	-0.116 ***	10	2	2	39.3	29.3-56	
MA-24	Northern Coalfield	3.3054	50.3581	7	109	3.114	0.567	0.513	3.31	-0.117 **	7	2	2	–	–	X
MA-25	Northern Coalfield	3.3059	50.3575	8	119	3.4	0.529	0.532	3.34	0.008	7	2	2	–	–	X
MA-26	Northern Coalfield	3.306	50.3584	13	130	3.714	0.549	0.554	3.34	0.008	13	4	3.79	–	–	X
MA-27	Northern Coalfield	3.3073	50.3585	7	117	3.343	0.527	0.539	3.24	0.026	7	3	3	–	–	
MA-28	Northern Coalfield	3.519	50.4135	9	105	3	0.517	0.501	2.85	-0.034	3	2	–	–	–	
MA-29	Northern Coalfield	3.5193	50.413	16	111	3.171	0.498	0.497	2.86	-0.001	10	3	3	54.7	27.8-333.5	
MA-30	Northern Coalfield	3.5208	50.4133	14	108	3.086	0.516	0.5	3.11	-0.034	11	3	2.99	–	–	
MA-31	Northern Coalfield	3.567	50.4044	10	115	3.286	0.503	0.501	3.01	-0.005	10	3	2.92	–	–	

MA-22	Northern Coalfield	3.2457	50.3751	38	142	4.057	0.545	0.529	3.27	0.003	13	5	4.53	51.2	38.6-72.3	
MA-23	Northern Coalfield	3.2899	50.3425	40	118	3.371	0.561	0.504	2.9	-0.116 ***	10	2	2	39.3	29.3-56	
MA-24	Northern Coalfield	3.3054	50.3581	7	109	3.114	0.567	0.513	3.31	-0.117 **	7	2	2	–	–	X
MA-25	Northern Coalfield	3.3059	50.3575	8	119	3.4	0.529	0.532	3.34	0.008	7	2	2	–	–	X
MA-26	Northern Coalfield	3.306	50.3584	13	130	3.714	0.549	0.554	3.34	0.008	13	4	3.79	–	–	X
MA-27	Northern Coalfield	3.3073	50.3585	7	117	3.343	0.527	0.539	3.24	0.026	7	3	3	–	–	
MA-28	Northern Coalfield	3.519	50.4135	9	105	3	0.517	0.501	2.85	-0.034	3	2	–	–	–	
MA-29	Northern Coalfield	3.5193	50.413	16	111	3.171	0.498	0.497	2.86	-0.001	10	3	3	54.7	27.8-333.5	
MA-30	Northern Coalfield	3.5208	50.4133	14	108	3.086	0.516	0.5	3.11	-0.034	11	3	2.99	–	–	
MA-31	Northern Coalfield	3.567	50.4044	10	115	3.286	0.503	0.501	3.01	-0.005	10	3	2.92	–	–	
MA-32	Northern Coalfield	3.5706	50.4076	34	133	3.8	0.519	0.517	2.74	-0.003	14	4	3.71	83.7	52.3-179.3	
NEF-01	Northern France	3.9334	49.5534	13	128	3.657	0.527	0.514	3.36	-0.027	11	4	3.98	-	–	
E-01	Eastern France	5.6668	47.5337	16	144	4.114	0.501	0.557	3	0.093 ***	14	2	2	68.9	34-764.6	X
E-02	Eastern France	5.0678	47.4317	23	146	4.171	0.471	0.572	3.46	0.153 ***	13	2	2	22.9	17.7-31.1	X
E-03	Eastern France	5.4833	46.75	20	94	2.686	0.345	0.373	3.57	0.078 **	12	1	1	62.7	26.2- ∞	X
E-04	Switzerland	6.3632	46.5355	17	109	3.114	0.436	0.49	1.92	0.111 ***	7	1	1	42.3	23.5-134	X
E-05	Switzerland	6.4196	46.5025	6	94	–	–	–	3.4	–	4	2	–	–	–	X
E-06	Lorraine, France	6.105	48.5509	10	108	3.086	0.501	0.49	3.45	-0.027	9	2	2	–	–	
SEF-01	Alpes, France	5.8811	44.2881	21	152	4.343	0.496	0.576	3.17	0.096 ***	10	1	1	1506.1	88.1- ∞	X
SWF-01	Lot, France	1.7075	44.8536	20	137	3.914	0.527	0.562	2.32	0.059 **	11	2	2	6.9	5.5-8.6	
WF-01	Western, France	-3.0028	48.8768	20	97	2.771	0.348	0.355	2.7	0.004	8	1	1	9	6.5-12.6	X
WF-02	Western, France	-3.0438	48.7871	21	75	2.143	0.31	0.327	2.65	0.031	7	1	1	40.5	16- ∞	X
WF-03	Western, France	-1.1667	48.2167	25	139	3.971	0.495	0.528	2.92	0.052 *	13	2	2	95.9	46.2-1805.	X
WF-04	Western, France	-1.992	48.016	22	112	3.2	0.449	0.444	2.43	-0.012	15	2	2	54.6	30.4-172.9	X
WF-05	Western, France	-1.95	47.8833	6	93	–	–	–	2.46	–	2	1	–	–	–	X
N-01	Sweden	13.4687	55.7106	10	52	1.486	0.144	0.157	3.32	0.089 *	10	1	1	–	–	

Laboratorial procedures and data quality control

Whole genomic DNA was extracted from swabs using Macherey-Nagel (Düren, Germany) NucleoSpin® 96 trace kits following the standard protocol outlined in the manufacturer's handbook. To assess nuclear DNA polymorphism, all samples were genotyped at 37 nuclear microsatellite loci (Table 1). Of these loci, 15 are described in Rowe et al. (1997, 2000) and Rogell et al. (2005) and 22 are recently developed loci, described in Faucher et al. (2016) (see Table S1). Amplification procedures (PCR), multiplexing and genotyping were carried out following the standard protocols described in Faucher et al. (2016). PCR products were analyzed using an ABI Prism 3730xl Analyzer (Applied Biosystems, Foster City, CA) and multilocus genotypes were manually scored using GENEMAPPER 3.7 software. Individuals that failed to amplify or that showed dubious genotypes underwent a second round of PCR. All the 1209 individuals were successfully genotyped at the 37 loci with an overall missing data rate of 1.95%.

Analyses of genetic diversity were applied to 62 populations (out of 68) characterized by a minimum sample size of seven genotyped individuals. The six remaining populations were only used for the Bayesian clustering analyses (see below). Linkage disequilibrium (LD) among pairs of loci and across populations was tested using a log-likelihood ratio statistic described in Raymond & Rousset (1995). Departures from Hardy–Weinberg (HW) equilibrium were tested for hypotheses of either excess or deficit in heterozygotes using a multisample score test, which defines a global

test across populations (Rousset & Raymond 1995). In both cases, the Markov Chain method implemented in the software GENEPOP v4.4.3 (Rousset 2008) provided unbiased estimates of probability using the following parameters: 10,000 dememorizations, 1000 batches, and 10,000 iterations per batch. Because multiple statistical tests were performed, we applied a sequential Bonferroni correction to P-values for a family-wise error rate of $\alpha = 0.05$ (Rice 1989).

Mitochondrial genetic diversity was also investigated using single nucleotide polymorphism (SNP) with a "kompetitive" allele-specific PCR (KASPar) genotyping assay method and primers defined by K Biosciences (Hoddesdon, UK), as described in Faucher *et al.* (2016). Fifteen SNPs located along a total of 2200 bp from the D-loop and the 16S gene and previously found to be polymorphic were surveyed using amplification procedures and detection of polymorphism as in Faucher *et al.* (2016). An average of 11 individuals per population were genotyped (see Table 1). Overall, the 15 SNPs were successfully amplified for 750 individuals on 855 chosen randomly (1.7% missing SNP data per locus). Individuals with SNP missing data were discarded from further analyses. Finally, each SNP combination was treated as a single haplotype because the mitochondrial genome is inherited as a single linkage unit.

Genetic diversity and contemporary Ne

Each population was described by calculating the total number of alleles observed in nuclear microsatellites (A_T), the average number of alleles per nuclear locus (A_n), the observed heterozygosity (H_O), the gene diversity (H_E), and the number of mitochondrial haplotypes (M_T). Similarly, allelic richness (A_r) and haplotypic richness (M_r) were estimated after accounting for population variation in sample sizes with a rarefaction method and a hierarchical sampling design implemented in the software package HPRARE (Kalinowski 2005). As a measure of individual inbreeding, the internal relatedness (*IR*) was estimated for each individual using the R package GENHET (Coulon 2009).

Long-term and contemporary N_e shapes current levels of genetic diversity and determines the strength of genetic drift (Ellegren & Galtier 2016). N_e was estimated using the LD method described in Waples (2006) and implemented in NeEstimator v2.0 (Do *et al.* 2014). This estimator is fairly robust to violations of the closed-population assumption when migration rates are moderate (Waples & England 2011). Alleles with frequency less than 5% were excluded and 95% confidence intervals were obtained by jackknifing across loci. N_e estimates are very sensitive

to reduced sample sizes, even when a large number of loci are employed (Wang 2016). Populations with sample sizes ≤ 15 were therefore excluded from analyses. Overlapping generations may also bias contemporary N_e estimates which should be interpreted with cautions (Waples *et al.* 2014). Nonetheless, LD estimations of N_e allow reliable relative comparisons across populations (*e.g.* Wang 2012; Lourenço *et al.* 2017).

Population genetic affiliation and levels of genetic differentiation

Population genetic affinities was first described with a neighbor-joining (NJ) tree based on pairwise genetic Cavalli-Sforza and Edwards (1967) distances (D_{CE}), using POPULATIONS 1.2.32 software (http://bioinformatics.org/~tryphon/populations/). For a visual assessment of the mitochondrial geographical distribution, haplotype distribution was mapped and genealogical relationships among haplotypes were estimated using a median-joining haplotype network drawn with POPART (http://popart.otago.ac.nz).

Next, we used a non-spatially explicit Bayesian clustering analysis, without any *a priori* population affiliation, as implemented in the program STRUCTURE v 2.3.4 (Pritchard *et al.* 2000; Hubisz *et al.* 2009). This analyze was applied on two data sets: (i) a full data set comprising 68 populations; and (ii) a data set comprising 53 populations sampled in the focal area of northern France (Fig. 1A). We used 500,000 burn-in Markov Chain Monte Carlo (MCMC) iterations followed by 5.10^6 MCMC iterations post-burn-in, carried out on 30 replicates for each value of the tested number of clusters (K) ranging from 2 to 53 or 68 (depending on the focal population data set surveyed). The most likely K was evaluated using the calculation of the *ad hoc* statistic ΔK following Evanno *et al.* (2005). CLUMPP v 1.1 (Jakobsson & Rosenberg 2007) and DISTRUCT v 1.1 (Rosenberg 2004) were used to identify the most likely clustering solution among replicated runs and for graphical display, respectively. As suggested by Puechmaille (2016), we controlled for the effect of uneven sampling schemes: additional analyses were performed with even sample sizes per geographic cluster previously observed. For this, 22 populations (mean 6 per genetic cluster, totalizing 376 individuals) were randomly chosen. As a complementary analysis free of any genetic assumptions, a spatial principal component analysis (sPCA, Jombart *et al.* 2008) was also performed as described in Fig. S5.

Genetic differentiation was assessed using *F*-statistics according to the Weir & Cockerham (1984) ANOVA procedure on nuclear data across loci and populations using SPAGEDI version 1.4

(Hardy & Vekemans 2002). Mono- and multilocus intra-population fixation indices (F_{IS}), and inter-population fixation indices (F_{ST}) were estimated and tested for significance using 10,000 permutations. Mitochondrial genetic differentiation was further assessed using G_{ST}, an estimator equivalent to F_{ST} (Pons & Petit 1996), also using SPAGEDI and tested for significance with 10,000 individual permutations among populations. To evaluate how observed genetic structure is shaped by contemporary gene flow or is a vestige of past historical events, mean mitochondrial G_{ST} estimates were compared to mean N_{ST} estimates, an index of genetic differentiation accounting for phylogenetic distances between mitochondrial haplotypes (Pons & Petit 1996). Likewise, F_{ST} estimates were also compared to the index of genetic differentiation R_{ST} which is based on variance in allelic size (Slatkin 1995). N_{ST} and R_{ST} estimates were compared to G_{ST} and F_{ST} estimates using 10,000 permutations of genetic distance between haplotypes and of allele size respectively, using SPAGEDI.

Statistical comparison of genetic parameters and spatial genetic structure (SGS) patterns

We tested for significant differences in mean genetic diversity (H_E, H_O, A_r, M_r), N_e estimates, mean *IR* and mean pairwise F_{ST} among the main inferred genetic clusters (see above) in the focal study area in northern France. One-way analysis of variance (ANOVA) was used in cases where the data followed normality and homoscedasticity, while a non-parametric Kruskal-Wallis test was used when assumptions were not met. When statistically significant differences were detected, a *post hoc* Tukey-Kramer HSD's test or a *post hoc* Kruskal multi-comparisons test was performed for ANOVA and Kruskal-Wallis test, respectively. For nuclear genetic diversity, a generalized linear model with loci set as random effect, followed by a multi comparison test on means by Tukey's contrast, was further used to test for significant differences among genetic clusters (R packages, "lme4", Bates *et al.* 2015 and "multcomp", Hothorn *et al.* 2008).

The occurrence of SGS was tested at the individual level by applying spatial autocorrelation analyses as implemented in SPAGEDI. Nason's kinship coefficient F_{ij} (Loiselle *et al.* 1995) was chosen as a pairwise estimator of individual genetic relatedness, as it has robust statistical properties (Vekemans & Hardy 2004). To visualize SGS, F_{ij} values were averaged over a set of 10 distance classes defined to obtain the same number of individual pairs within each distance class. Standard errors of F_{ij} were estimated using a jackknifing procedure over loci. Significance of F_{ij} estimates was calculated by 10,000 permutations of individual locations for the nuclear and the

mitochondrial haplotype data set. To quantify and compare the strength of SGS among populations without arbitrarily setting geographical distance intervals, we used the *Sp* statistics (Vekemans & Hardy 2004). Caution is also required when interpreting the extent of positive SGS displayed by the first x-intercept of correlograms (Peakall *et al.* 2003; Favre-Bac *et al.* 2016). To provide an approximation of the true extent of detectable positive SGS, we further performed analyses of cumulative distance classes by increasing distance class sizes spanning the minimum distance between samples to the maximum distance of sampling, as described in Peakall *et al.* (2003). Finally, isolation by distance (IBD) patterns were investigated using polynomial regression and Mantel tests of association as described in Oden & Sokal (1986) and Smouse *et al.* (1986) to estimate the significance of relationships between pairwise D_{CE} genetic distances and pairwise geographical distances among populations.

Evolutionary history of populations in the focal study area

Based on the genetic clusters depicted in the data set, approximate Bayesian computations (ABC) analyses were used to determine the historical process involved in the settlement of populations. We evaluated whether the observed genetic variability in northern France resulted from (i) an eastern continental origin, (ii) a western coastal origin, or (iii) another unsampled native locality. The objective was also to discriminate between recent human-mediated translocations and colonization from populations historically located in inland habitats in the neighborhood of the mining area. Divergence times were estimated to gain further insights into evolutionary history chronology.

Scenarios were evaluated in a two-step analysis. A first set of scenarios (set A) aimed to analyze the origin of each target genetic unit in northern France by using the nuclear genetic units depicted outside the focal study area in northern France as putative sources (Fig. 2A). Based on the results yielded by set A, a second set of scenarios (set B) aimed to depict the relationship among the different target genetic units and to estimate their divergence times taking into account all the observed genetic units (Fig. 2B). Overall, three demographic scenarios were defined for each analysis step following each of the three hypotheses on population origins. ABC analyses were computed using the software DIYABC v.2 (Cornuet *et al.* 2014).

For all scenarios, an identical set of microsatellite loci was simulated. Due to computational limitations and the infinite number of possible scenarios when numerous populations are

considered, inferences were based on finite set of geographically delimited groups. These genetically homogenous groups were defined using Bayesian clustering analysis (see above). In each group of the focal study area, populations with the smallest nuclear pairwise genetic differentiation index (F_{ST} < 0.075) were selected (N=95, N=132 and N=65 for coastal, western coalfield and eastern coalfield populations, respectively) to represent the whole group for comparisons with data simulated under the different scenarios, as suggested in Lombaert *et al.* (2014). For the genetic units outside the focal study area, all individuals belonging to each external genetic clusters were included to guarantee sufficient sample size (N= 94 in western lineage and N= 103 in eastern lineage). The origin(s) of target genetic units in the focal study area was then tested based on adjacent groups and by simplifying the scenarios as much as possible to avoid poor estimation of parameters (Bertorelle *et al.* 2010). Assuming post-glacial recolonization hypothesis with leading-edge expansions along the Atlantic coastline (Rowe *et al.* 2006), scenarios were set with divergence time in generations (t_1 to t_4), putative ancestral effective population size (N_a), current effective population size estimates (N_1 to N_5) and effective sizes of founder populations (N_{f1}, N_{f2}, N_{f3}, N_{f4}) (see Table S2). Analyses were performed with historical, demographic and mutational parameter values drawn from the prior distributions described in Table S2 (*e.g.*, Cornille *et al.* 2013; Dussex *et al.* 2014; Lombaert *et al.* 2014). Detailed steps of each analysis are described in Table S2.

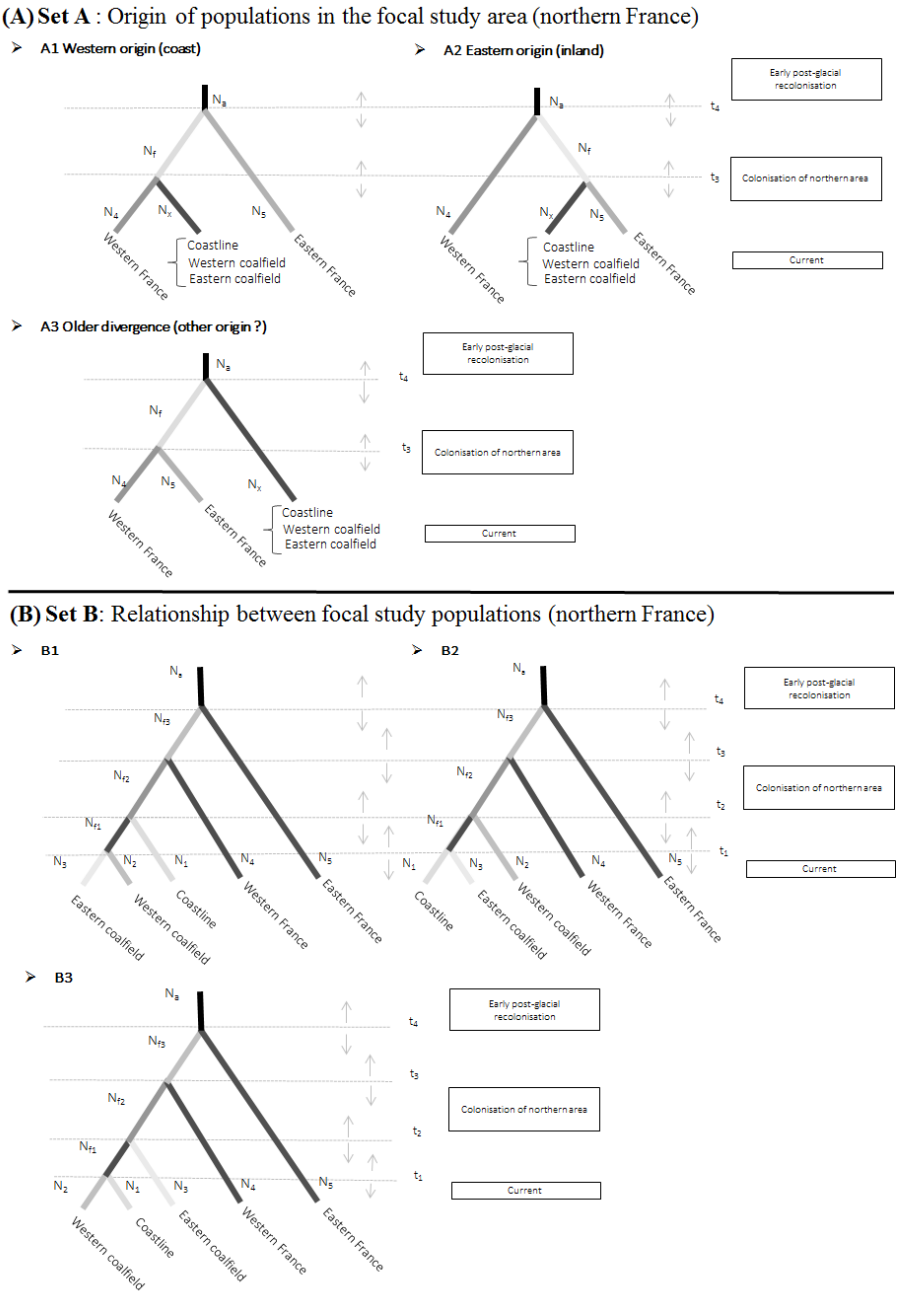


Figure 2: Graphical representation of competing scenarios of the settlement of *Bufo [Epidalea] calamita* in the focal study area in northern France, tested by approximate Bayesian computation. N_a, N_1 to N_5, N_{f1} to N_{f4} refer to effective population sizes of putative ancestral, standing populations and past transitory populations, respectively, and, t_1 to t_4 refer to divergence times (prior settings of model parameters are given in Table S2). The first set of scenarios (A) aims to determine the origins of each northern lineage by evaluating the relationships of each northern target unit with respect to lineages located outside northern France (eastern and western France lineages). The second set of scenarios (B) aims to determine the northern France lineage relationships taking into account all the lineages observed. All hypotheses assumed a shared ancestor and ancient divergence of western and eastern France lineages from the last climate cooling (see Results on genetic affiliations).

Results

Genetic diversity and contemporary N_e

Among the 37 nuclear microsatellite loci, 2 were monomorphic (Buca3 and BC38, see Table S1). Exact tests demonstrated no LD among loci. Two loci, Bcalμ9 and Bc22, were excluded due to technical artifacts (*i.e.,* null alleles and/or allele drop-out). The remaining loci perfectly matched expectations of HW proportions, with A_r ranging from 1.59 to 5.01 (Table S1). Levels of intra-population genetic diversity were very variable with the Swedish population being characterized by the lowest level of nuclear and mitochondrial diversity (Table 1). Based on the 35 nuclear microsatellite loci used in subsequent analyses, multilocus A_T ranged from 52 alleles to 157 alleles per population. Sixteen out 62 populations showed an overall significant deviation from HW equilibrium, which was due to only few significant single-locus effects, involving either excess or deficit in heterozygotes (Table 1). Effective population sizes ranged from 7 to 1506, but were mostly comprised between 10 and 60, with a mean of 93 and a median of 45 (Table 1).

In terms of mitochondrial genetic diversity, 21 haplotypes were found with three major haplotypes accounting for 58% of the mitochondrial diversity, and other haplotypes differing from one of these major haplotypes by one to four mutation steps. Haplotypes were closely related because only 16% of the haplotype pairwise comparisons indicated haplotypic divergence of more than five mutation steps (Fig. 3A). Over the whole dataset, populations exhibited from one to five different haplotypes (Table 1 and Fig. 3B). Only one major haplotype (Hap 11) was shared by the focal study populations in northern France and the other populations (Fig. 3B). Interestingly, of the 12 haplotypes which were exclusively observed in the focal study area, 8 haplotypes occurred only in coalfield populations (Fig. 3B).

Population genetic affiliation and levels of genetic differentiation

We found significant genetic differentiation over all sampled populations, with a mean multilocus F_{ST} of 0.177 ±0.008 ($P < 0.001$) for the nuclear data and a mean G_{ST} of 0.487 ($P < 0.001$) for the mitochondrial data. 98.6% of nuclear pairwise F_{ST} and 89% of mitochondrial pairwise G_{ST} were significant, ranging from 0.00 to 0.582 and from 0.00 to 1.000, respectively. Additionally, a significant phylogeographical signal was detected for both nuclear and mitochondrial data when using a random balanced subset of 22 populations over the whole sampled area to avoid bias due to uneven sampling (see Table S3).

Clear genetic partitioning was detected over the whole data set of sampled populations and within the focal study area in northern France. Bayesian clustering, sPCA and NJ tree results concurred, indicating the occurrence of five distinct genetic clusters. Bayesian clustering analysis suggested two levels of structure in the whole data set because the K vs ΔK distribution was multimodal, with a hierarchical structure at $K = 2$ and at $K = 5$ (Fig. S2A). When $K = 2$, the inferred genetic clusters consisted of one group composed of individuals from the focal study area (northern France) and one group of individuals from other regions. This result arose due to uneven sampling efforts. Additional runs involving randomly chosen, more balanced data sets showed a clear partitioning, splitting the subset into three genetic clusters according to geographical location (Fig. S3). The second modal $K = 5$ obtained in the analysis of the whole data set showed genetic clusters in clear concordance with geographical location (Fig. 1B). Very few admixed individuals occurred within populations: more than 88% of individuals from western France (WF-1 to WF-5) and eastern France (E-1 to E-6) were assigned to their respective cluster with at least 90% membership probability. In the focal study area (northern France), individuals were split among three genetically distinct clusters according to their spatial distribution: coastline populations, western and eastern coalfield populations with admixed populations in the western coalfield cluster (Fig. 1B). Genetic affinities among populations depicted by tree topology (Fig. S4) and sPCA (Figure S5A) unambiguously confirmed this pattern. Furthermore, mitochondrial data mirrored spatial affiliation patterns depicted by nuclear data: each of the five clusters identified were characterized by a specific distribution of haplotypes (Fig. 3B).

When focusing only on the focal study area in northern France, Bayesian analyses confirmed the genetic divergence among coastline, western coalfield and eastern coalfield populations and yielded identical patterns of clustering, with two major modes at $K = 2$ (coastal *versus* coalfield populations) and $K = 3$ (coastal, western and eastern coalfield populations) (Fig. S1B). The sPCA synthetic map of the first three global scores distinguished identical population affiliation along with a further genetic break between southern and northern littoral populations (Fig. S5B). An overall significant genetic differentiation still occurred with a mean multilocus F_{ST} estimate of 0.110 ±0.009 ($P < 0.001$) and $G_{ST} = 0.403$ ($P < 0.001$). No phylogeographical signal was detected within the focal study area in northern France: R_{ST} (0.109; $P < 0.001$) and N_{ST} (0.413; $P < 0.001$) did not significantly differ from the F_{ST} and G_{ST} (Table S3).

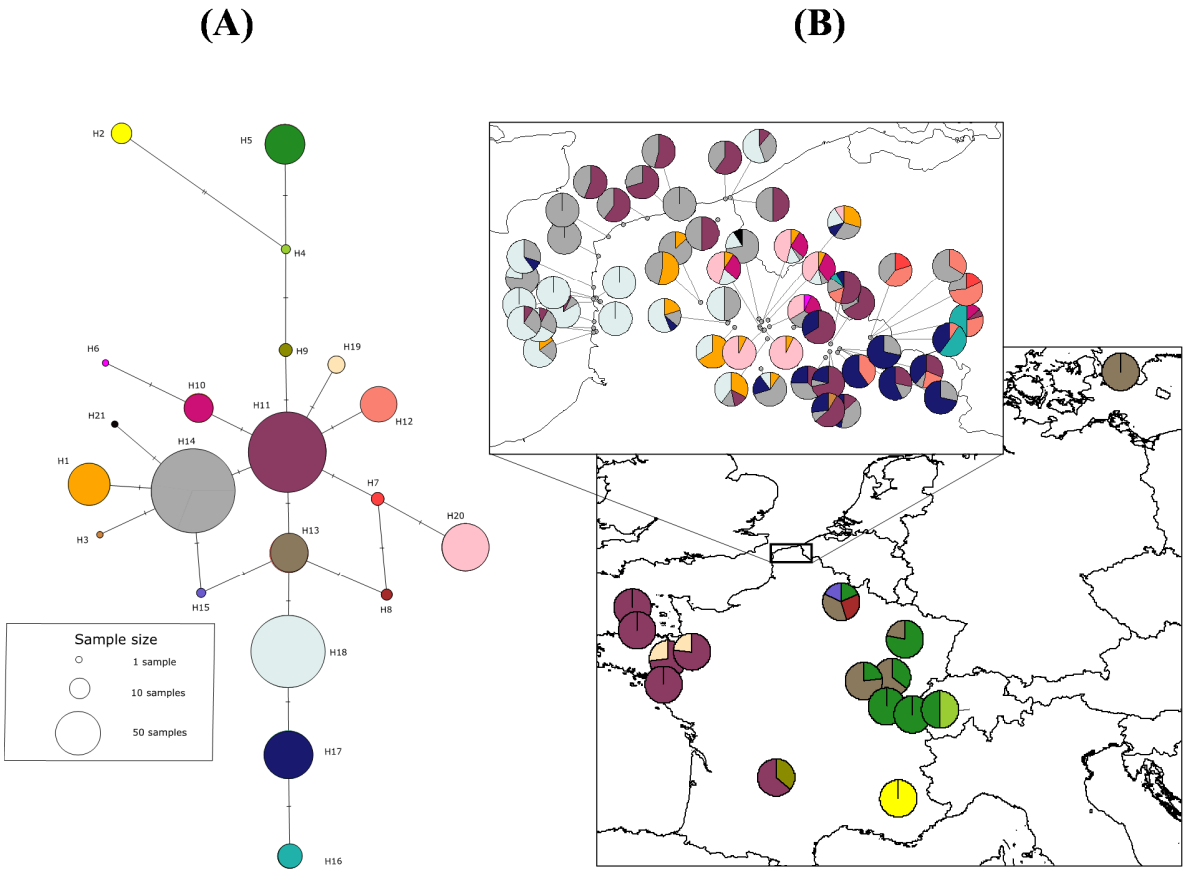


Figure 3: (A) Median joining network depicting the relationships among 21 haplotypes (H1 to H21) observed in 68 sampled populations of *Bufo [Epidalea] calamita*. Mitochondrial diversity is based on a combination of 15 mitochondrial DNA SNPs coded as a single haplotype. Each black trait represents one nucleotide difference. (B) Maps showing the geographical distribution of mitochondrial diversity for 68 sampled populations of *B. calamita*. Inset map refers to the fine-scaled geographical distribution of populations sampled in the focal study area in northern France. Colors refer to the haplotype identities given in the haplotype network.

Statistical comparison of genetic parameters and SGS patterns

Levels of genetic diversity and inbreeding were compared among the three groups defined within the focal study area in northern France. No significant difference occurred regarding F_{IS} or A_r (Fig. 4 A, B). Similarly, no significant difference was observed in terms of single-sample N_e estimates (Fig. 4C, Kruskal Wallis; $\chi^2 = 6.2472$; $P > 0.05$). Nonetheless, as showed by H_e, H_o and M_r estimates, coastal populations appeared to be significantly less genetically diverse compared with eastern coalfield populations or both eastern and western coalfield populations (ANOVA, χ^2 =27.664, df=2, $P < 0.001$, ANOVA, χ^2 =24.305, df=2, $P < 0.001$, Kruskal Wallis; $\chi^2 = 11.441$; $P <$ 0.01 respectively, Fig. 4 D, E and F). Coastal populations also showed a significantly higher - individual inbreeding level (ANOVA; $F = 9.184$; $P < 0.001$) compared with coalfield populations

(Fig. 4 G). Trends for lower mean pairwise nuclear F_{ST} but higher mean mitochondrial pairwise G_{ST} were observed in coastal populations (ANOVA; $F = 3.624$; $P < 0.05$ and Kruskal Wallis; $\chi^2 = 16.637$; $P < 0.001$, Fig. 4 H and I).

SGS depicted by the correlograms in Figure 5 showed a consistent pattern of decreasing genetic relatedness with increasing geographical distance. Strength of SGS was 5 to 60-fold higher for mitochondrial DNA data set compared with the nuclear DNA data set (Fig. 5, A, B, C, D, E, and F). There was a considerable difference in the extent of SGS between coastal populations and inland populations. Nuclear kinship coefficient dropped to zero at 40 km until reaching significant negative value around 80 km in the case of coastal populations whereas this drop occurred at much shorter spatial scale (≅5 km) in the two coalfield clusters (Fig. 5, A, B and C). Similar patterns hold for mitochondrial data, especially for coalfield populations for which a more compact spatial-autocorrelation pattern was found (Fig. 5, D, E and F). Analyses of cumulative distance classes unambiguously confirmed this contrasting extent of positive SGS (Fig. S6). A significant IBD both along the coastline and within the mining area was observed ($r_z = 0.514$ along the coastline, $r_z = 0.693$ in western coalfield and $r_z = 0.516$ in eastern coalfield; all at $P < 0.001$, Fig. 5 G, H and I). Linear relationships clearly hold only at very short geographical distances in the coalfield area (< 5 km) whereas IBD extends over larger geographical distances along the coastline.

Evolutionary history of populations in the focal study area

Pre-evaluation of prior-parameter combinations and results of model checking can be found in Table S5 and Fig. S7. For the first sets of scenarios aiming to evaluate the putative origins of the genetic units observed in northern France, the relative posterior probabilities calculated for each scenario provided the strongest statistical support for scenario A1, suggesting a common western origin for the three northern lineages (P (L) = 0.976, 95% CI [0.967-0.985], P (L) = 0.948, 95% CI [0.936 – 0.960], P (L) = 0.949, 95% CI [0.908-0.990] for the coastal, the western and the eastern coalfield areas, respectively) (see Table S4 for detailed results). The low type II error rate calculated for these most likely scenarios suggest that these probability estimations are robust, with a low probability of choosing a false scenario.

The relative posterior probabilities calculated for each scenario of simulation set B, that aimed at deciphering the relationships among the different genetic units of Northern France, provided the strongest statistical support for scenario B1 (P(L) = 0.913, 95% CI [0.892-0.934]),

which also showed the lowest type II error rate (Fig. 2 B). This scenario suggests that the eastern cluster diverged 300,000 years ago (Table S4 and S5), followed by the western cluster, during the last glacial maximum (138,000 years ago). Finally, the coastal genetic cluster divergence was dated to 52,000 years ago and the within-mining area divergence shaping the eastern and western genetic differentiation was estimated to 14,000 years ago.

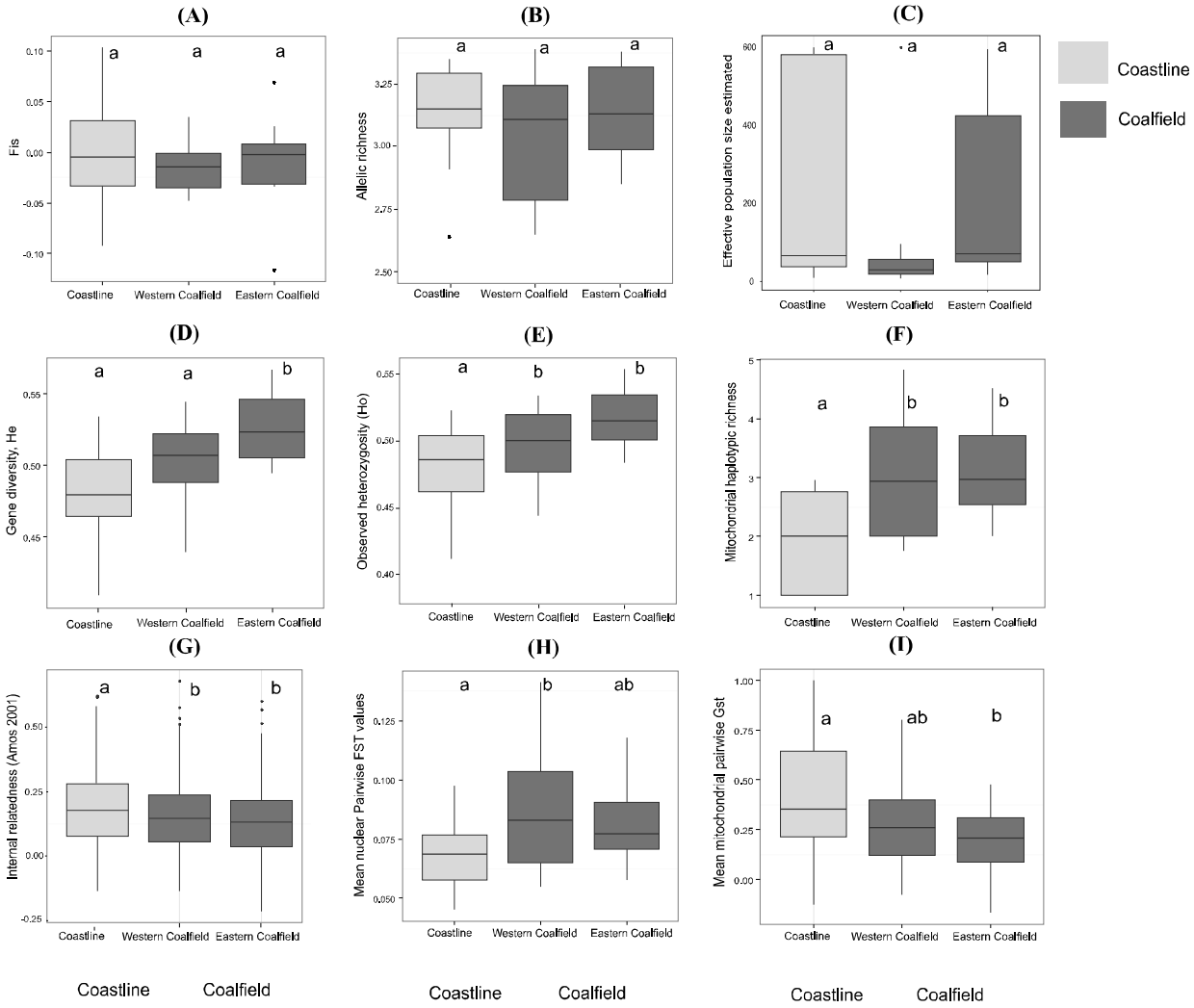


Figure 4: Variation in (A) mean multilocus intra-population fixation index (F_{IS}), (B) nuclear allelic richness (A_r), (C) effective population size estimates on nuclear data, (D) gene diversity (H_e) and (E) observed heterozygosity (H_o) among the three genetic clusters depicted within the focal study area in northern France: coastal population cluster, western coalfield population cluster and eastern coalfield population cluster. Also indicated are (F) mitochondrial haplotypic richness, (G) internal relatedness *IR*, (H) mean pairwise F_{ST} estimates and (I) mean mitochondrial pairwise G_{ST}. Significant differences tested by a Tukey-Kramer HSD test or Kruskal Wallis multi-comparison test are indicated by letters [a], [b] and [c].

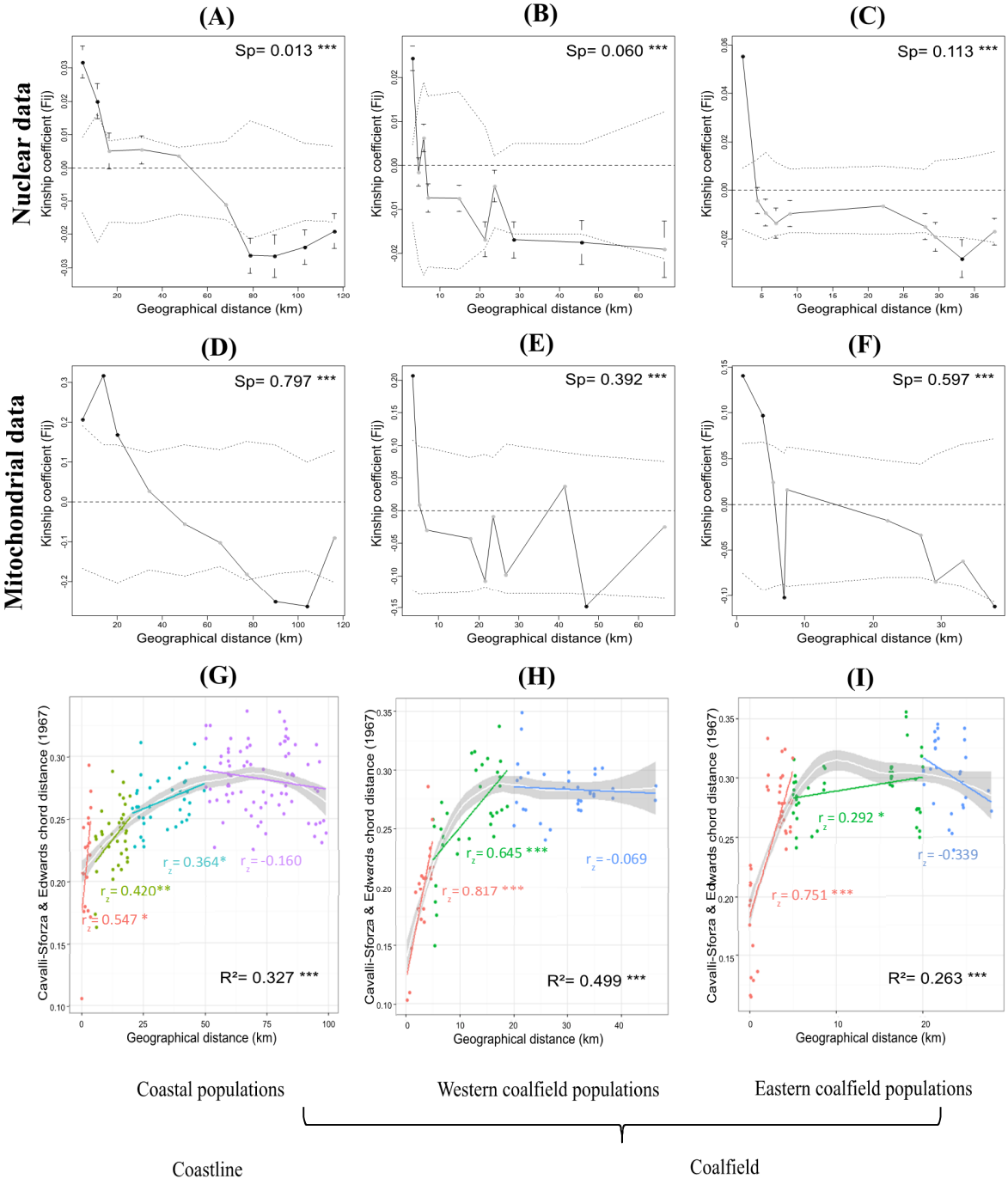


Figure 5: Spatial genetic structure in *Bufo [Epidalea] calamita* populations located within the focal study area in northern France and defined as belonging to three distinct genetic clusters: coastline populations, western coalfield populations and eastern coalfield populations. Average pairwise kinship coefficients (F_{ij}, Loiselle *et al.* 1995) between individuals are plotted against geographical distance for nuclear microsatellites (A, B, C) and mitochondrial haplotypes (D, E, F) for the three clusters of populations. Dashed lines indicate the upper and lower 95% CI limits of no significant spatial genetic structure (● $P < 0.05$, ● $P > 0.05$). Also indicated for each cluster are the *Sp* statistics and their statistical significance, allowing comparisons of the strength of spatial genetic structure. Scatterplots of chord genetic distance (D_{CE}; Cavalli-Sforza & Edwards 1967) against pairs of Euclidean geographical distance for the three genetic clusters (G, H, I). Polynomial regression lines are plotted

for spatial scales indicated in color, orange: < 5 km, green:] 5 km, 20 km], blue:] 20 km, 50 km [and purple: >50 km. The number of pairwise comparisons is comprised between 18 and 78. Mantel coefficients r_z computed for each spatial scale and regression adjusted R-squared (R^2) are also indicated. $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Discussion

Intraspecific genetic variation is a key component of biodiversity: it determines ecological and evolutionary outcomes such as the potential of a given species to respond to environmental changes associated with human activities (DiBattista 2008; Alberti 2015; Mimura *et al.* 2017). Recent studies investigated the impact of anthropogenic land conversion, like urbanization or quarrying activities, on genetic diversity (*e.g.* Flavenot *et al.* 2015; Beninde *et al.* 2016; Lourenço *et al.* 2017). However, the effect of post-industrial areas such as coalfields on patterns of genetic structure remains to be investigated. To the best of our knowledge, among the very few population genetics studies performed in such human-modified environments (*e.g.* Brock *et al.* 2007; Esfeld *et al.* 2008), this work is the first describing in detail current genetic patterns for a vertebrate species. We found that coalfield areas shelter distinct genetic lineages of the natterjack toad, and that these populations are also characterized by different patterns of genetic structure compared to populations found in natural habitats. Our findings and their evolutionary and conservation implications are discussed below.

Level and spatial arrangement of genetic variability in anthropogenic and natural areas

Populations established in urbanized fragmented and/or human-modified colonized habitats may display lower levels of genetic diversity, higher inbreeding and lower contemporary N_e, especially for amphibian species (McCartney-Melstad & Shaffer 2015; Rivera-Ortíz *et al.* 2015). For instance, Munshi-South *et al.* (2013) found a significant decline in genetic diversity in urban populations of the Dusky salamander in New York City. In contrast, Furman *et al.* (2016) found no evidence of genetic erosion and increased genetic differentiation in natural and artificially constructed wetlands in a wood frog, despite extensive urbanization. In the same way, isolated urban populations of the Fire salamander did not exhibit lower levels of genetic diversity and showed similar N_e compared to rural populations (Lourenço *et al.* 2017).

In our study, natural coastal populations of natterjack toad met a regional gene flow/genetic drift equilibrium with a continuous IBD pattern over increasing spatial scales. Conversely, a positive SGS was observed at much shorter spatial scales in inland coalfields areas. This highlights a different regional dynamics, with genetic drift being much more influential than gene flow

beyond a spatial scale of 5 km. Such a departure from gene flow/drift equilibrium is found in a wide array of taxa (*e.g.* Hutchison & Templeton 1999; Austin *et al.* 2004; Hänfling & Weetman 2006; Phillipsen *et al.* 2015). Besides, our results are supported by higher levels of nuclear and mitochondrial genetic differentiation among coalfield populations compared with the coastal populations over a similar geographical area. Overall, this pattern supports a mosaic of genetically distinct populations with genetic breaks occurring over a few kilometers probably due to the unfavorable habitat (highly urbanized or intensive farming) surrounding spoil heaps in coalfield areas. Such pronounced SGS is a common feature in reptilian and amphibian populations where natural or human-induced dispersal barriers are prevailing factors in disrupting gene flow (Vos *et al.* 2001; Arioli *et al.* 2010; Munshi-South *et al.* 2013; Sotiropoulos *et al.* 2013; Beninde *et al.* 2016).

However, despite coalfield populations being isolated due to habitat fragmentation, we observed no decrease in the levels of gene diversity in natterjack toad. In contrast, coalfield populations showed higher levels of genetic diversity and lower inbreeding levels compared with natural coastal populations. Moreover, we did not detect any differences in terms of contemporary N_e between natural coastal and coalfield populations, a result also found in Lourenço *et al*. (2017) when comparing Fire salamanders in rural and urbanized environments. Our N_e estimates ranged from 7 to 600 and were comparable to effective population sizes observed in other pond-breeding amphibians located outside urbanized environments, like Yosemite toads or Fire salamanders (Wang *et al.* 2011; Wang 2012; Lourenço *et al.* 2017). Our N_e estimates were also higher than those estimated by Beebee (2009) in a range of natural British and European continental *B*. *calamita* populations. Inland coalfield estimated values of N_e were even in the upper range of N_e for amphibian populations whose values are usually expected to be under 100 and frequently closer to 10 (reviewed in Phillipsen *et al.* 2011). Altogether, our results support the conservation value of post-industrial areas because they host populations with substantial levels of genetic diversity and consistent contemporary N_e. Our results are very similar to the few studies conducted on similar environments: plant species living in former mining areas, like Orchidaceae, Compositaea, Chenopodiacea or Brassicaceae, also exhibited a high genetic variation that did not differ from natural populations (Brock *et al.* 2007; Esfeld *et al.* 2008; Prinz *et al.* 2009; Pauwels *et al.* 2012).

Whereas lower levels of genetic diversity and higher inbreeding levels in natural coastline populations can be attributed to postglacial leading-edge colonization that involves strong genetic

drift (Rowe *et al.* 2006; Slatkin & Excoffier 2012), the high levels of genetic diversity within the coalfield area may be a result of independent colonization events from surrounding populations. Indeed, colonizers from multiple sources are known to enhance intra-population genetic diversity (Whitlock & McCauley 1990). Among the few available comparable studies, this process of colonization has been suggested in *Suaeda maritima*, a halophyte plant species that repeatedly colonized former potash mining dumps in Germany (Prinz *et al.* 2009, see also Esfeld *et al.* 2008). In our case study, *B. calamita* is known to be a pioneer species with high dispersal capabilities (Miaud *et al.* 2000; Sinsch *et al.* 2012). Thus, the most parsimonious hypothesis is that toads successfully dispersed from natural populations initially established in former semi-natural habitats (humid meadows or pastures) located in the vicinity of the coalfields (Godin 2002). Additionally, accidental human-mediated introductions of individuals from unknown sources, during mining activities or currently by individual translocations, cannot be ruled out, as observed in the European tree frog (Andersen *et al.* 2004). The exceptional levels of mitochondrial diversity found in the two coalfield entities reinforce the likelihood of multiple waves of immigration events from diverse sources. Together, the occurrence of distinct maternal lineages and the absence of equilibrium between gene flow and drift suggest an establishment of colonizers from previous semi-natural areas surrounding the coalfields, potentially facilitated by trade operations related to mining activities.

Population evolutionary history

Local interactions between the micro-evolutionary processes of gene flow and genetic drift do not exclude biogeographic historical processes as key factors shaping the current genetic structure (*e.g.*, Austin *et al.* 2004; Dussex *et al.* 2014). Indeed, it is now well established that Pleistocene climate cycling had a major role in driving the present patterns of genetic structure for several taxa in Europe (Hewitt 1999; Stewart *et al.* 2010). The divergence time estimates in each most likely ABC scenario were consistent and our results suggest a common history but distinct subsequent re-colonization pathways of each evolutionary unit located in the focal area (northern France). The three northern genetic units and the eastern and the western genetic units located south of the focal study area diverged during the Pleistocene from around -300,000 to -14 000 years ago. Different glacial refugia during climatic cooling drove population genetic divergence following a south-to-north pathway: firstly an eastern genetic unit split, secondly a western genetic unit split,

thirdly a northern coastal genetic unit split and finally a recent split of the eastern and western coalfield genetic units in the last 14 000 years. Several studies in a wide array of taxa have already raised the hypothesis of multiple European glacial refugia during the Pleistocene (Hewitt 1999; Cornille *et al.* 2013; Havrdovà *et al.* 2015; Ursenbacher *et al.* 2015). Rowe *et al.* (2006) suggest that local but relatively recent and short-lived glacial refugia, located in northern France and central-eastern Europe from 20,000 to 9000 years ago, explain the lineages observed throughout the natterjack toad geographical range distribution. These long-term isolation events in different refugia may explain the observed divergence between northern genetic units and the western genetic unit. Our salient result was the common ancestry of the three genetic units depicted in the focal study area, dating back around 50,000 years ago. Former inland semi-natural populations that subsequently established in the neighboring coalfield areas have therefore a common history with natural coastal populations. The timescale is however too short to get more precise inferences using ABC analyses. It should also be noted that the occurrence of multiple mtDNA lineages in this post-glacial expansion area may be suggestive of an impact of recent human activities that mixed different post-glacial lineages, which may inflate time divergence estimates (Lombaert *et al.* 2014). Future studies involving assignment tests and a comprehensive sampling across the species' range are required to gain further insights into the evolutionary history of observed lineages.

Evolutionary and conservation implications

Ignoring the nature of intraspecific genetic variation in management decisions can lead to negative consequences for biodiversity conservation (Allentoft & O'Brien 2010; McCartney-Melstad & Shaffer 2015; Mimura *et al.* 2017). We depicted clear genetic boundaries in contrasting habitats, which can be used for defining conservation and management units. Our results (large levels of genetic diversity and N_e) suggest that the erosive effect of genetic drift is not currently a major threat for natterjack populations in coalfield areas. Based on the few comparable studies involving plant populations established in post-mining areas, Prinz *et al.* (2009) argued that this may be the rule rather than an exception. Hence, former mining areas can act as refugia areas and offset natural habitat loss for pioneer species living in early successional habitats (Esfeld *et al.* 2008). Altogether, these results highlight the value of former mining areas for the long-term conservation of animal and plant pioneer species, as found for other non-urbanized areas such as quarries (*e.g.* Flavenot *et al.* 2015). Highly genetically diverse natterjack populations established

in the mining areas have probably benefited from accidental human-mediated transport, lower predation risk and/or reduced competition with the common toad (Beebee & Denton 1996). For natural coastal populations, additional investigations are needed to determine whether population reinforcement is required. Nevertheless, transfers across such ecologically well-differentiated habitats (coastal *vs* coalfield habitats) should be discouraged to avoid possible outbreeding depression. Indeed, local adaptation to coalfield environment conditions might already have been established (*e.g*. Pauwels *et al.* 2012 in Brassicaceae). The challenge is now to replace the past colonizer and migrant flow by ongoing effective regional gene flow because we clearly demonstrated non-equilibrium conditions with gene flow occurring only at short spatial scales (less than 10 km). The protection of spoil heaps and surrounding open habitats is essential for the conservation of viable populations. However, the whole landscape configuration remains a key factor for the long-term management of connecting pathways among populations living in such exceptional areas inherited from European industrial history.

Acknowledgments:

We are grateful to Stephen Spear and three anonymous reviewers for many comments that greatly improved the quality of the manuscript. We would like to thank the following organizations for their invaluable help in tissue collection: EDEN 62, CG59, LPO Bourgogne, LPO Lot, Bretagne Vivante, CEN Rhone-Alpes, CEN Franche Comté, Université de Lausanne, CEN Lorraine and CEN Picardie. José Godin and Claire Blaise provided very helpful informations on the regional species history. We also wish to express our gratitude to Clément Mazoyer for his assistance in computer analyses. Numerical results presented in this paper were carried out using the European Grid Infrastructure (http://www.egi.eu) with the Biomed virtual organization (http://lsgc.org/en/Biomed:home) via DIRAC portal (http://diracgrid.org) supported by France Grille (http://www.france-grilles.fr/). This work was funded by the AMPHIDIV project (Région Nord-Pas De Calais). This work is also a contribution to the CPER research project CLIMIBIO. The authors thank the French Ministry for Higher Education and Research, the Hauts de France Regional Council and the European Regional Development Fund for their financial support of this project.

Data Accessibility

- Nuclear mitochondrial DNA sequences: GenBank accessions KX237573–KX237594.

- Mitochondrial DNA sequences used to define SNP markers: GenBank accessions KX237595–KX237630.

- Coordinates of sample locations, nuclear and mitochondrial genetic data: DRYAD: doi:10.5061/dryad.b89k7.

Author Contributions

JFA, CV, SR and JJ conceived and designed the study. LF, SR, RQ, JJ, LH and JFA conducted the sampling. LF, LH and CG genotyped the samples. LF, CG and JFA designed the new nuclear microsatellite loci and mitochondrial SNPs. LF, JFA, SG and JJ analyzed the data. LF and JFA wrote the manuscript. All authors provided editorial comments.

References

Alberti M. (2015) Eco-evolutionary dynamics in an urbanizing planet. *Trends in Ecology and Evolution*, **30**, 114–126.

Allentoft ME, O'Brien J. (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, **2**, 47–71.

Allentoft ME, Siegismund HR, Briggs L, Andersen LW. (2009) Microsatellite analysis of the natterjack toad (*Bufo calamita*) in Denmark: populations are islands in a fragmented landscape. *Conservation Genetics*, **10**, 15–28.

Amici V, Landi S, Frascaroli F *et al.* (2015) Anthropogenic drivers of plant diversity: perspective on land use change in a dynamic cultural landscape. *Biodiversity and Conservation*, **24**, 3185–3199.

Andersen LW, Fog K, Damgaard C. (2004) Habitat fragmentation causes bottlenecks and inbreeding in the European tree frog (*Hyla arborea*). *Proceedings of the Royal Society of London, Series B*, **271**, 1293–1302.

Arioli M, Jakob C, Reyer H-U. (2010) Genetic diversity in water frog hybrids (*Pelophylax esculentus*) varies with population structure and geographic location. *Molecular Ecology*, **19**, 1814–1828.

Austin JD, Lougheed SC, Boag PT. (2004) Controlling for the effects of history and nonequilibrium conditions in gene flow estimates in northern bullfrog (*Rana catesbeiana*) populations. *Genetics*, **168**, 1491–1506.

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C. (2013) Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, **88**, 310–326.

Bates D, Mächler M, Bolker B, Walker S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1–48.

Beebee TJC. (2009) A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: genetic compensation and population size-genetic diversity correlations. *Molecular Ecology*, **18**, 4790–4797.

Beebee TJC, Denton JS. (1996) *The natterjack toad conservation handbook (Species recovery program)*. English Nature.Northminster House, Peterborough PE1 1UA.

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

Beninde J, Feldmeier S, Werner M *et al.* (2016) Cityscape genetics: structural vs. functional connectivity of an urban lizard population. *Molecular Ecology*, **25**, 4984–5000.

Bertorelle G, Benazzo A, Mona S. (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Brock J, Aboling S, Stelzer R, Esch E, Papenbrock J. (2007) Genetic variation among different populations of Aster tripolium grown on naturally and anthropogenic salt-contaminated habitats: implications for conservation strategies. *Journal of Plant Research*, **120**, 99–112.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L. (2006) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics*, **8**, 509–511.

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

Cornille A, Giraud T, Bellard C *et al.* (2013) Postglacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Molecular Ecology*, **22**, 2249–2263.

Cornuet J-M, Pudlo P, Veyssier J *et al.* (2014) DIYABC v2.0 : a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187–1189.

Coulon A. (2009) GENHET: an easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources*, **10**, 167–169.

Denton JS, Hitchings SP, Beebee TJC, Gent A. (1997) A recovery program for the natterjack toad (*Bufo calamita*) in Britain. *Conservation Biology*, **11**, 1329–1338.

DiBattista JD. (2008) Patterns of genetic variation in anthropogenically impacted populations. *Conservation Genetics*, **9**, 141–156.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (*Ne*) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Dussex N, Wegmann D, Robertson C. (2014) Postglacial expansion and not human influence best explains the population structure in the endangered kea (*Nestor notabilis*). *Molecular Ecology*, **23**, 2193–2209.

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

Esfeld K, Hensen I, Wesche K *et al.* (2008) Molecular data indicate multiple independent colonizations of former lignite mining areas in Eastern Germany by *Epipactis palustris* (Orchidaceae). *Biodiversity and Conservation*, **17**, 2441–2453.

Evanno G, Regnaut S, Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Faucher L, Godé C, Arnaud J-F. (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo* (*Epidalea*) *calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). *Journal of Heredity*, **107**, 660–665.

Favre-Bac L, Mony C, Ernoult A, Burel F, Arnaud J-F. (2016) Ditch network sustains functional connectivity and influences patterns of gene flow in an intensive agricultural landscape. *Heredity*, **116**, 1–13.

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

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

Furman BLS, Scheffers BR, Taylor M, Davis C, Paszkowski CA. (2016) Limited genetic structure in a wood frog (*Lithobates sylvaticus*) population in an urban landscape inhabiting natural and constructed wetlands. *Conservation Genetics*, **17**, 19–30.

Godin J. (2002) Degré de rareté, évolution de la distribution et particularités de l'herpétofaune de la région Nord - Pas de Calais. *Bulletin de la Société Herpétologique de France*, **104**, 16–35.

Hänfling B, Weetman D. (2006) Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the river sculpin, *Cottus gobio*. *Genetics*, **173**, 1487–1501.

Hardy OJ, Vekemans X. (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Havrdovà A, Douda J, Krak K *et al.* (2015) Higher genetic diversity in recolonized areas than in refugia of *Alnus glutinosa* triggered by continent-wide lineage admixture. *Molecular Ecology*, **24**, 4759–4777.

Hewitt GM. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hothorn T, Bretz F, Westfall P. (2008) Simultaneous inference in general parametric models. *Biometrical Journal*, **50**, 346–363.

Hubisz MJ, Falush D, Stephens M, Pritchard JK. (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

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.

Jakobsson M, Rosenberg NA. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

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

Kalinowski ST. (2005) HP-RARE 1.0 : a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Lemoine G. (2012) Interêt, gestion, protection et valorisation des terrils miniers dans le Nord–Pas-de-Calais. *Annales de la Société Géologique du Nord*, **19**, 1–8.

Loiselle BA, Sork VL, John N, Graham C. (1995) Spatial genetic structure of a tropical understory shurb, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

Lombaert E, Guillemaud T, Lundgren J *et al.* (2014) Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, **23**, 5979–5997.

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

McCartney-Melstad E, Shaffer HB. (2015) Amphibian molecular ecology and how it has informed conservation. *Molecular Ecology*, **24**, 5084–5109.

Miaud C, Sanuy D, Avrillier J-N, Sanuy D, Miaud C. (2000) Terrestrial movements of the natterjack toad *Bufo calamita* (Amphibia, Anura) in a semi-arid, agricultural landscape. *Amphibia-Reptilia*, **21**, 357–369.

Mimura M, Yahara T, Faith DP *et al.* (2017) Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evolutionary Applications*, **10**, 121–139.

Munshi-South J, Zak Y, Pehek E. (2013) Conservation genetics of extremely isolated urban populations of the northern dusky salamander (*Desmognathus fuscus*) in New York City. *PeerJ*, **1**, e64.

Oden NL, Sokal RR. (1986) Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology*, **35**, 608–617.

Pauwels M, Vekemans X, Godé C *et al.* (2012) Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **193**, 916–928.

Peakall R, Ruibal M, Lindenmayer DB. (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. *Evolution*, **57**, 1182–1195.

Phillipsen IC, Funk WC, Hoffman EA, Monsen KJ, Blouin MS. (2011) Comparative analyses of effective population size within and among species: ranid frogs as a case study. *Evolution*, **65**, 2927–2945.

Phillipsen IC, Kirk EH, Bogan MT, Mims MC, Julian D. (2015) Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects moderate dispersal. *Molecular Ecology*, **24**, 54–69.

Pons O, Petit EJ. (1996) Measuring and testing genetic differentiation with ordered vs. unordered alleles. *Genetics*, **144**, 1237–1245.

Prinz K, Weising K, Hensen I. (2009) Genetic structure of coastal and inland populations of the annual halophyte *Suaeda maritima* (L.) *dumort*. in Central Europe, inferred from amplified fragment length polymorphism markers. *Plant Biology*, **11**, 812–820.

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

Puechmaille SJ. (2016) The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, **16**, 608–627.

Raymond M, Rousset F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

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

Rivera-Ortíz FA, Aguilar R, Arizmendi MDC, Quesada M, Oyama K. (2015) Habitat fragmentation and genetic variability of tetrapod populations. *Animal Conservation*, **18**, 249–258.

Rodrigues ASL, Andelman SJ, Bakarr MI *et al.* (2004) Effectiveness of the global protected area network in representing species diversity. *Nature*, **428**, 640–643.

Rogell B, Gyllenstrand N, Höglund J. (2005) Six polymorphic microsatellite loci in the Natterjack toad, *Bufo calamita*. *Molecular Ecology Notes*, **5**, 639–640.

Rosenberg NA. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

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

Rousset F, Raymond M. (1995) Testing heterozygote excess. *Genetics*, **140**, 1413–1419.

Rowe G, Beebee TJC, Burke T. (1997) PCR primers for polymorphic microsatellite loci in the anuran amphibian *Bufo calamita*. *Molecular Ecology*, **6**, 401–402.

Rowe G, Beebee TJC, Burke T. (2000) A further four polymorphic microsatellite loci in the natterjack toad *Bufo calamita*. *Conservation Genetics*, **1**, 371–372.

Rowe G, Harris DJ, Beebee TJC. (2006) Lusitania revisited: a phylogeographic analysis of the natterjack toad *Bufo calamita* across its entire biogeographical range. *Molecular Phylogenetics and Evolution*, **39**, 335–346.

Sinsch U, Oromi N, Miaud C, Denton JS, Sanuy D. (2012) Connectivity of local amphibian populations: modelling the migratory capacity of radio-tracked natterjack toads. *Animal Conservation*, **15**, 388–396.

Slatkin M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Slatkin M, Excoffier L. (2012) Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, **191**, 171–181.

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

Sotiropoulos K, Eleftherakos K, Tsaparis D *et al.* (2013) Fine scale spatial genetic structure of two syntopic newts across a network of ponds: implications for conservation. *Conservation Genetics*, **14**, 385–400.

Stewart JR, Lister AM, Barnes I, Dalen L. (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London, Series B*, **277**, 661–671.

Ursenbacher S, Guillon M, Cubizolle H *et al.* (2015) Postglacial recolonization in a cold climate specialist in western Europe: patterns of genetic diversity in the adder (*Vipera berus*) support the central – marginal hypothesis. *Molecular Ecology*, **24**, 3639–3651.

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

Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJ. (2001) Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity*, **86**, 598–608.

Wake DB, Vredenburg VT. (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11466–11473.

Wang IJ. (2012) Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. *Diversity and Distributions*, **18**, 1033–1041.

Wang J. (2016) A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*, **25**, 4692–4711.

Wang IJ, Johnson JR, Johnson BB, Shaffer HB. (2011) Effective population size is strongly correlated with breeding pond size in the endangered California tiger salamander, *Ambystoma californiense*. *Conservation Genetics*, **12**, 911–920.

Waples RS. (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167–184.

Waples RS, Antao T, Luikart G. (2014) Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, **197**, 769–780.

Waples RS, England PR. (2011) Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, **189**, 633–644.

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

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

Wilkinson JW, Griffiths RA. (2013) Amphibian declines and conservation in Britain. In: Amphibian biology. Status of conservation and decline of amphibians: Eastern hemisphere. Western Europe (eds Heatwole H, Wilkinson JW), pp. 48–57. Pelagic Publishing.Exeter, UK

Supplementary files

Table S1: Estimates of nuclear genetic variation for 37 nuclear microsatellite loci over 62 populations (N>7) of *Bufo [Epidalea] calamita* from Western Europe: total number of alleles (A_T), allelic richness (A_r), observed heterozygosity (H_O), expected heterozygosity (H_E) and mean fixation indices (F_{IS}, F_{ST}, according to Weir & Cockerham 1984). Significance of F_{ST} was tested using an exact G-*test* with 10,000 iterations for each locus over all populations. Heterozygosity excess and deficit (estimated on F_{IS}) were tested using the multiscore sample U-test. Both tests were corrected using Bonferroni adjustment (Rice 1989).* $P < 0.05$; *** $P < 0.001$

Locus	Reference*	Multiplex	Dye	A_T	A_r	H_O	H_E	F_{IS}	F_{ST}
Bcal7	1	1	NED	8	2.333	0.359	0.36	-0.019	0.193 ***
Bcal1	1	1	FAM	18	5.011	0.763	0.74	-0.034 *	0.145 ***
Bcal2	1	1	PET	7	2.517	0.373	0.421	0.106 ***	0.171 ***
Bcal11	1	2	FAM	10	2.689	0.397	0.418	0.036	0.173 ***
Buca2	2	2	VIC	7	3.179	0.576	0.578	-0.002	0.196 ***
Buca3	2	2	PET	1	1.000	0	0	-	-
Buca5	2	2	NED	9	1.801	0.196	0.194	0.024	0.306 ***
Bcal3	2	3	FAM	16	3.804	0.561	0.610	0.106 ***	0.186 ***
Bcal4	2	3	PET	12	2.795	0.548	0.535	-0.02	0.188 ***
Bcal5	2	3	FAM	9	2.623	0.424	0.413	-0.027	0.234 ***
Bcal6	2	3	VIC	11	3.495	0.625	0.623	-0.024 *	0.144 ***
Bcal10	3	4	NED	10	3.382	0.600	0.62	0.011	0.2 ***
Bcal9	3	4	FAM	11	2.762	0.292	0.426	0.316 ***	0.299 ***
Buca6	2	4	FAM	16	3.670	0.63	0.644	0	0.173 ***
Buca1	2	5	NED	8	2.201	0.300	0.338	0.103 ***	0.205 ***
BC05	4	6	NED	8	2.692	0.496	0.493	-0.016	0.224 ***
BC08	4	6	HEX	21	4.172	0.654	0.66	0.017	0.137 ***
BC11	4	6	HEX	11	4.835	0.744	0.741	-0.023	0.127 ***
BC19	4	6	FAM	7	2.443	0.401	0.404	-0.002	0.261 ***
BC22	4	6	FAM	9	3.510	0.495	0.623	0.213 ***	0.187 ***
BC29	4	6	NED	9	3.259	0.547	0.575	0.031	0.158 ***
BC46	4	6	FAM	10	3.303	0.568	0.557	-0.013	0.188 ***

BC38	4	6	PET	1	1.000	0	0	-	-
BC01	4	7	FAM	14	2.943	0.486	0.479	-0.007	0.204 ***
BC02	4	7	FAM	12	4.376	0.679	0.693	0.009	0.13 ***
BC04	4	7	NED	8	2.965	0.491	0.484	-0.026	0.226 ***
BC24	4	7	NED	6	2.944	0.587	0.591	-0.004	0.132 ***
BC25	4	7	PET	3	1.975	0.282	0.267	-0.067	0.192 ***
BC28	4	7	HEX	4	2.643	0.544	0.495	-0.131 ***	0.206 ***
BC34	4	7	HEX	5	2.247	0.463	0.451	0.010	0.133 ***
BC35	4	7	PET	10	1.595	0.148	0.148	-0.021	0.232 ***
BC15	4	8	FAM	8	1.810	0.183	0.159	0.128***	0.434***
BC18	4	8	FAM	16	4.120	0.684	0.685	-0.007	0.141***
BC37	4	8	HEX	10	4.269	0.708	0.669	0.022*	0.123***
BC39	4	8	HEX	9	3.821	0.641	0.660	-0.032	0.145***
BC45	4	8	PET	10	4.664	0.712	0.716	-0.011	0.162***
BC09	4	8	NED	11	4.143	0.673	0.662	-0.006	0.156***
Mean				9.595	3.054	0.492	0.508	0.014	0.181***

***1**: (Rowe *et al.* 1997), **2**: (Rogell *et al.* 2005), **3**: (Rowe *et al.* 2000a), **4**: (Faucher *et al.* 2016)

Table S2: Prior settings of historical and demographic parameters used in approximate Bayesian computations to simulate and compare scenarios of *Bufo* [*Epidalea*] *calamita* coalfield colonization in northern France, described in Figure 2 and based on the resulting reference table and the simulated pseudo-observed genetic data sets. Timing of events is in generations, assuming a generation time of 3.5 years in *B. calamita*. *B. calamita* reaches maturity at a minimum of 2 years of age for males and a maximum of 5 years for females (Stevens *et al.* 2003; Rowe & Beebee 2004; Leskovar *et al.* 2006)**. The sex-ratio was set to an equal number of males and females** (Denton *et al.* 1997; May *et al.* 2011)**. The microsatellite loci were assumed to follow a generalized stepwise mutation model with three parameters: the mean mutation rate (mean μ), the mean parameter of the geometric distribution (mean P) of allele length in terms of the number of repeated mutation events and the mean mutation rate for single nucleotide instability (mean μ_{SNI}). Because all loci have range of less than 40 contiguous allelic states, we set their ranges to 40 as suggested by Cornuet *et al.* (2014). Each locus is characterized by individual μ_loc drawn from a gamma distribution (mean = mean μ and shape = 2), P_loc drawn from a gamma distribution (mean = mean P and shape = 2) and μ_{SNI}_loc drawn from a gamma distribution (mean = mean μ_{SNI} and shape = 2).**

Parameter	Parameter name	Prior distribution	
		Type	Interval
Ancestral effective population size	N_a	uniform	{10-100 000}
Effective population size before founding event	$N_{f[1-4]}$	uniform	{10-100 000}
Effective population size of the northern coastal populations	N_1	uniform	{10-100 000}
Effective population size of the western coalfield populations	N_2	uniform	{10-100 000}
Effective population size of the eastern coalfield populations	N_3	uniform	{10-100 000}
Effective population size of the western populations	N_4	uniform	{10-100 000}
Effective population size of the eastern populations	N_5	uniform	{10-100 000}
Number of generations since the first population divergence event	t_4	uniform	{10-200 000}
Number of generations since the second population divergence event	t_3	uniform	{10-100 000}
Number of generations since the third population divergence event	t_2	uniform	{10-100 000}
Number of generations since the fourth population divergence event	t_1	uniform	{10-80 000} $t_4 > t_3$ $t_3 > t_2$ $t_2 \geq t_1$
Mutation model parameters			
Mean mutation rate	μ	uniform	$\{10^{-9}\text{-}10^{-3}\}$
Mean parameter of the geometric distribution of the number of repeats	P	uniform	{0.001-0.3}
Mean single nucleotide insertion/deletion mutation rate	μ_{sni}	uniform	$\{10^{-8}\text{-}10^{-5}\}$

Summary statistics, prior distributions and simulations

A first set of one-sample summary statistics (the mean number of alleles and the mean expected heterozygosity in the population and pairwise summary statistics, the population genetic differentiation index and the shared allele distance) allowed us to compute the probability of scenarios, to discriminate them and to draw the posterior distribution of the parameters based on simulations.

A pre-evaluation of scenario-prior combinations was applied using a principal component analysis (PCA) on 1000 simulated datasets to evaluate if at least some parameter combinations provided simulated datasets that fit the observed dataset and, then, if necessary, to adjust the scenario design.The reference table was built on 3.10^6 simulated genetic datasets (around 10^6 per scenario). The posterior probabilities of the competing scenarios were then estimated by polychotomous logistic regression (Cornuet *et al.* 2008, 2014) on the 1% of the 10^6 simulated datasets closest to the observed dataset under the given scenario. The selected scenario was the one with the highest posterior probability value. The robustness and relevance of inferences were assessed with methods based on the analysis of pseudo-observed simulated data sets (Pods) (Bertorelle *et al.* 2010; Cornuet *et al.* 2010). The type I error rate (risk of excluding the focal scenario when it is true) and type II error rate (risk of selecting the focal scenario when it is false) were calculated to evaluate the robustness of the scenario choice. Posterior model checking was performed on the final scenario of each analysis to evaluate the goodness of fit between the inferred evolutionary history and the real data based on a second set of summary statistics, as suggested by Cornuet *et al*. (2010). Summary statistics were the following: one-sample summary statistics, the mean allelic size variance in the population and three two-sample summary statistics, the mean individual assignment likelihoods of population *i* assigned to population *j*, the mean expected heterozygosity pooling samples from two given populations, the shared allele distance between two given populations.

References:

Bertorelle G, Benazzo A, Mona S. (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Cornuet J-M, Pudlo P, Veyssier J, Dehne-garcia A, Marin J, Estoup A, Gautier M. (2014) DIYABC v2.0 : a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187–1189.

Cornuet J-M, Ravigné V, Estoup A. (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, **11**, 401.

Cornuet J-M, Santos F, Beaumont MA, Robert CP, Marin J, Balding DJ, Guillemaud T, Estoup A. (2008) Inferring population history with DIY ABC : a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Denton JS, Hitchings SP, Beebee TJC, Gent A. (1997) A recovery program for the natterjack toad (*Bufo calamita*) in Britain. *Conservation Biology*, **11**, 1329–1338.

Leskovar C, Oromi N, Sanuy D, Sinsch U. (2006) Demographic life history traits of reproductive natterjack toads (*Bufo calamita*) vary between northern and southern latitudes. *Amphibia-Reptilia*, **27**, 365–375.

May S, Zeisset I, Beebee TJC. (2011) Larval fitness and immunogenetic diversity in chytrid-infected and uninfected natterjack toad (*Bufo calamita*) populations. *Conservation Genetics*, **12**, 805–811.

Rowe G, Beebee TJC. (2004) Reconciling genetic and demographic estimators of effective population size in the anuran amphibian *Bufo Calamita*. *Conservation Genetics*, **5**, 287–298.

Stevens VM, Wesselingh RA, Baguette M. (2003) Demographic processe in a small, isolated population of Natterjack Toads (*Bufo calamita*) in southern Belgium. *Herpetological Journal*, **12**, 59–67

Table S3: Mean genetic differentiation estimated (i) from the whole data set of 62 *Bufo [Epidalea] calamita* populations (ALL POP), (ii) from a random subset of 22 populations (22 POP) and (iii) from populations only located within the focal study area in northern France (Focal study area). (A) Mean G_{ST} estimated on mitochondrial data, and mean N_{ST} accounting for the phylogenetic distances between mitochondrial haplotypes (Pons & Petit 1996). (B) Mean F_{ST} estimated on nuclear data and mean R_{ST} based on variance in allelic size (Slatkin 1995). N_{ST} and R_{ST} estimates were calculated and compared to G_{ST} and F_{ST} estimates obtained after 10,000 phylogenetic distances or allele size permutations respectively, using SPAGeDi version 1.4 (Hardy & Vekemans 2002).* $P < 0.05$; *** $P < 0.001$

(A)	ALL POP	22 POP	Focal study area
G_{ST}	0.487	0.564	0.403
N_{ST}	0.525	0.660	0.413
$N_{ST}>G_{ST}$	*NS*	*	*NS*

(B)	ALL POP	22 POP	Focal study area
F_{ST}	0.177	0.267	0.110
R_{ST}	0.184	0.532	0.109
$R_{ST} > F_{ST}$	*NS*	***	*NS*

Table S4: Detailed results of the approximate Bayesian computations analysis to trace back the evolutionary history of lineages of *Bufo [Epidalea] calamita* in northern France. See Table S2 for details on the test statistics. Probabilities are given with 95% confidence intervals in brackets.

		Prior evaluation of scenarios					Analysis of Pods		Model checking on selected scenario		
ABC analysis	Target genetic unit	Total number of test statistics	Proportion of test statistics with P < 0.05 or P > 0.95	Proportion of test statistics with P < 0.01 or P > 0.99	Scenario	Probabilities [95% CI]	Type I error	Type II error	Total number of test statistics	Proportion of test statistics with $P <$ 0.05 or $P >$ 0.95	Proportion of test statistics with $P <$ 0.01 or $P >$ 0.99
Set A: Origin of populations in the focal study area (northern France)	Northern coastline populations	15	0.467	0.2003	Western origin	0.9761 [0.967-0.985]	0.123	0.130	18	0.167	0.000
		15	0.533	0.133	Eastern origin	0.0174 [0.0096-0.094]	0.133	0.140	18	0.167	0.000
		15	0.533	0.133	Older lineage	0.0067 [0.003-0.010]	0.139	0.125	18	0.167	0.000
	Western coalfield populations	15	0.533	0.133	Western origin	0.9481 [0.936-0.960]	0.123	0.129	18	0.167	0.000
		15	0.533	0.133	Eastern origin	0.0291 [0.020-0.038]	0.121	0.121	18	0.167	0.000
		15	0.533	0.133	Older lineage	0.0227 [0.016-0.030]	0.133	0.140	18	0.056	0.111
	Eastern coalfield populations	15	0.533	0.133	Western origin	0.9542 [0.943-0.965]	0.122	0.130	18	0.056	0.111
		15	0.533	0.133	Eastern origin	0.0325 [0.023-0.042]	0.122	0.137	18	0.167	0.000
		15	0.533	0.133	Older lineage	0.0133 [0.009-0.018]	0.136	0.113	18	0.111	0.056
Set B: Relationship between focal study populations	All northern lineages	40	0.450	0.250	Coastal lineage is more ancient	0.9129 [0.892 0.934]	0.208	0.195	51	0.098	0.000
		40	0.450	0.250	Western coalfield lineage is more ancient	0.0382 [0.026 0.050]	0.183	0.189	51	0.059	0.039
		40	0.450	0.250	Eastern coalfield lineage is more ancient	0.0489 [0.034-0.064]	0.188	0.195	51	0.059	0.039

Table S5: Posterior distributions of parameters for scenarios with the highest posterior probability when comparing three conflicting scenarios of postglacial colonization of *Bufo [Epidalea] calamita* in northern France. Timing of events is in years, assuming a generation time of 3.5 years.

Parameters	Scenario A1 Northern coastal unit	Scenario A1 Northern western coalfield unit	Scenario A1 Northern eastern coalfield unit	Scenario B1 Northern lineages historical relationships
	Mean [Quantile 5%-95%]	Mean [Quantile 5%-95%]	Mean [Quantile 5%-95%]	Mean [Quantile 5%-95%]
N_1	5.17^4[1.85^4-8.81^4]	—	—	7.64^4 [5.21^4-9.51^4]
N_2	—	3.71^4 [1.05^4 – 7.84]	—	3.71^4 [1.19^4- 7.54^4]
N_3	—		4.31^4 [1.27^4 – 8.33^4]	4.32^4[1.40^3 - 8.24^4]
N_4	4.38^4 [1.35^4 - 8.35^4]	4.47^4 [1.44^4 -8.39^4]	4.70^4 [1.51^4 – 8.55^4]	6.11^4 [3.43^4 - 8.80^4]
N_5	6.99^4 [3.90^4 – 9.44^4]	7.41^4 [4.45^4 - 9.53^4]	7.47^4 [4.51^4 – 9.55^4]	8.20^4 [5.73^3 - 9.60^4]
N_{f1}	$4\ .56^3$ [2.44^4 - 9.11^4]	5.68^4 [1.22^4 - 9.51^5]	5.62^4 [1.24^4 - 9.47^4]	—
N_{f2}	—	—	—	6.80^4 [2.41^4 - 9.71^4]
N_{f3}	—	—	—	5.83^4 [1.28^4 - 9.57^4]
N_{f4}	—	—	—	5.29^4 [9.13^3 - 9.42^4]
N_a	4.59^4 [4.40^3-9.21^4]	4.72^4 [4.86^3 - 9.30^4]	4.89^4 [5.75^3 - 9.34^4]	4.80^4 [5.93^3 – 9.32^3]
t_1 (generations)	—	—	—	4.13^3 [3.71^2 - 1.13^4]
(time)				*14,455 years*
t_2 (generations)	—	—	—	1.94^4[3.26^3 - 3.66^4]
(time)				*52,150 years*
t_3 (generations)	2.03^4 [2.41^3-5.56^4]	2.63^4 [3.70^3 - 6.59^4]	2.58^4 [3.41^3 - 6.45^4]	3.95^4 [1.05^3 - 8.35^4]
(time)	*71,050 years*	*92,050 years*	*90,300 years*	*138,250 years*
t_4 (generations)	6.301^4 [1.50^4-1.41^5]	7.01^4 [1.88^4 - 1.51^5]	7.05^4 [1.83^4 - 1.51^5]	9.23^4 [3.39^4- 1.72^5]
(time)	*220,500 years*	*245,350 years*	*246 ,750 years*	*323,050 years*
μ	1.70^{-5} [3.61^{-6}-4.26^{-1}]	1.58^{-5} [3.39^{-6} - 3.91^{-5}]	1.49^{-5} [3.28^{-6} – 3.75^{-5}]	1.20^{-5} [2.94^{-5}- 2.82^{-5}]
P	1.43^{-1} [1.96^{-2}-2.75^{-1}]	1.29^{-1} [1.51^{-2} - 2.66^{-1}]	1.27^{-1} [1.46^{-2} - 2.62^{-1}]	1.53^{-1} [2.22^{-2} - 2.80^{-1}]
SNI	1.48^{-6} [1.32^{-7}-5.44^{-6}]	1.18^{-6} [9.70^{-8} - 4.63^{-6}]	1.24^{-7} [1.01^{-7} - 4.73^{-6}]	8.77^{-7} [7.92^{-8} - 3.44^{-6}]

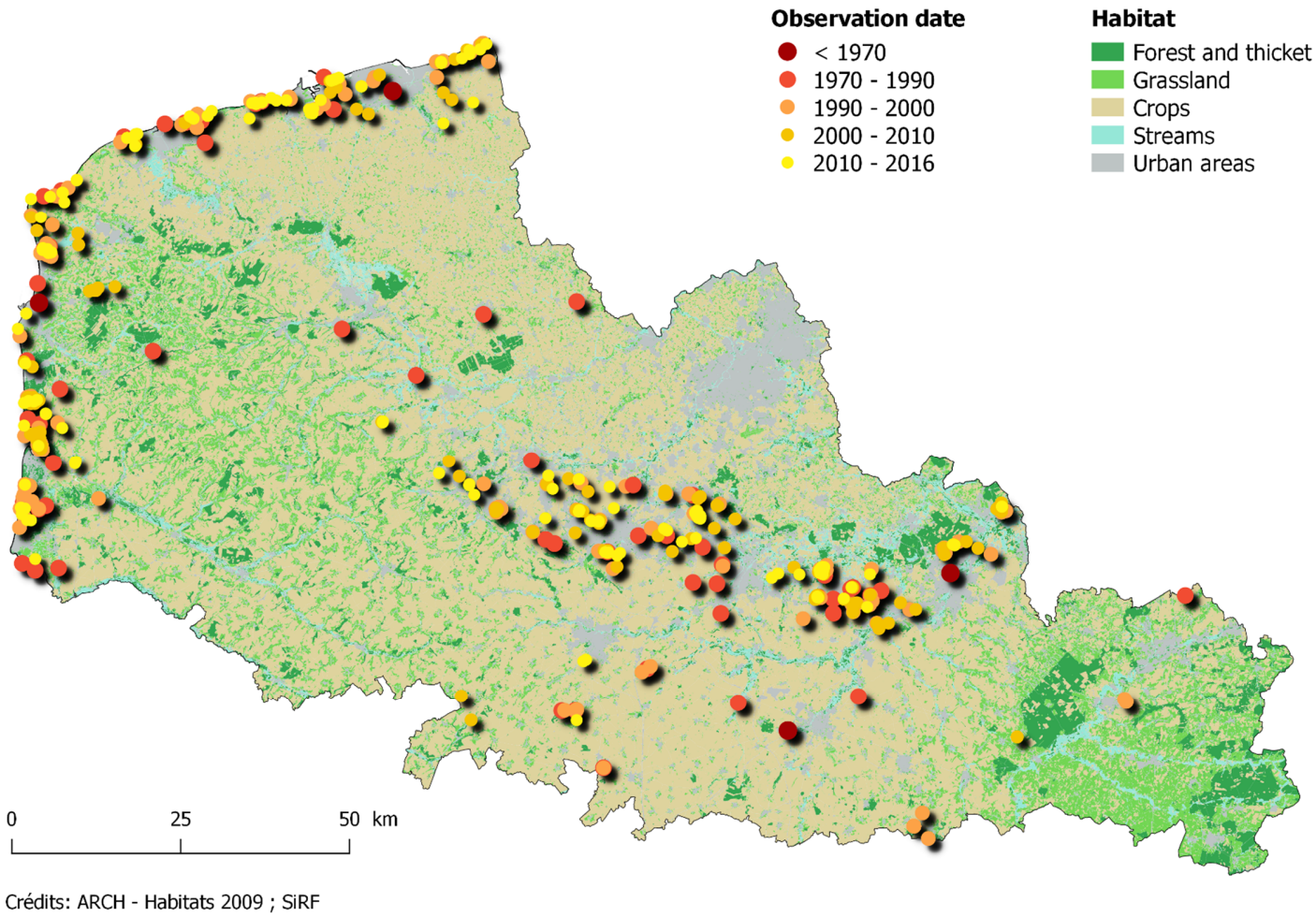


Figure S1: Geographical map featuring land use in northern France and localizing the historical occurrence of *Bufo* [*Epidalea*] *calamita* dating back to 1878. Sources : SiRF (Système d'information régional sur la faune) [http://www.sirf.eu/]. Groupe ornithologique et naturaliste du Nord – Pas de Calais (2017/05/03) for data observation and ARCH project for land use layer [open data].

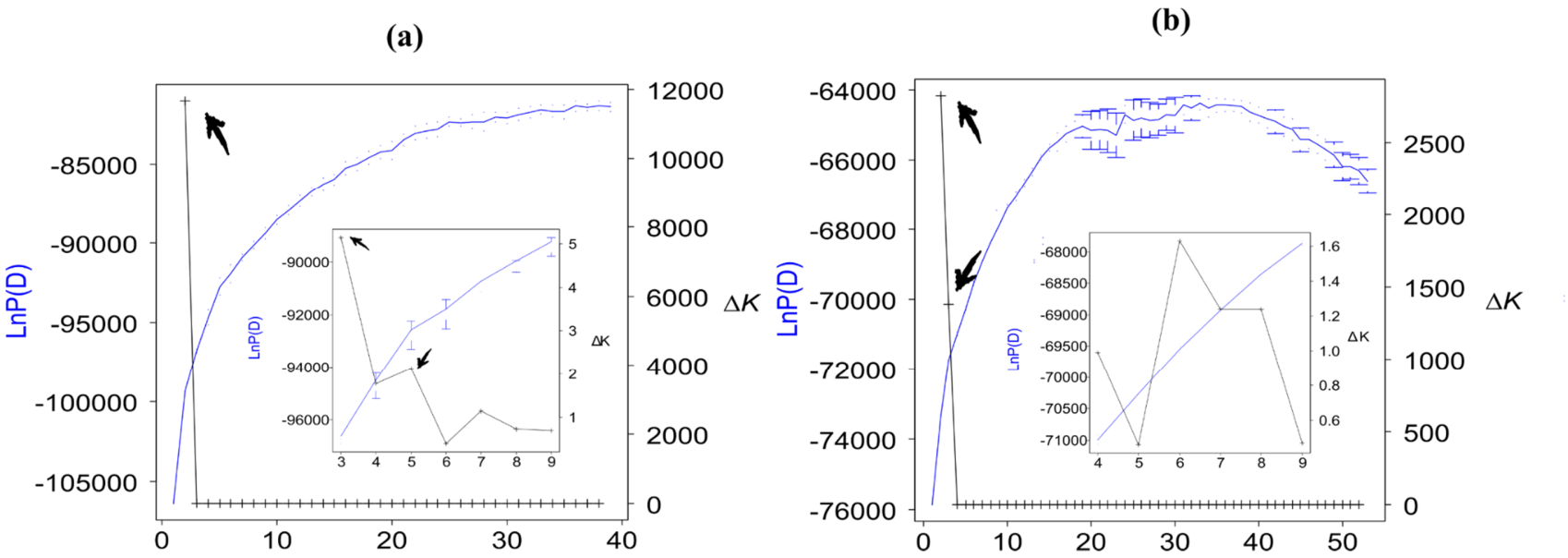


Figure S2: Ln(P(X|*K*)) and the ad hoc statistic Δ*K* (Evanno et al. 2005) obtained over 30 replicated runs and plotted against the putative number of *K* clusters using (A) the whole dataset of *Bufo* [*Epidalea*] *calamita* populations; (B) only the individuals located in the focal study area in northern France. The most likely number of K was detected by the highest variation of the log probability of data between two successive *K* values.

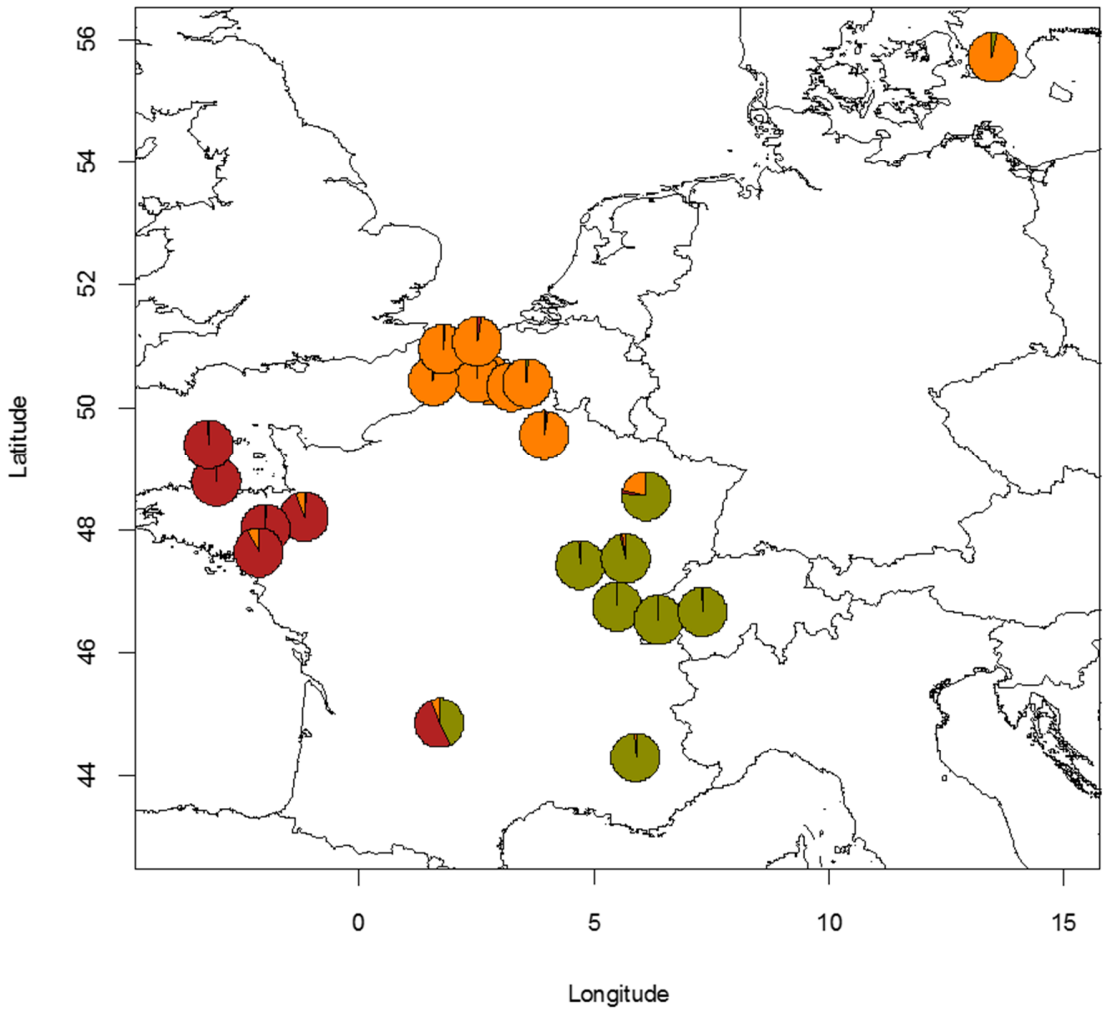


Figure S3: Bayesian clustering assignments performed, following Pritchard *et al.*, (2000), on a balanced sampling scheme of 22 *Bufo [Epidalea] calamita* populations based on 35 microsatellites. The diagrams on the map show the mean assignment probability of populations of belonging to one of the $K = 3$ genetic clusters detected (Pritchard *et al.* 2000)**.**

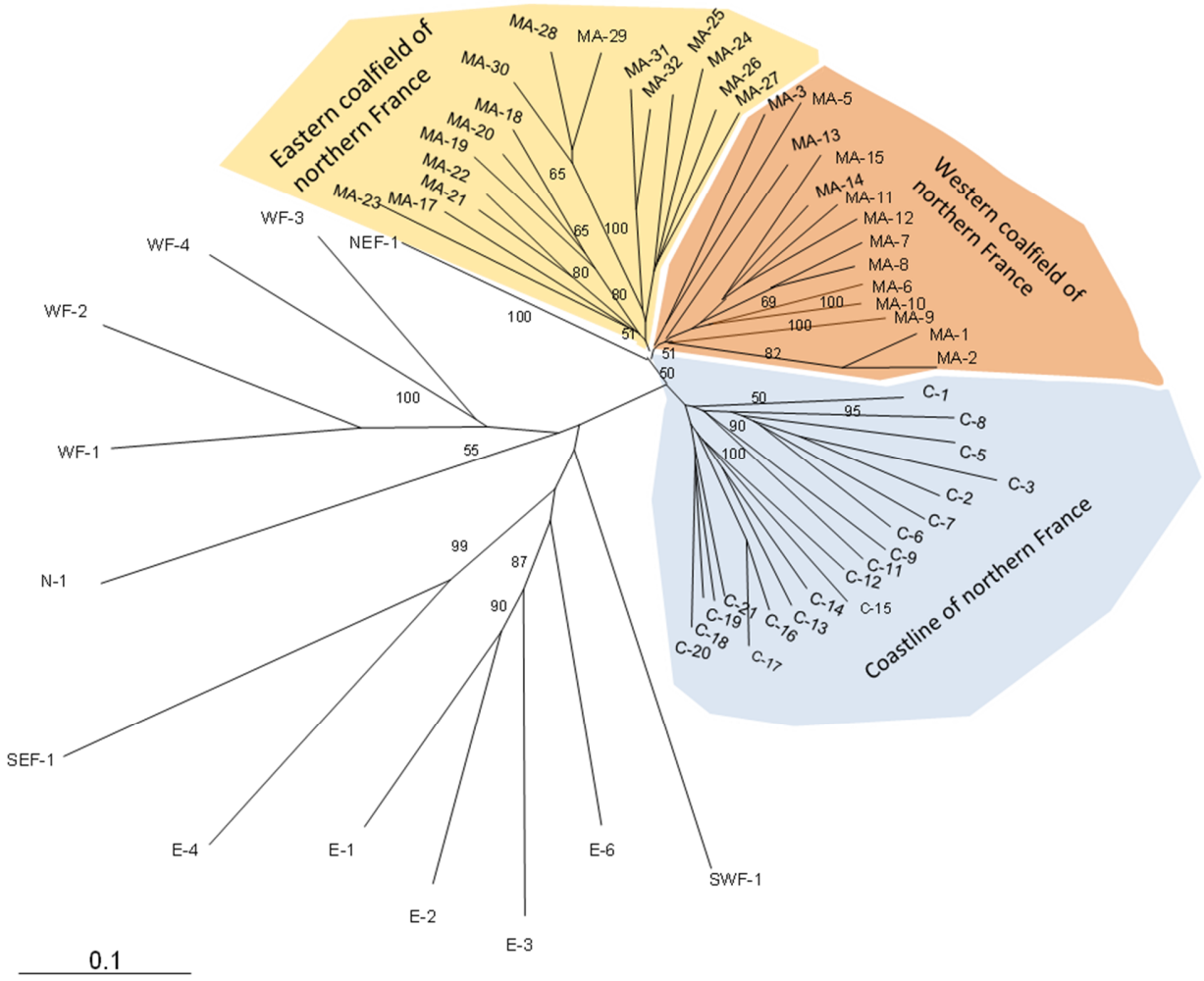


Figure S4: Neighbor-joining tree based on Cavalli-Sforza & Edwards (1967) chord distance (D_{CE}) on nuclear data. Only bootstraps >50% are indicated. Colored circles around populations indicated sampling areas of populations within the focal study area in northern France. Population codes are provided in Table S1.

(A)

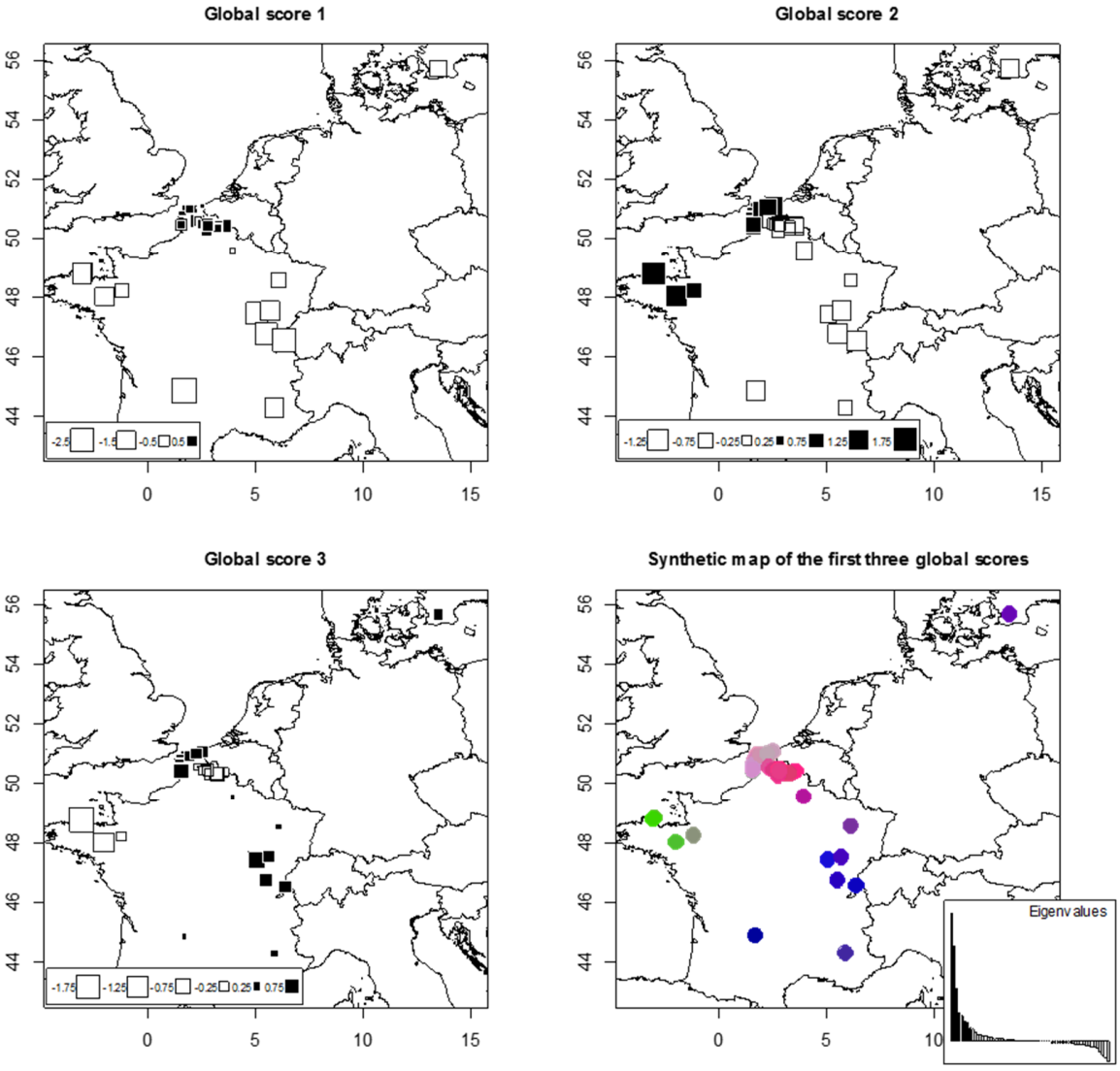
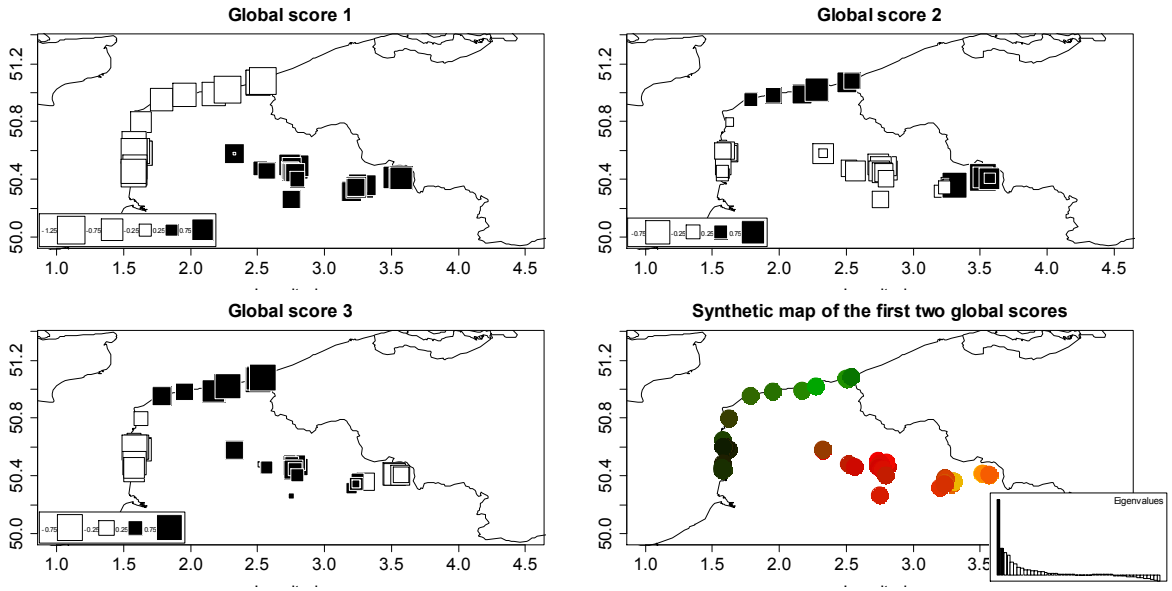


Figure S5: Maps of the first three sPCA global scores observed on (A) the whole data set of *Bufo* [*Epidalea*] *calamita* and (B) populations located in the focal studied area of northern populations. Each square represents a population. Populations that are more closely related in the multivariate space share the same color of squares. Sizes of squares indicate the magnitude of spatial positive autocorrelation. In the synthetic maps, each population was mapped by color coding the three sPCA lagged scores folowing a color channel as in Menozzi *et al.* (1978). A minimum spanning tree connected graph was used as connection network to perform sPCA.

(B)



As a complementary analysis free of any genetic assumptions, a spatial principal component analysis (sPCA, reviewed in Jombart et al. 2009) was performed to investigated population genetic structure. sPCA relies on a modification of PCA such that not only the variance of principal components, but also their spatial autocorrelation (based on Moran's I, see Sokal & Oden 1978), are optimized. A minimum spanning tree connected graph fitted the most with genetic data (data not shown) and was therefore used as connection network to perform sPCA. Genetic variability can stem from (i) global structure, displaying positive spatial autocorrelation and depicting IBD or related cline processes, or (ii) local structure, displaying negative spatial autocorrelation and reflecting stronger genetic differences among nearest neighbors (Jombart *et al.* 2008). sPCA and tests for statistical significance of global and local structures were performed using the R ADEGENET package (Jombart 2008). Synthetic maps of population scores were drawn to facilitate visual assessment of spatial genetic structure (Menozzi *et al.* 1978). As above, analyses were conducted on the whole data set of populations and only on populations belonging to the focal area in northern France.

References:

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

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

Jombart T, Pontier D, Dufour A-B. (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**, 330–341.

Menozzi P, Piazza A, Cavalli-Sforza LL. (1978) Synthetic maps of human gene frequencies in Europeans. *Science*, **201**, 786–792.

Sokal RR, Oden NL. (1978) Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean Society*, **10**, 199–228.

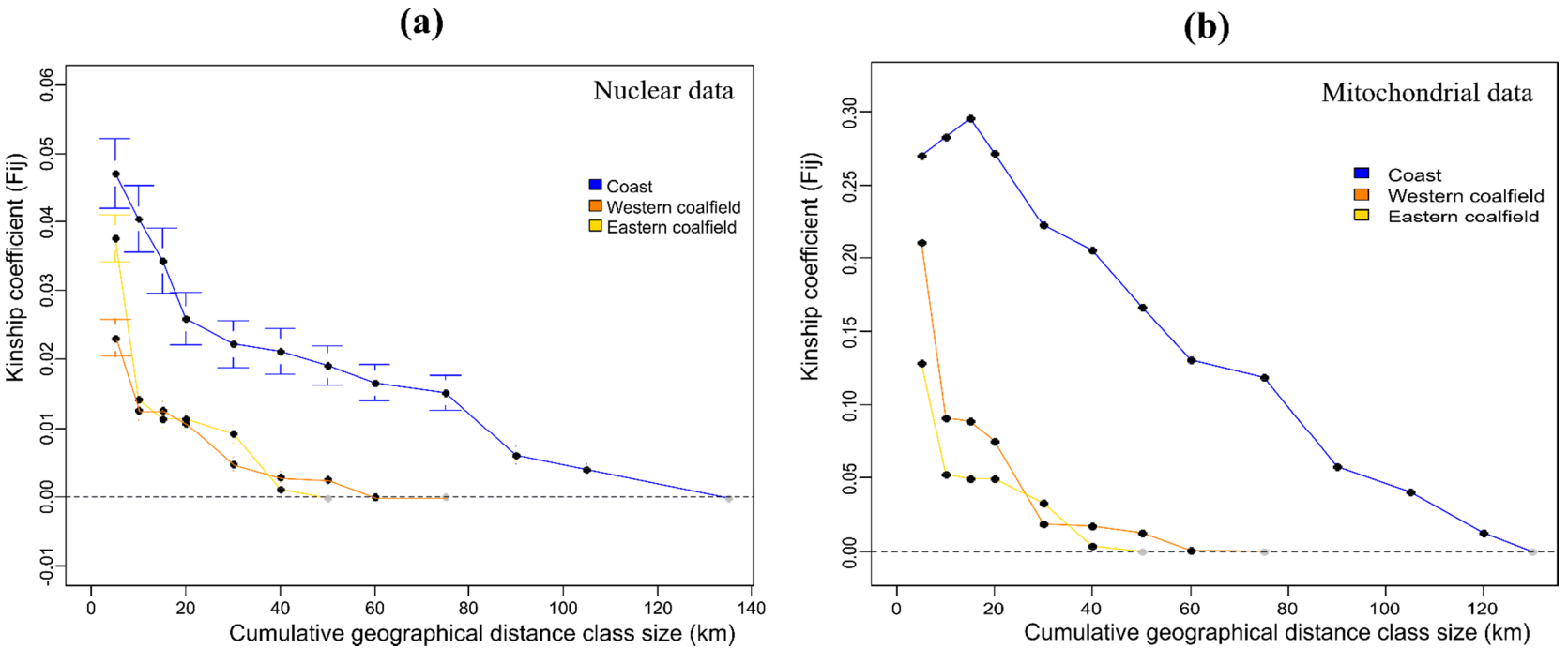


Figure S6: Spatial genetic structure in *Bufo [Epidalea] calamita* populations located within the focal study area in northern France and defined as belonging to three distinct genetic clusters: coastline populations, western coalfield populations and eastern coalfield populations. Average pairwise kinship coefficients (F_{ij}, Loiselle *et al.* 1995) among individuals are plotted against cumulative geographical distance classes (see Peakall *et al.* 2003) for nuclear microsatellites (A) and mitochondrial haplotypes (B). Standard error of nuclear F_{ij} were estimated using jackknifing procedure over loci. Statistical significance of estimated coefficients was tested by 1 000 permutations of individuals among localities (● $P < 0.05$, ● $P > 0.05$).

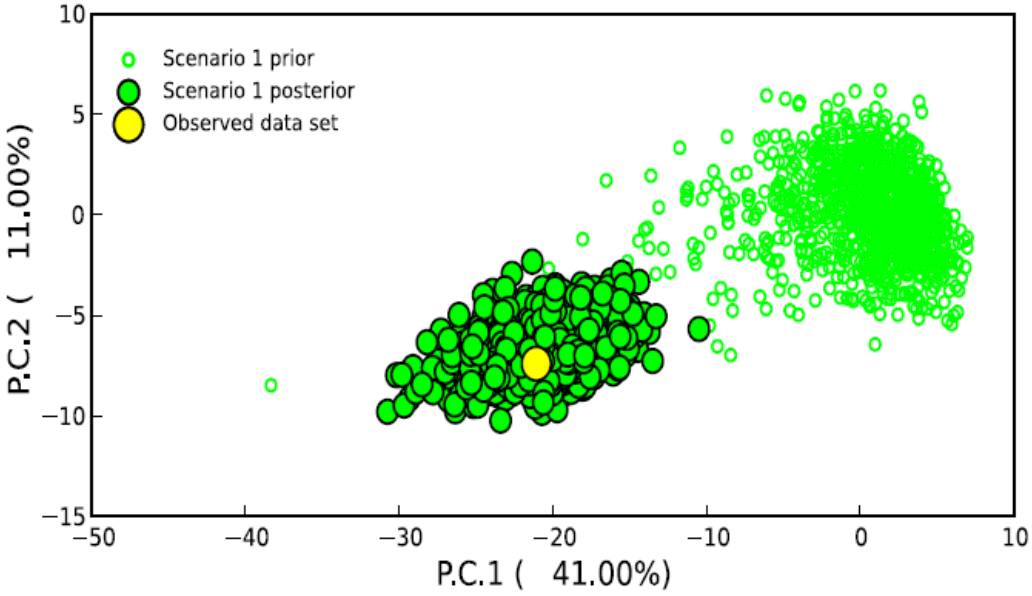


Figure S7: Principal component analysis (PCA) of test quantities when processing model checking for the most likely scenario B1 detailed in Figure 2. The pseudo-observed test data set was simulated under scenario B1. PCA was processed on the test quantities corresponding to summary statistics detailed in the legend of Table S2 and not used to discriminate among scenarios and compute the posterior distributions of parameters given in Table S5

CHAPITRE III

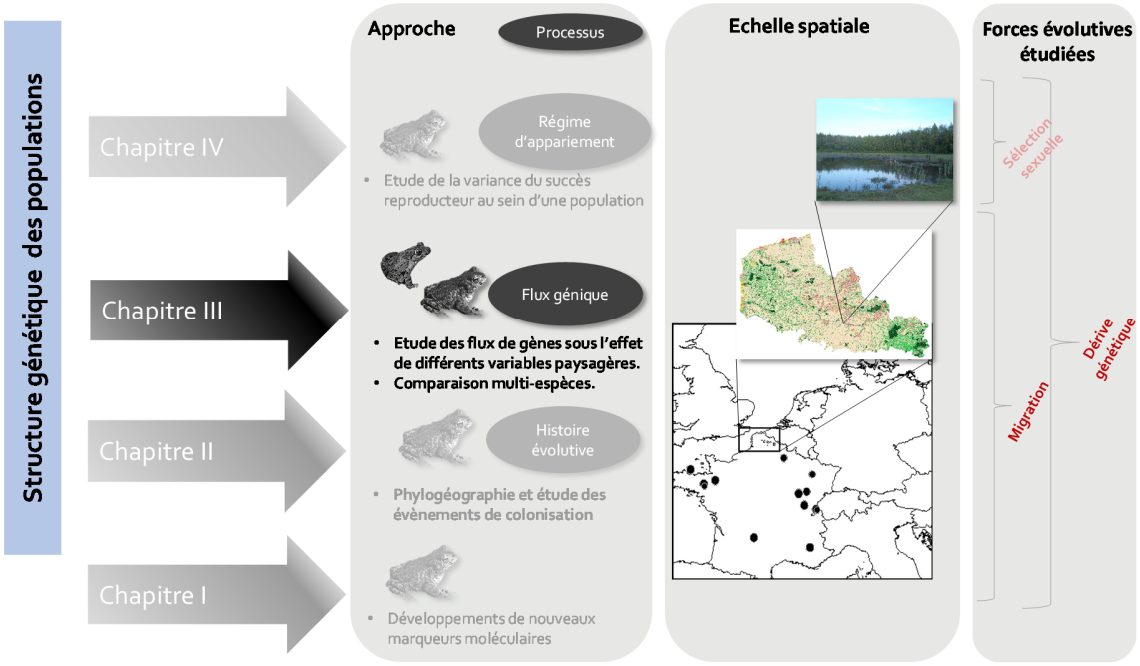
Impact du paysage et des tailles efficaces sur les patrons de flux géniques chez le Pélodyte ponctué et le Crapaud calamite établis dans des zones fortement anthropisées

© Guillaume Lemoine

Photographie de Pélodytes ponctués sur le flanc d'un terri

La connectivité du paysage est un des paramètres clés modulant la dynamique des populations. Dans ce chapitre, nous avons entrepris d'évaluer l'impact de la connectivité du paysage pour deux espèces d'amphibiens pionnières, le Crapaud calamite et le Pélodyte ponctué. Ces deux espèces occupent des habitats similaires sur le littoral et dans le bassin minier du nord de la France, mais présentent des traits d'histoire de vie, et notamment des capacités de dispersion différentes. Le Crapaud calamite forme d'assez grandes populations sur l'ensemble du nord de la France tandis que le Pélodyte ponctué présente des effectifs beaucoup plus restreints, ceci pouvant s'expliquer par le fait qu'il soit en limite nord d'aire de répartition géographique. L'estimation des flux géniques à partir de la divergence génétique des populations, ou à partir des niveaux d'apparentement entre individus, nécessite de prendre également en compte les effets de la dérive génétique. Des paires de populations échangeant des flux géniques à la même intensité ne présenteront pas la même divergence génétique selon qu'elles aient de grandes tailles efficaces ou de petites tailles efficaces de population. Plus la population est d'effectif réduit, plus la dérive génétique entrainera une perte aléatoire importante de diversité génétique induisant une plus forte différenciation génétique entre populations.

Notre approche a ainsi consisté à évaluer l'effet de la matrice paysagère sur les flux de gènes en étudiant la relation entre les discontinuités génétiques observées entre populations et une série de distances écologiques mesurées à partir de différents modèles de paysages, tout en prenant en compte de potentiels effets d'asymétrie dans les tailles efficaces.



L'objectif général de ce chapitre de thèse était de déterminer si les patrons de structure génétique spatiale des populations de Crapauds calamites observée dans le chapitre II pouvaient être expliqués par la composition et la structure du paysage. Autrement dit, nous voulions tester s'il existait une moins bonne connectivité paysagère liée à l'existence de barrières aux flux de gènes ou à la perte de corridors au sein du bassin minier, zone fortement anthropisée.

Plusieurs travaux s'intéressant à la dispersion du Crapaud calamite en fonction de différentes variables paysagères ont déjà été menés et nous ont permis de faire plusieurs hypothèses sur l'effet du paysage sur les flux de gènes chez cette espèce (Stevens *et al.* 2004, 2006a,b; Flavenot *et al.* 2015). Ainsi, nous avons prédit une certaine perméabilité de variables paysagères suivant un gradient de couverture végétale et un gradient d'intensité de pressions anthropogéniques pour chacune des espèces, sachant qu'elles partagent une niche écologique similaire. Les hypothèses reposaient sur un effet barrière des zones urbaines et des terres agricoles, et un effet corridor des zones enherbés ouvertes comme les prairies et les bords de routes, ceci pour chacune des deux espèces. Les effets des variables paysagères ont été modélisés en leur attribuant des valeurs de résistances reflétant l'impact qu'on leur attribue *a priori* (*i.e.* augmente ou diminue les probabilités de dispersion des individus au sein de la matrice paysagère, cf Figure III.1).

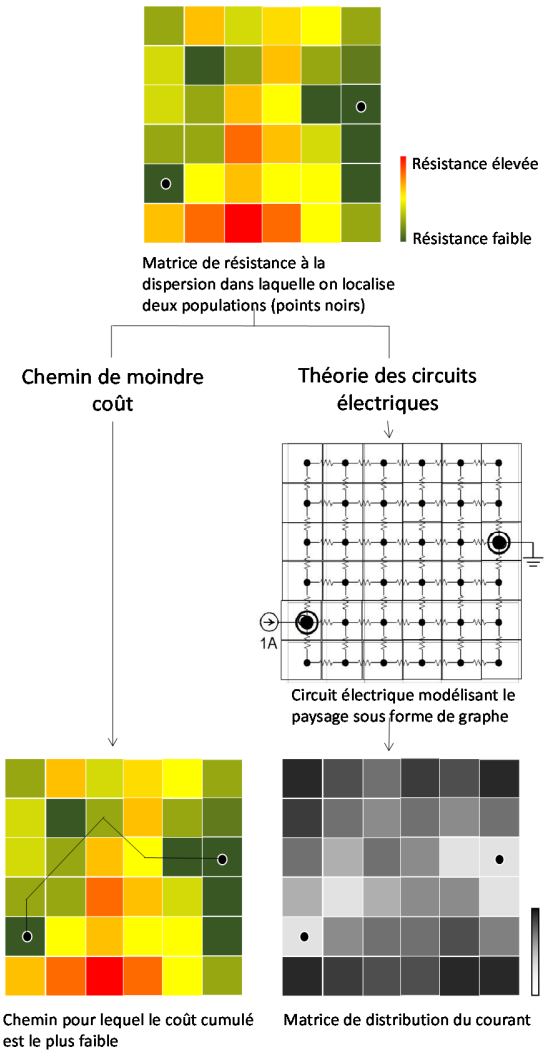


Figure III 1: Schéma illustrant la démarche de modélisation du paysage et d'estimation des distances écologique selon la méthode du chemin de moindre coût et selon la méthode fondée sur la théorie des circuits électriques

Le Pélodyte ponctué étant situé en limite d'aire de répartition géographique, il était attendu que les niveaux de différenciation génétique soient plus élevés entre les populations de cette espèce qu'entre populations de Crapaud calamite (Bertorelle *et al.* 2009; Allendorf *et al.* 2012; Frankham *et al.* 2014).

De même, la distribution du Pélodyte ponctué étant plus restreinte dans le bassin minier, nous avons fait l'hypothèse que la connectivité paysagère soit moins bonne encore pour cette espèce et se répercute en des flux de gènes plus restreints spatialement, induisant une structure génétique spatial à plus courte distance géographique.

Au final, en plus des 53 populations de Crapaud calamites, 267 Pélodytes ponctués répartis sur 28 sites ont donc été échantillonnés pour cette approche comparative, et génotypés sur 22 marqueurs microsatellites. Les distances écologiques ont été mesurées sur différents modèles de paysages hypothétiques et de différentes façon*, i.e.* fondées sur la distance géographique euclidienne qui ne suppose pas d'effet distinct de chacune des variables du paysage, fondées au contraire sur le calcul de chemins de moindre coût, ou enfin, fondées sur la théorie des réseaux de circuits électriques qui mesure les distances en prenant en compte les résistances spécifiques de chacune des variables environnementales (Figure III 1)

Cette confrontation des données de génétique des populations et des mesures de distances écologiques nous a ainsi permis d'identifier les seuils de connectivité moyen des populations en deçà desquels les événements de migration deviennent trop rares pour contrer la dérive génétique. Sur le littoral nos résultats indiquent que les populations se structurent selon un schéma classique d'isolement reposant sur la distance Euclidienne. En revanche, dans le bassin minier, les résultats indiquent un effet du paysage significatif sur les flux de gènes. Comme pour le Crapaud calamite, dans le bassin minier les flux de gènes opèrent sur une échelle spatiale inférieure à 5kms pour le Pélodyte ponctué. En prenant en compte les effets des différentes variables paysagères, la connectivité des populations est fortement diminuée au regard de la simple prise en compte de la distance géographique Euclidienne. De plus, les populations de Pélodyte ponctué montrent des niveaux d'apparentement génétique entre individus et des niveaux de différenciation inter-populationnelles plus élevés que chez le Crapaud calamite. Ces résultats mettent en évidence une connectivité très restreinte entre populations du bassin minier dont certaines se retrouvent totalement isolées des autres. Cette étude comparative met en évidence les enjeux de la connectivité paysagère dans les zones post-industrielles fortement perturbées comme le bassin minier. Dans ce paysage, peu d'habitats réellement favorables sont disponibles et les populations sont réellement tributaires d'espaces ouverts fournis par les friches industrielles des sites miniers. Cette étude souligne l'importance du maintien des sites miniers comme unité de conservation et la nécessité de gestion du paysage en vue d'améliorer la connectivité entre populations, notamment en ce qui concerne quelques îlots de

populations qui semblent complètement déconnectées d'autres populations pourtant spatialement très proches.

Références

Allendorf FW, Luikart GH, Aitken SN (2012) *Conservation and the Genetics of Populations*. Wiley-Blackwell, Oxford (UK).

Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C (2009) *Population Genetics for Animal Conservation*.

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

Frankham R, Bradshaw CJ a., Brook BW (2014) Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, **170**, 56–63.

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, Leboulengé 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.

IMPACT OF LANDSCAPE FEATURES AND EFFECTIVE POPULATION SIZE ON PATTERNS OF GENE FLOW IN POND-BREEDING AMPHIBIAN SPECIES ESTABLISHED IN ANTHROPOGENICALLY-IMPACTED AREAS

Leslie Faucher[1], Anne Duputié[1], Stéphanie Rondel[3], Cédric Vanappelghem[1,2], José Godin[1,4], Clément Mazoyer[1], Cécile Godé[1], Laura Hénocq[1], Thomas Brom[1], Julie Jaquiéry[5] & Jean-François Arnaud[1]

[1] Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France
[2] Conservatoire d'espaces naturels du Nord et du Pas de Calais, Lillers, France
[3] Centre Permanent d'Initiatives pour l'Environnement - Chaîne des Terrils, Loos-en-Gohelle, France
[4] Groupe ornithologique et naturaliste du Nord – Pas-de-Calais, Lille, France
[5] Université de Rennes 1, UMR CNRS 6553 – ECOBIO, 35042 Rennes Cedex, France

Keywords: amphibians, gene flow, landscape connectivity, spatial genetic structure, *Pelodytes punctatus, Bufo [Epidalea] calamita*, spoil heaps

Abstract

Gene flow is a key evolutionary process that influences the level of intraspecific genetic diversity necessary for adaptive evolution in changing environments. Loss of habitat connectivity and low effective population size can exacerbate the loss of genetic diversity in isolated fragments and drive population extinction, especially for remnant populations localized in fragmented urban and industrial areas. To understand the influence of human-modified land cover on patterns of amphibian gene flow, we carried out a comparative multispecies survey of genetic structure within semi-natural, urban and post-industrial environments in northern France. To this end, we investigated whether landscape features impacted the genetic structure of two pioneering pond-breeding amphibians, the parsley frog (*Pelodytes punctatus*) and the natterjack toad (*Bufo* [*Epidalea*] *calamita*). We designed *a priori* landscape models to test for the effect of landscape features and local N_e on genetic divergence using Mantel tests and generalized mixed effect models. We expected (i) higher levels of genetic divergence in the parsley frog because the study area constitutes its leading edge distribution range and (ii) reduced negative landscape effects to dispersal for coastal areas which are less impacted by human-induced fragmentation than the coal basin. We depicted marked local genetic discontinuities across neighboring genetic lineages within urbanized and reclaimed coalfield areas. We found an increased effects of genetic drift for smallest populations and a greater population genetic divergence for the parsley frog, suggesting drift effects enhanced by leading-edge expansions and low dispersal capabilities. Whereas Euclidean distance best explained the genetic connectivity for semi-natural areas located along the coastline, open fields and grasslands may facilitate gene flow in urban and post-mining areas for which even nearest-neighbor populations appeared to be genetically disconnected. Surprisingly, roads also appeared to act as efficient corridors in parts of the studied areas. Our findings provide recommendations for conservation and management of very fine-scaled disconnected amphibian metapopulations.

Introduction

Nowadays human activities have an unprecedented impact on ecosystems (Foley *et al.* 2005). Native habitat areas drastically decrease in a fragmented matrix of anthropogenic environments, often resulting in decreasing population size and reduced dispersal among populations. Whereas fragmentation effects, independently of habitat loss, often imply significant positive effects (Fahrig 2017), a lack of connectivity among populations have long-standing demographic and genetic implications. This includes inbreeding depression, compromised immune response, low population size resulting in vulnerability to face environmental and demographic stochasticity and an erosion of the adaptive potential (Fahrig 2003; Mimura *et al.* 2017). Indeed, reduced levels of gene flow among populations imply a loss of gene diversity which is acknowledged as a major driver of the current mass extinction crisis (Keller & Waller 2002; Allentoft & O'Brien 2010; Miraldo *et al.* 2016). The levels of intraspecific genetic diversity, that is the amount of genetic variation among individuals, have key consequences for ecosystem functioning, the composition of communities, the viability of spatially-structured populations, the individual fitness and, as such, provide the critical basis for adaptive evolution in changing environments (Banks *et al.* 2013; Miraldo *et al.* 2016; Mimura *et al.* 2017). In this context, population viability greatly depends of individual dispersal capabilities and of the arrangement of suitable landscape components allowing efficient gene flow (Baguette *et al.* 2013). Dispersal insures recruitment and gene flow among populations and may counterbalance a population size decrease by stochastic demographic events and a loss of genetic diversity by genetic drift effects. Overall, life-history traits, species habitat requirements and landscape features have a major impact on dispersal, specific landscape features either impeding or promoting individual dispersal (e.g. Row *et al.* 2010; Gubili *et al.* 2017).

By combining landscape ecology, population genetics and spatial statistics, it is possible to identify broad or species-specific barriers and corridors which is the focus of landscape management in conservation plans and helps managers making appropriate conservation decisions regarding monitoring programs (e.g. Michels *et al.* 2001; Weckworth *et al.* 2013; Frei *et al.* 2016; Richardson *et al.* 2016; Gubili *et al.* 2017). Rooted in a classical evolutionary and population genetics framework (see Dyer 2015), this approach proved to be successful for a wide range of animal and plant taxa, including both above- or belowground terrestrial organisms and marine species as well (e.g. Villemey *et al.* 2016; Buonomo *et al.* 2017; Dupont *et al.* 2017). However,

gene flow is not the only parameter influencing population genetic differentiation. Indeed, the effective population size (N_e), *i.e.* the number of adults in an idealized population under the Wright-Fisher model, is an additional critical key parameter that determines the rate of loss of gene diversity induced by genetic drift (Ellegren & Galtier 2016). Hence, it appears important to consider both N_e and landscape to better estimate the amount of genetic differentiation and the loss of genetic diversity directly linked to a lack of landscape connectivity from those induced by decreasing N_e (Serrouya *et al.* 2012; Weckworth *et al.* 2013; reviewed in Lowe *et al.* 2017).

Pond-breeding amphibians are ideal models for investigating the impact of habitat use and land cover features on patterns of gene flow and levels of N_e within a mosaic of human-modified landscapes. Indeed, they have moderate to low dispersal capabilities, strong breeding site fidelity, small N_e and discontinuous distribution of their suitable aquatic breeding habitats (Beebee & Griffiths 2005; Stevens *et al.* 2006a, b; Phillipsen *et al.* 2011; McCartney-Melstad & Shaffer 2015). Furthermore, they share various aquatic and terrestrial habitats that must be connected for seasonal migration between breeding and overwinter habitats. Many amphibian species occur in temporary ponds, inducing patchy breeding assemblage, with high rates of extinction and colonization that necessitate connection among breeding habitats to maintain regional persistence (Cushman 2006; Janin *et al.* 2009; Goldberg & Waits 2010; Cosentino *et al.* 2012). Finally, amphibian species are among the most threatened taxa (Stuart *et al.* 2004; Allentoft & O'Brien 2010). They suffer from a world-wide loss of wetlands and from habitat fragmentation, especially in western Europe which is characterized by the lowest levels of genetic diversity for amphibian species, notably in anthropogenic ecosystems (Cushman 2006; Miraldo *et al.* 2016).

This study focused on a strongly fragmented post-industrial area, inherited from three centuries of coal extraction and located in northern France, Western Europe. Coal activities led to the formation of hundreds of slag heaps along with an extensive urban development. Today, this mining area consists of a severely fragmented landscape dominated by urban habitats and farming areas and crossed by several chains of slag heaps constituting new habitats of substitution for pioneering amphibian species (Godin 2002; Faucher *et al.* 2017). In this study area, we investigated the effect of land cover on population genetic divergence of pioneering amphibian species by taking into account the conjoint effect of N_e. Comparative analysis of spatial genetic structure using a multispecies approach on the same landscape is required to expand the scope of inference and to provide more generalized management decisions and cost-effective conservation measures

(Richardson *et al.* 2016; Favre-Bac *et al.* 2016; Gutiérrez-Rodríguez *et al.* 2017). To this end, we investigated whether landscape features impact the genetic structure of two pioneering pond-breeding amphibians, the common parsley frog (*Pelodytes punctatus*) and the natterjack toad (*Bufo* [*Epidalea*] *calamita,* Laurenti 1768). Both species inhabit pioneering habitats, breed in ephemeral ponds but likely differ in their dispersal and colonization abilities, although life-history traits related to dispersal remain little studied in the parsley frog. Reclaimed mine environments of northern France have been colonized by both species from previous remnant semi-natural inland habitats (Godin 2002). In the natterjack toad, we recently showed that these former mining areas offset the ongoing loss of inland natural habitats and host highly genetically diverse populations (Faucher *et al.* 2017). Nonetheless, despite high levels of genetic diversity, dispersal events seem to occur only over very short spatial scales (< 10 km) in mining areas, as compared to more natural habitats located along the coastline (Faucher *et al.* 2017). This points out the need to characterize the landscape components impeding gene flow exchanges among these newly established populations. Besides, this industrial area constitutes the northern leading edge of the parsley frog species' distribution. This also points the need for a thorough investigation of connecting pathway features driving the patterns of gene flow events because leading edge populations are generally thought to be genetically depauperate and of small N_e (Allendorf *et al.* 2012; Frankham *et al.* 2014).

To understand the influence of human-modified land cover on patterns of amphibian gene flow and drivers of population connectivity, we thus carried out a comparative survey of genetic divergence within semi-natural, urban and post-industrial environments of northern France. To this end, we compared the levels of genetic differentiation at both population and individual levels. We used *a priori* landscape models to test for the effects of landscape features and local N_e on the genetic divergence against the null hypothesis of isolation by distance (IBD, Wright 1943) where the Euclidean geographical distance accounts for the genetic divergence observed. To investigate whether local genetic discontinuities can be attributed to landscape effects, more ecologically realistic geographical distances, relying on animal landscape perception during dispersal, were additionally estimated. First, we estimated classical least cost paths (LCP) that assume that organisms disperse following a single and optimal route. However, LCP measures cannot consider multiple optimal dispersal pathways across the landscape, which often involves historical gene flow spanning several generations (Slatkin 1985). Consequently, we also used a model borrowing to electrical circuit theory, that evaluates the total landscape resistance between sampling sites

(McRae 2006). This model proved to allow correct inferences on patterns of gene flow because it accounts for multiple pathways and shape and width of habitat swaths connecting structured populations (e.g. McRae & Beier 2007; Row *et al.* 2010). Next, to assess the most likely landscape model, we evaluated the relationships between pairwise genetic distances, Euclidean geographical distances, and each of the more realistic modified ecological distances using Mantel and partial Mantel tests by taking into account the effect of local N_e. We also tested for the effects of each landscape feature on patterns of gene flow using generalized mixed effect models, as described in Cleary *et al.* (2017). Our expectations were the following: (1) higher levels of genetic divergence among populations in the parsley frog because this species occurs in a cryptic way in areas of northern France which constitute the leading edge of its distribution range, (2) reduced effects of barriers to gene flow for the natterjack toad which is thought to be a good disperser compared to the parsley frog (Sinsch 1997; Stevens *et al.* 2004; Frei *et al.* 2016) (3) a less stringent effect of landscape barriers to dispersal within the semi-natural habitats located along the coastline because this area is less impacted by human-induced fragmentation, (4) contrasting effects of forests and crops on levels of gene flow, with crops acting as a barrier to dispersal and forests not impeding among-population connectivity for the natterjack toad (Stevens *et al*. 2006a,b).

Material & Methods

Study species

The natterjack toad and the common parsley frog are European anurans mostly breeding in temporary ponds and occurring in early successional open habitats like dunes, marshes or human-made habitats such as cultivated areas, flooded quarries or slag heaps (Beebee & Denton 1996; Denton *et al.* 1997; Wilkinson & Griffiths 2013; Arntzen *et al.* 2017). Both species are protected by national and sub-national legislation throughout much of their range (IUCN 2016). In northern France, the natterjack toad is commonly distributed along the coastline and in human-modified inland habitats (Godin 2002; Arntzen *et al.* 2017). The parsley frog shares the same habitats, but is rarer because northern France constitutes the leading edge of the species geographical distribution (Godin 2002).

With respect to life-history traits related to dispersal capabilities, the natterjack toad is known to cover substantial dispersal distances, as much as several kilometers, with a dispersal capacity of more than 2.5 km (Sinsch 1997). In terms of dispersal behavior, adult natterjacks show a strong fidelity to breeding areas. Toadlets are known to be the dispersal stage of the species and

the movement ability of juvenile natterjack toads is affected by landscape components (Stevens *et al.* 2004; Frei *et al.* 2016). Puzzling results were found on the effect of landscape matrix on dispersal ability in toadlets. In contrast to sandy soil and roads, areas such as forests or agricultural fields decrease the movement efficiency. However, by using acoustic and olfactory cues, natterjack toadlets are able to actively select different environments and, in contrast to expectations, exhibited significant preference for forested environments despite their high resistance to movements (Stevens *et al.* 2006a). Overall, the avoidance of agricultural habitats and a preference towards forest elements may drive the patterns of gene flow in the natterjack toad but this needs to be thoroughly investigated over a large landscape (Stevens *et al.* 2006b). Very few data are available on the parsley frog dispersal capabilities. This species is known to display high phenotypic plasticity allowing successful exploitation of diverse temporal niches in unpredictable environments (Jourdan-Pineau *et al.* 2012). However, studies on breeding site fidelity and mean dispersal range remain to be carried out.

Study area

Northern France is a broad open plain with a highly fragmented landscape dominated by arable lands and urbanized areas (Figure 1). Urban areas cover more than 16% of the area and agricultural land cover more than 70% of the area (20% grasslands and pasture, 80% intensive cropping). Conversely, forests, semi-natural habitats and wetlands represent only 10% of the landscape. Most of the early successional natural habitats such as dunes or heaths are located along the shoreline. In inland areas, the landscape was profoundly modified since the end of World War 2. Most remnant semi-natural habitats occur within a matrix of agricultural lands or in abandoned coalfield areas where mining activities had shaped a large number of slag heaps since the last three centuries (Godin 2002; Lemoine 2012). Nowadays, these reclaimed mine environments constitute artificial open habitats that shelter ponds with rich species diversity (Lemoine 2012). However, the patchy distribution of the breeding ponds may impede the extent of gene exchanges and questions the long-term survival of populations (Arntzen *et al.* 2017; Faucher *et al.* 2017). Therefore, this area provides a unique opportunity to study the impact of landscape heterogeneity on patterns of gene flow within patchy metapopulations of pioneering amphibian species.

Sampling procedure, DNA extraction and DNA genotyping

53 natterjack populations were sampled in the study area, totalizing 959 natterjack toads (mean=18 ±9 individuals per population, see Figure 1 and Table S2). Part of this dataset was analyzed for colonization history of coalfields in a wider biogeographical context (Faucher et al. 2017). We further collected 28 populations of the parsley frog, totalizing 267 individuals with a mean of 9 ±8 individuals per population (Figure 1, Table S2). Individuals were caught by night during the reproductive period (from March to June) and DNA was non-invasively sampled using plain sterile 15SC Copan (Brescia Italia) swabs as described in Faucher et al. (2016). Individuals were immediately released after buccal swabbing. Each locality was sampled once in 2013 or 2014.

Detailed information on genotyping can be found in Faucher *et al.* (2016). Briefly, whole genomic DNA was extracted from swabs using Macherey-Nagel (Düren, Germany) NucleoSpin® 96 trace kits following the standard protocol outlined in the manufacturer's handbook. To assess nuclear DNA polymorphism in the natterjack toad, we used 37 nuclear microsatellite loci described in Rowe *et al.* (1997, 2000), Rogell *et al.* (2005) and Faucher *et al.* (2016). In the same way, 22 nuclear microsatellite loci (Jourdan-Pineau *et al.* 2009; Van De Vliet *et al.* 2009) were used to characterize nuclear DNA diversity in the parsley frog (see Table S1). PCR products were analyzed using an ABI Prism 3730xl Analyzer (Applied Biosystems, Foster City, CA) and multilocus genotypes were manually scored using GENEMAPPER 3.7 software. Individuals showing dubious genotypes underwent a second round of PCR. Overall missing data rate was 1.14% for natterjack toad and 5.00% in the parsley frog.

Analyses of genetic data

Basic statistics

Linkage disequilibrium (LD) among pairs of loci and across populations, and departures from Hardy–Weinberg (HW) equilibrium for each locus were tested using, respectively, a log-likelihood ratio statistic and a multisample score test with the Markov Chain method implemented in the software GENEPOP v4.4.3 (Rousset 2008) (10,000 dememorizations, 1000 batches, and 10,000 iterations per batch). Significance was adjusted using Bonferroni correction (Rice 1989). Within-population diversity was classically assessed for both species by estimating the total number of alleles observed (A_T), the observed heterozygosity (H_O) and the gene diversity (H_E). The allelic richness (A_r), an estimator of the allele number free of sampling bias, was estimated using

the software package HPRARE (Kalinowski 2005). We also used GENEPOP to estimate mean genetic differentiation among populations as expressed by F_{ST} estimates and using 10,000 permutations of multilocus genotypes for significance of results. These analyses were carried out only for population with a sampling size of N>=7 individuals.

Genetic distances

Each population was defined as a group of individuals belonging to the same pond. Analysis were carried both at population and individual level to optimize inferences on gene flow. Genetic divergence was quantified using the D_{CE} chord distance (Cavalli-Sforza & Edwards 1967) between pairwise populations. This genetic distance makes no assumption regarding constant population size or mutation rates among loci, assumes genetic drift as the main evolutionary driver of divergence among populations, and appears to be the most efficient in correctly depicting the genetic affinities among closely related populations using microsatellite loci (Takezaki & Nei 1996; Weckworth *et al.* 2013; Séré *et al.* 2017). Pairwise individual kinship coefficients were also estimated to analyze the spatial genetic structure at the scale of individuals. Using SPAGEDI version 1.4 (Hardy & Vekemans 2002), Nason's kinship coefficient F_{ij} (Loiselle *et al.* 1995) was calculated as it is a relatively unbiased estimator with low sampling variance.

Effective population size

The effective population size is a key factor driving the extent of genetic drift within populations (Ellegren & Galtier 2016). N_e was therefore used to provide a quantitative factor that may affect the patterns of genetic differentiation, independently from landscape factors (e.g. Serrouya *et al.* 2012; Weckworth *et al.* 2013). N_e was estimated using the LD method described in Waples (2006) and implemented in NEESTIMATOR v2.0 (Do *et al.* 2014). We reported results for populations with sampling sizes of at least 7 individuals per population, after excluding alleles with frequency less than 0.05 and with 95% confidence intervals derived from a jackknife approach. To investigate how asymmetry in N_e could affect the level of genetic divergence among nearest neighboring populations, pairwise harmonic mean of N_e between nearest-neighbours were estimated. The relationship between D_{CE} genetic distance, adjusted for geographical distance, and harmonic means of N_e was then tested using a nonlinear regression as in Serrouya *et al.* (2012).

Spatial genetic structure

To describe broad population affinities, we used a non-spatially explicit Bayesian clustering analysis without any *a priori* population affiliation, as implemented in the program STRUCTURE v 2.3.4 (Hubisz *et al.* 2009). We used an admixture model with 100,000 burn-in Markov Chain Monte Carlo (MCMC) iterations followed by 5.10^6 MCMC iterations post-burn-in. 30 replicates were generated for each number (K) of clusters tested, ranging from 1 to the total number of sampled populations. As suggested by Wang (2017) we set an alternative ancestry prior, by setting the parameter alpha to $1/N$ (N being the number of sampled populations) which allows for unequal representations of source populations. This parameterization proved to improve individual assignments even if sampling is highly unbalanced. The most likely number of genetic cluster was then inferred using the calculation of the *ad hoc* statistic ΔK following Evanno *et al.* (2005). CLUMP v 1.1 (Jakobsson & Rosenberg 2007) and DISTRUCT v 1.1 (Rosenberg 2004) were finally used for graphical display of the most likely clustering solution.

To further contrast areas of continuous genetic variation from those depicting genetic discontinuities due to strong barriers to gene flow, we additionally used a Bayesian kriging of the rate of local variation of genetic differentiation, as described in Duforet-Frebourg & Blum (2014). This explanatory analysis is not hypothesis-driven but is especially designed for conservation studies dealing with land management of highly fragmented populations (Duforet-Frebourd & Blum 2014). Nonstationary patterns of isolation by distance (IBD) was determined at the scale of individuals using a matrix of pairwise estimates of individual kinship F_{ij} and using the software LocalDiff, available at http://membres-timc.imag.fr/Michael.Blum/LocalDiff.html.

Finally, the extent of spatial genetic structure (SGS) was evaluated at the individual level by applying spatial autocorrelation analyses. To visualize the patterns of SGS, pairwise individual kinship coefficients F_{ij} were averaged over distance classes defined to contain the same number of pairs of individuals. Standard errors of F_{ij} were estimated using a jackknifing procedure over loci and significance of mean F_{ij} estimates was calculated by 10,000 permutations of individual locations. The spatial scale over which gene flow occur, defining a genetic neighborhood in the broad sense, was considered as the geographical distance scale for which mean F_{ij} significantly dropped under zero (Sokal & Wartenberg 1983; Favre-Bac *et al.* 2016). To compare the strength of SGS between spatial autocorrelograms, we calculated the S_p statistic (Vekemans & Hardy 2004): $S_p = -b_F/(1- F_1)$, where b_F is the slope of the regression of the kinship coefficient against the

logarithm of the distance, and F_1 is the mean kinship coefficient of the shortest distance class. These statistics were obtained using SPAGeDi.

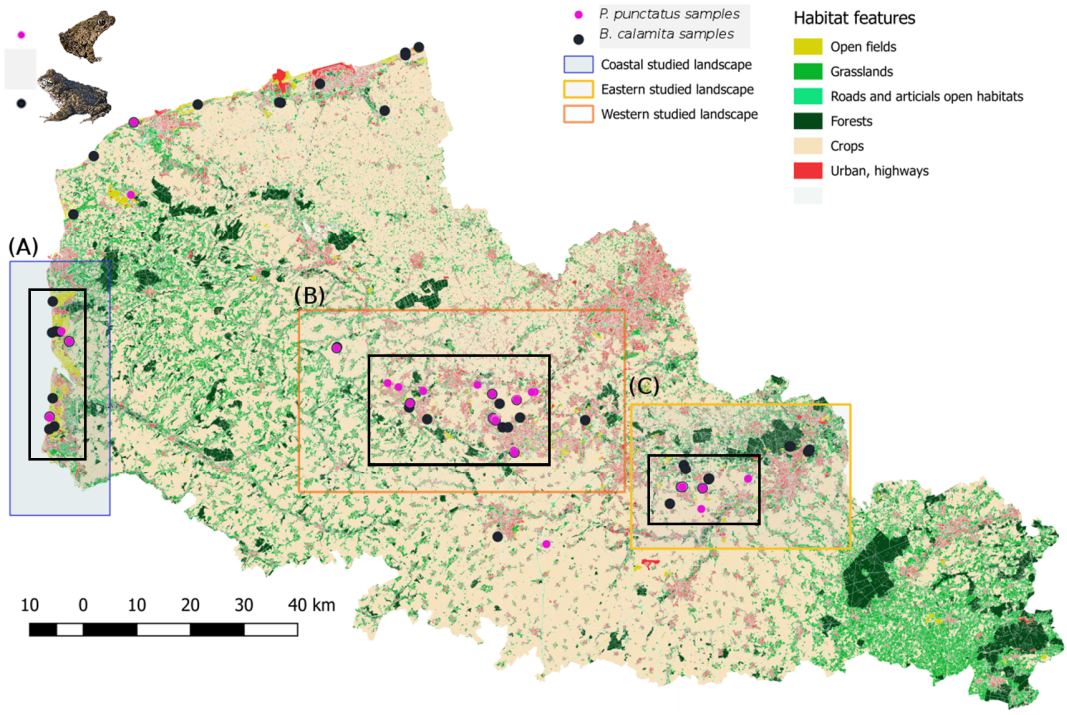


Figure 1: Maps of the three studied areas and population locations, (A) Coast, (B) Western coalfield, (C) Eastern coalfield. The six categories aggregate habitats with *a priori* the same effect on the common parsley frog (*Pelodytes punctatus)* and the natterjack toad (*Bufo* [*Epidalea*] *calamita*): 1) open habitats and other littoral natural habitats suitable for the species, 2) low vegetation habitats, 3) roads and other anthropogenic areas with low vegetation, 4) forests, 5) crops, 6) urban areas, highways and deep water, the first landscape category being the more suitable for species dispersal and the last one being the more impermable for species dispersal. Black dots indicate *B. calamita* sampled populations and red dotes indicate *P. punctatus* sampled populations. The black square defined the sub-areas for which circuit-derived resistance computed were estimated using circuit-based method.

Landscape models

Six main landscape features were retained based on the 2009 land cover map provided by the ARCH project (http://www.archnature.eu/) and describing 63 habitat categories at the scale of 1/10 000 in northern France (Nord-Pas de Calais district). The six landscape categories were defined by taking into account amphibian life-history features, like the probability of mortality, the landscape permeability to individual movements and the habitat suitability to successfully settle and breed. We clustered the habitats with an *a priori* similar effect on the natterjack toad and on

the parsley frog: (1) open habitats, (2) low vegetation habitats (grasslands and gardens), (3) roads, (4) forests, (5) crops, (6) urban areas, highways and deep waters (Table 1). These choices were based on knowledge about basic species ecology and movement behavior on diverse substrates in the natterjack toad, and assuming similar effects for the parsley frog (Stevens *et al.* 2004, 2006a,b, Arntzen *et al.* 2017). The ARCH data set did not distinguish local roads and highways. Highway were therefore manually added with the support of road maps (scale of 1: 200 000) provided by the Institut Géographique National (http://www.ign.fr/). The simplified landscape was converted into a raster map with a grid cell size of 5 meters. Landscape data were managed using the software QGIS 2.18.7 with GRASS 7.2.0 (Figure 1).

We defined six landscape models, five models accounting for different effect of each landscape feature on individual movements and the null model where distance is measured as the Euclidean distance. The first and the second landscape models attributed resistance values from 1 to 6 and from 1 to 600 following a linear distribution for the six habitats defined above (Table 1 and Figure S1). The third model attributed resistance values from 1 to 100 following a logarithmic distribution and the fourth model attributed resistance values from 1 to 6 with an exponential distribution (Table 1). These two distributions assume a stronger effect for corridor features (logarithmic distribution) or barriers features (exponential distribution; Figure S2). These different resistance scales were used because LCP and circuit-based resistance calculations are sensitive to the range of values used to defined the landscape model (Koen *et al.* 2012). Finally, we tested a last model where forests had a lower resistance value than the grasslands and roads, following a linear distribution of values from 1 to 6 (see Stevens *et al.* 2006b).

We additionally built six landscape models specific to each landscape feature with the given landscape feature set as a corridor (with the lowest resistance value=1) while all other landscape features had higher resistance value of 600 (Table 1 and Figure S2). Hence, we had 6 models setting each given landscape feature as a corridor for which pairwise LCP-derived resistances between samples were assessed.

Table 1: Eleven models of landscape resistances based on land cover features that could affect the patterns of gene flow among populations of the natterjack toad (*Bufo* [*Epidalea*] *calamita)* and the parsley frog (*Pelodytes punctatus)*. Six main categories of landscape features clustered together different habitats acknowledged to have a similar effect on species dispersal.

	Euclidean	Resistance_1	Resistance_2	Resistance_3	Resistance_4	Resistance_5	Open habitats	Grasslands	Roads and gardens	Forests	Crops	Urban and deep water
Open habitats	1	1	1	0.3	2.72	1	1	*600*	*600*	*600*	*600*	*600*
Grasslands	1	2	120.8	0.64	82.86	3	*600*	1	*600*	*600*	*600*	*600*
Roads	1	3	240.6	0.98	163	4	*600*	*600*	1	*600*	*600*	*600*
Forests	1	4	360.4	1.32	243.145	2	*600*	*600*	*600*	1	*600*	*600*
Crops	1	5	480.2	1.66	323.287	5	*600*	*600*	*600*	*600*	1	*600*
Urban and deep water	1	6	600	2	403	6	*600*	*600*	*600*	*600*	*600*	1

Ecologically-derived distance inferences

To analyze the effect of landscape features on historical patterns of gene flow, we calculated cumulative cost distance of least-cost paths (LCP) between pairwise populations for the different landscape models defined above. LCPs take into account the heterogeneity of the landscape along a single and optimal pathway and were estimated using the *R* package gdistance (van Etten 2017). We further evaluated the total landscape resistance between pairwise localities by using the circuit theory-derived total resistance as described in McRae (2006). Resistance among population locations were evaluated by nodal analyses using the software CIRCUITSCAPE that converts raster habitats by replacing habitat cells with nodes and connecting nearest-nodes with resistors (McRae 2006; McRae & Beier 2007). Results of Bayesian clustering based on neutral genetic markers split the whole data set of both species into distinct areas (see Results). Consequently, ecological distances were measured for three geographic clusters of populations: the coastal region, the western and the eastern colafield areas (areas A-C on Figure 1). Both analyses were performd in a rectangular area encompassing all populations plus a border band 7 km wide for the LCP analysis, and 3 km wide for the circuit-based resistance. This results in ecological distances measured over 886 km^2 and 161km² for the coastal area A, 2073 km^2 and 654km² for the western inland coalfield area B, and 1105 km^2 and 302km² for the eastern inland coalfield area C (Figure 1). Studied areas included 89 to 142 parsley frogs and 66 to 315 natterjack toads. When a given studied area contained fewer than 6 populations, no result was reported.

Effect of landscape and population size on population genetic divergence

Mantel tests were performed to evaluate the correlation between pairwise D_{CE} genetic distances and pairwise Euclidian or ecological distances (Smouse *et al.* 1986; Sokal *et al.* 1989). We also performed partial Mantel tests to investigate the relationships between pairwise genetic distances and ecologically-modified geographical distances by accounting for Euclidean geographical distance and N_e effects. This last approach provides a measure of the strength of the landscape resistance after removing the effect of Euclidean geographical distance and/or N_e. Finally, significance of the relationship between genetic distance and the distance derived from the best models was confirmed by carrying out partial Mantel tests on Euclidean geographical distance controlling for the model distance derived. All Mantel and partial Mantel tests were conducted using the spatial statistic software PASSAGE 2 and significance of Mantel coefficients (r_z) was assessed using 1000 permutations. Only models that induced non-significant correlations between Euclidean geographical distance and genetic distance by controlling for the derived distance were considered better than Euclidean distance (see Cushman & Landguth 2010).

To complement Mantel tests, we also performed linear mixed effect models fitted by the maximum likelihood population effects (MLPE) parameterization (Clarke *et al.* 2002). Generalized least squares were used with Restricted Maximum Likelihood (REML) estimation, as implemented in the package nlme (Bates *et al.* 2015; Pinheiro *et al.* 2017). Pairwise genetic distances were set as the response variable, and pairwise Euclidean and ecological distances were set as explanatory variables. Pairwise samples were set as random effects and MPLE method accounts for the non-independence among pairwise distances. The random effect term includes a covariate structure where a proportion of the total variance is the result of the correlation between two pairwise distances involving a common population (Clarke *et al.* 2002). All explanatory variables were centered reduced. As explanatory variables showed high level of collinearity, we only carried out univariate models. Support for each explanatory variable was assessed using AIC.

Results

Genetic diversity, levels of genetic differentiation, population genetic affiliation and N_e

Two and 7 microsatellite loci were monomorphic in the natterjack toad and the parsley frog, respectively (Table S1). In both species, exact tests demonstrated no evidence of LD among loci. In natterjack toad, two loci (Bcalμ9 and Bc22) were excluded due to recurrent and significant departures from HW expectations likely resulting from null alleles and/or allele drop-out. In the parsley frog, two loci (PPU5 and ppu12) showed a significant excess in homozygotes at the global level, with inconsistent patterns across populations. The remaining loci matched HW expectations, with A_r ranging from 1.455 to 5.311 and from 1.405 to 4.991 for the natterjack toad and the parsley frog, respectively (Tables S1 and S2). Both species exhibited similar levels of genetic diversity with H_e ranging from 0.379 to 0.679. In terms of mean level of genetic differentiation, the parsley frog populations exhibited a higher level of genetic differentiation than natterjack toad populations, with a mean multilocus F_{ST} of 0.268 (±0.019; $P < 0.001$) and of 0.110 (±0.050; $P < 0.001$) respectively.

For both species, we detected similar and concordant genetic discontinuities among individuals. We observed a continuous variation of gene flow along the coastline in accordance with stationary patterns of IBD. In contrast, strong local genetic differentiation was observed in coal basin, a pattern suggestive of dispersal barriers limiting the among-population connectivity (Figure 2A, B). In terms of broad population genetic affiliation, the ΔK statistics showed two modal K in accordance with two and three genetically distinct clusters for the natterjack populations (Figure 2F, 2G and 2H and Figure S3B). For the Natterjack toad, genetic affiliation matched with geographical areas: coastline areas and inland coalfield areas with a further subdivision into a western and an eastern coalfield genetic cluster, with more than 86% of individuals being assigned with individual probability of membership > 0.9. A slightly different pattern emerged in the parsley frog, with a hierarchical structure in 2 and 6 genetic clusters (Figure 2C, 2D and 2E and Figure S3A). 93% and 56% of individual membership probabilities were > 0.9 for the modal K of 2 and 6, respectively. Overall, parsley frog populations split into two genetically distinct clusters including coastline and western inland populations on the first hand and eastern coalfield populations on the other hand. Further genetic delimitations into 4 or 2 sub-clustering can be visualized for both main clusters (Figure 2D and 2E).

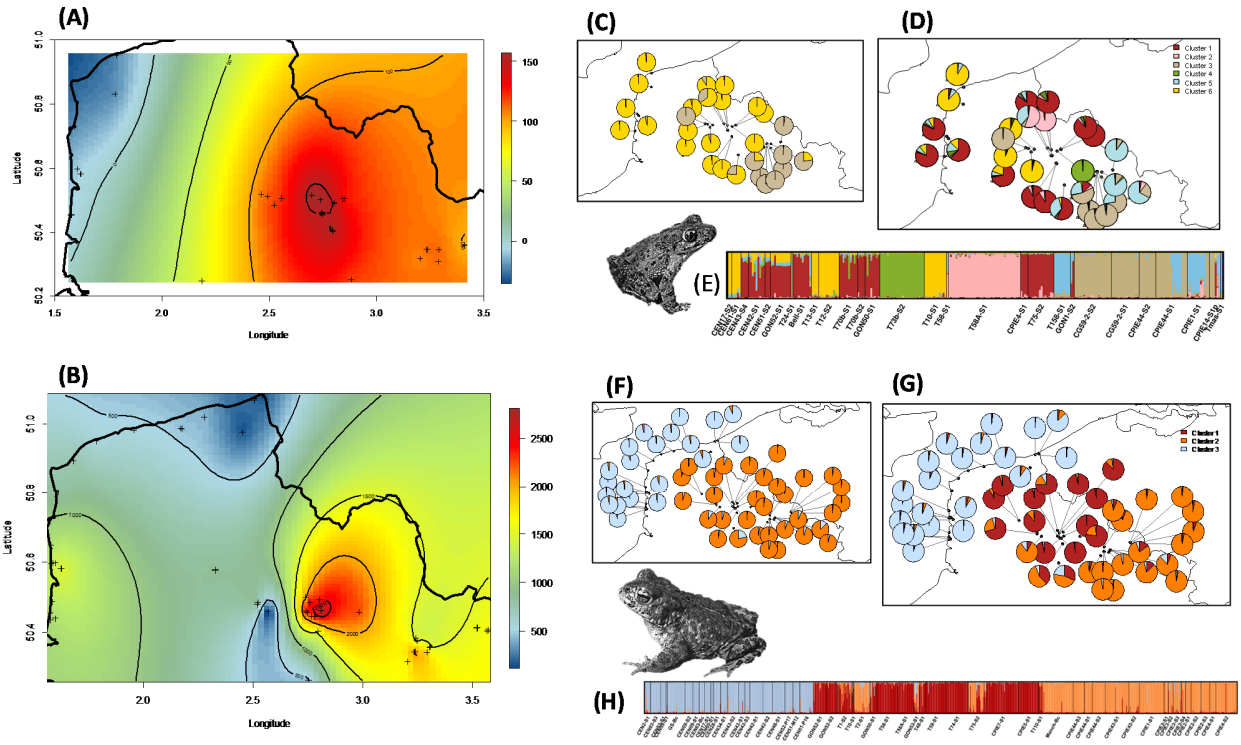


Figure 2: Maps showing the patterns of genetic discontinuities and population affiliation in *Pelodytes punctatus* (on the top) and *Bufo* [*Epidalea*] *calamita* (on the bottom) in the studied area located in northern France. (A) and (B) depict local patterns of genetic differentiation by Bayesian kriging based on pairwise individual kinship coefficients. Figures from (C) to (H) depict the results of Bayesian clustering assignment probability of populations giving their membership in *K*= 2 and *K*= 6 genetic clusters for the parsley frog and *K*= 2 and *K*=3 genetic clusters for the natterjack toad. Each population is represented by a diagram partitioned into *K* colors displaying the mean population estimated membership fractions in *K* clusters. (E) and (G): barplots of individual estimated membership fractions in *K* clusters.

LD-based N_e estimates were lower in the parsley frog ($\widehat{Ne}$ =27) compared to the natterjack toad ($\widehat{Ne}$ =85). A trend for a negative exponential relationship between pairwise harmonic means of N_e and genetic divergence, adjusted or not for geographical distance, was observed for natterjack toad populations located along the coastline and in western coalfield area (Figure 3). However, this genetic drift effect caused by asymmetry in N_e between nearest neighbouring populations was not observed for the parsley frog populations (data not shown).

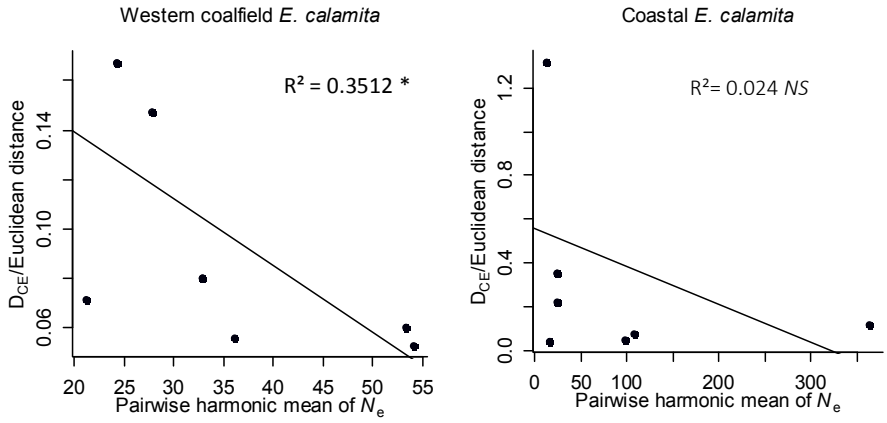


Figure 3: Chord genetic distance (D_{CE}) corrected by Euclidean geographical distance as a function of harmonic mean of pairwise effective population size (N_e) for nearest-neighboring populations. Relationship was modeled using linear regression.

Landscape models of gene flow

Euclidean geographical distance was significantly correlated with genetic distance both at individual and population scale in each area. Euclidean geographical distance appeared to be the most salient factor shaping the extent of gene flow for natterjack toad populations located along the coastline: patterns of IBD were, in most cases, not improved using either LCP or circuit-based resistances in this area (Table 2). Furthermore, partial Mantel tests that controlled for the effect of simple Euclidean geographical distance did not improve the results, whatever the individual or population scale considered. However, for analyses that controlled for the effect of harmonic means of N_e, the Mantel r_z were greater for coastline natterjack populations and for inland coalfield parsley frog populations for both LCP and circuit theory-based resistances.

Table 2: Mantel test and partial Mantel test results (r_z) showing the relationship between genetic divergence and Euclidean geographical distance or ecologically-derived distance estimated by least cost path or circuit theory method using the first six models of landscape resistance described in Table 1. Significance was tested using 1000 permutations of individuals or populations. Estimates of kinship coefficient (F_{ij}) were used at the individual scale and chord genetic distance (D_{CE}) was used at the population scale. Matrices of Euclidean geographical distance and matrices of pairwise harmonic means of N_e were used as control variables for partial Mantel tests on ecologically derived distance. Ecologically derived-distances on best models were used as control variables for partial Mantel tests on Euclidean geographical distance. Analyses were carried out in eastern and western part of coalfield areas for both the parsley frog (*Pelodytes punctatus*) and the natterjack toad (*Bufo* [*Epidalea*] *calamita*), and along the coastline for the natterjack toad. Partial (G.R|E) and partial (G.E| LCP_ Resistance1 or G.E| Circuits_ Resistance1): results from partial Mantel tests between genetic distance and ecological distances controling for Euclidean geographical distance and *vice-versa*; partial (G.R|N_e): results from partial Mantel tests between genetic distance and ecological derived-distances controling for pairwise harmonic mean of effective population size (N_e). ─ non analysed.

	Coastline		West		East			
INDIVIDUALS LEVEL								
Species *N*	*B.calamita* *150 individuals*		*P. punctatus* *142 individuals*		*B.calamita* *324 individuals*		*P. punctatus* *89 individuals*		*B. calamita* *315 individuals*	
	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*
Euclidean	-0.195	*0.011*	-0.387	*0.001*	-0.223	*0.001*	-0.214	*0.001*	-0.286	*0.001*
LEAST COST PATH										
LCP_Resistance 1	-0.199	*0.011*	-0.443	*0.001*	-0.235	*0.001*	-0.234	*0.001*	-0.301	*0.001*
LCP_ Resistance 2	-0.131	*0.011*	-0.386	*0.001*	-0.176	*0.001*	-0.286	*0.001*	-0.223	*0.001*
LCP_ Resistance 3	-0.198	*0.011*	-0.445	*0.001*	-0.236	*0.001*	-0.237	*0.001*	-0.302	*0.001*
LCP_ Resistance 4	-0.147	*0.011*	-0.491	*0.001*	-0.211	*0.001*	-0.310	*0.001*	-0.311	*0.001*
LCP_ Resistance 5	-0.199	*0.011*	-0.442	*0.001*	-0.236	*0.001*	-0.234	*0.001*	-0.303	*0.001*
Partial (G.R|E)										
LCP_ Resistance1	-0.043	*0.001*	-0.431	*0.001*	-0.105	*0.001*	-0.191	*0.001*	-0.246	*0.001*
LCP_ Resistance 2	-0.019	*0.012*	-0.345	*0.001*	-0.126	*0.001*	-0.232	*0.001*	-0.176	*0.001*
LCP_ Resistance 3	-0.032	*0.001*	-0.486	*0.001*	-0.108	*0.001*	-0.202	*0.001*	-0.243	*0.001*
LCP_ Resistance 4	-0.019	*0.011*	-0.359	*0.001*	-0.126	*0.001*	-0.234	*0.001*	-0.180	*0.001*
LCP_ Resistance 5	-0.041	*0.001*	-0.380	*0.001*	-0.109	*0.001*	-0.190	*0.001*	-0.228	*0.001*
Partial (G.E| LCP_ Resistance1)	*NS*		*NS*		*NS*		*NS*		*NS*	
N	66 individuals		142 individuals		318 individuals		89 individuals		232 individuals	
	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*
Euclidean	-0.214	*0.001*	–		-0.229	*0.001*	–		-0.376	*0.001*
CIRCUIT THEORY										
Circuits_ Resistance 1	-0.251	*0.001*	-0.501	*0.001*	-0.286	*0.001*	-0.249	*0.001*	-0.424	*0.001*
Circuits_ Resistance 2	-0.150	*0.001*	-0.361	*0.001*	-0.223	*0.001*	-0.216	*0.001*	-0.374	*0.001*
Circuits_ Resistance 3	-0.249	*0.001*	-0.492	*0.001*	-0.283	*0.001*	-0.247	*0.001*	-0.421	*0.001*
Circuits_ Resistance 4	-0.169	*0.001*	-0.377	*0.001*	-0.232	*0.001*	-0.217	*0.001*	-0.377	*0.001*
Circuits_ Resistance 5	0.045	*0.041*	-0.499	*0.001*	-0.285	*0.001*	-0.242	*0.001*	-0.413	*0.001*
Partial (G.R|E)										
Circuits_ Resistance 1	-0.156	*0.001*	-0.345	*0.001*	-0.179	*0.001*	-0.132	*0.001*	-0.217	*0.001*
Circuits_ Resistance 2	0.096	*0.001*	-0.149	*0.001*	-0.112	*0.001*	-0.048	*0.001*	-0.065	*0.001*
Circuits_ Resistance 3	-0.151	*0.001*	-0.331	*0.001*	-0.175	*0.001*	-0.127	*0.001*	-0.210	*0.001*
Circuits_ Resistance 4	0.074	*0.001*	-0.169	*0.001*	-0.117	*0.001*	-0.051	*0.001*	-0.076	*0.001*
Circuits_ Resistance 5	*NS*		-0.343	*0.001*	-0.179	*0.001*	-0.118	*0.001*	-0.186	*0.001*
Partial (G.E| Circuits_ Resistance1)	-0.039	*0.034*	*NS*		-0.037	0.001	*NS*		*NS*	

POPULATION LEVEL										
N	*10 populations*		*8 populations*		*11 populations*		*6 populations*		*16 populations*	
	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*
Euclidean	0.417	*0.009*	0.439	*0.035*	0.749	*0.002*	0.904	*0.012*	0.516	*0.001*
LEAST COST PATH										
LCP_ Resistance 1	0.394	*0.014*	0.506	*0.013*	0.725	*0.003*	0.913	*0.015*	0.540	*0.001*
LCP_ Resistance 2	*NS*		0.432	*0.033*	*NS*		0.323	*0.046*	0.379	*0.001*
LCP_ Resistance 3	0.396	*0.014*	0.507	*0.010*	0.723	*0.005*	0.914	*0.023*	0.542	*0.001*
LCP_ Resistance 4	*NS*		0.507	*0.013*	*NS*		0.602	*0.036*	0.540	*0.001*
LCP_ Resistance 5	0.399	*0.010*	0.499	*0.014*	0.729	*0.002*	0.914	*0.013*	0.548	*0.001*
Partial (G.R|E)										
LCP_ Resistance 1	*NS*		0.442	*0.024*	*NS*		*NS*		0.423	*0.001*
LCP_ Resistance 2	*NS*		*NS*		*NS*		*NS*		0.313	*0.002*
LCP_ Resistance 3	*NS*		0.441	*0.027*	*NS*		*NS*		0.414	*0.001*
LCP_ Resistance 4	*NS*		0.441	*0.040*	*NS*		*NS*		0.319	*0.001*
LCP_ Resistance 5	*NS*		*NS*		*NS*		*NS*		0.544	*0.000*
Partial (G.E| LCP_ Resistance1)	*NS*		*NS*		*NS*		*NS*		*NS*	
Partial (G.R|Ne)										
LCP_ Resistance 1	0.398	*0.003*	0.556	*0.008*	0.715	*0.005*	0.952	*0.001*	0.540	*0.001*
LCP_ Resistance 2	*NS*		*NS*		*NS*		*NS*		0.369	*0.001*
LCP_ Resistance 3	0.406	*0.008*	0.556	*0.007*	0.712	*0.004*	0.952	*0.001*	0.541	*0.001*
LCP_ Resistance 4	*NS*		0.556	*0.014*	*NS*		0.632	*0.021*	0.534	*0.001*
LCP_ Resistance 5	0.406	*0.019*	0.548	*0.015*	0.719	*0.003*	0.953	*0.001*	0.547	*0.001*
N	4 populations		8 populations		11 populations		6 populations		11 populations	
			r_z	*P*	r_z	*P*	r_z	*P*	r_z	*P*
Euclidean	–		0.439	*0.035*	0.749	*0.002*	0.904	*0.012*	0.623	*0.001*
CIRCUIT THEORY										
Circuits_ Resistance 1	–		*NS*		0.470	*0.047*	*NS*		0.733	*0.001*
Circuits_ Resistance 2	–		*NS*		*NS*		0.964	*0.030*	0.693	*0.001*
Circuits_ Resistance 3	–		*NS*		*NS*		0.956	*0.031*	0.732	*0.001*
Circuits_ Resistance 4	–		*NS*		*NS*		0.964	0.028	0.705	0.001
Circuits_ Resistance 5	–		*NS*		*NS*		0.953	*0.021*	0.710	*0.001*
Partial (G.R|E)										
Circuits_ Resistance 1	–		*NS*		0.511	*0.021*	0.646	*0.037*	0.553	*0.001*
Circuits_ Resistance 2	–		*NS*		*NS*		0.842	*0.041*	0.400	*0.011*
Circuits_ Resistance 3	–		*NS*		0.498	*0.034*	*NS*		0.549	*0.001*
Circuits_ Resistance 4	–		*NS*		*NS*		0.840	0.038	0.435	0.001
Circuits_ Resistance 5	–		*NS*		0.481	*0.041*	*NS*		0.458	*0.004*
Partial (G.E| Circuits_ Resistance1)	–		0.536	*0.017*	*NS*		*NS*		*NS*	
Partial (G.R|Ne)										
Circuits_ Resistance 1	–		*NS*		0.525	*0.018*	0.988	*0.004*	0.724	*0.001*
Circuits_ Resistance 2	–		*NS*		*NS*		0.992	*0.003*	0.673	*0.001*
Circuits_ Resistance 3	–		*NS*		0.515	*0.033*	0.988	*0.004*	0.723	*0.001*
Circuits_ Resistance 4	–		*NS*		*NS*		0.992	*0.001*	0.687	*0.001*
Circuits_ Resistance 5	–		*NS*		0.495	*0.038*	0.988	*0.006*	0.697	*0.001*

Analyses of relationships between genetic divergence and ecologically-derived geographical distances derived from the 6 landscape models revealed contrasting patterns among studied areas (Table 2). At the individual level, all geographical variables significantly correlated with pairwise individual kinship estimates and ecologically-derived resistance outperformed Euclidean geographical distance in explaining genetic distance. In parsley frog individuals, Mantel r_z values always increased for LCP-derived resistance compared to Euclidean distance, with a better fit for landscape model 3 which emphasizes prevailing corridors effects in the western coalfield, while landscape model 4 emphasizing prevailing barriers effects performed better in the eastern coalfield area (Table 2). For natterjack toad individuals, LCP-derived resistances yielded inconsistent results among areas. In contrast, the use of circuit-based resistances always increased the signal of IBD pattern, with the strongest correlation occurring for model 1 in the parsley frog (r_z = -0.501 and -0.249 for western and eastern coalfield areas, respectively, all at $P < 10^{-3}$) and in the natterjack toad (r_z = -0.286 and -0.424 for western and eastern coalfield areas, respectively, all at $P < 10^{-3}$) (Table 2). There were no significant correlations between Euclidean geographical distances and F_{ij} estimates using partial Mantel test controlling for landscape model 1.

At the population level, LCP-derived resistances outperformed the use of Euclidean distance for parsley frog populations, with a trend for increased r_z values for landscape model 3 that highlights potential corridor effects. Partial Mantel tests controlling for the effect of N_e improved these IBD signatures, with Mantel r_z values ranging from 0.556 to 0.953 (Table 2). In the natterjack toad populations, the use of LCP and circuit-based resistances increased the relationship with D_{CE} genetic distance for only eastern coalfield populations, especially when controlling for the effect of N_e. LCP-based distances gave the better fit with landscape model 5 in the eastern area but no trend emerged in the western area. In accordance with results at individual level, the use of circuit-based resistances showed the strongest IBD pattern for landscape model 1 in coalfield areas for the natterjack toad.

Keeping in mind that AIC is not the best suited method to compare fixed effects when using REML (see Clarke *et al.* 2002), linear mixed models using maximum-likelihood population effect gave further insights and consistent results into the relative effects of landscape features on genetic divergence. In the parsley frog, the best-supported mixed effect landscape models were crops set as corridors in the eastern area (ρ = 0.437 and 0.112 at the population and individual level respectively, $P < 0.001$), grasslands set as corridors at the population scale in the western area (ρ =

0.320 $P < 0.001$), and roads set as corridors at the individual scale in the western area ($\rho = 0.113$, $P < 0.001$) (Table 3). In the natterjack toad, open natural habitats ($\rho = 0.414$, $P < 0.001$) and roads ($\rho = 0.058$, $P < 0.001$) were best supported at the population and individual levels respectively for eastern coalfield area. Roads also appeared to be the best supported landscape variable in shaping genetic divergence at both individual and population scale in western coalfield area ($\rho = 0.251$ and 0.054 respectively, $P < 0.001$) and along the coastline ($\rho = 0.169$ and 3.08^{-10}, $P < 0.001$). Concordantly with Mantel test results, the Euclidean geographical distance was the second best landscape model explaining the patterns of genetic divergence for coastal area.

Table 3: Results from maximum likelihood population effects models for the first three best univariate models obtained on chord distance (D_{CE}) or kinship coefficient (F_{ij}) at population and individuals scale, respectively. Euclidean geographical distances, LCP-derived distances from landscape models caracterizing one landscape feature as a corridor (Table 1) or effective population size (at population scale) were used as explanatory variables. Distribution and collinearity of variables can be visualised Figure S4A and B.

		POPULATIONS						INDIVIDUALS					
		Feature	AIC	Rho	Coeff.	Residual standard error	Δ AIC	Feature	AIC	Rho	Coeff.	Residual standard error	Δ AIC
B. calamita	Eastern coalfield	Open habitats	-404.96	0.414	0.07	0.08	0	Grasslands	-1.28E+05	0.058	-0.035	0.058	0
		Urban	-387.14	0.093	0.03	0.05	17.82	Roads	-1.28E+05	0.023	0.029	0.023	328.4
		Crops	-385.57	0.106	0.03	0.05	19.39	Crops	-1.27E+05	0.019	-0.027	0.019	1033.8
	Western coalfield	Roads	-219.41	0.251	0.031	0.032	0	Roads	-135003.5	0.054	-0.033	0.07	0
		Euclidean	-213.47	0.276	0.237	0.035	5.93	Crops	-134248.8	0.053	-0.032	0.07	754.7
		Crops	-212.2	0.279	0.028	0.036	7.2	Euclidean	-134217.4	0.058	-0.033	0.071	786.1
	Coastline	Roads	-172.22	0.169	0.016	0.03	0	Roads	-2.93E+04	3.08E-10	-0.013	0.065	0
		Euclidean	-171.61	0.172	0.016	0.03	0.61	Euclidean	-2.93E+04	3.80E-10	-0.013	0.065	19.2
		Crops	-171.03	0.167	0.016	0.03	1.19	Crops	-2.93E+04	3.17E-10	-0.013	0.065	37.13
P. punctatus	Eastern coalfield	Crops	-40.15	0.437	0.098	0.013	0	Crops	-6189.52	0.112	-0.071	0.121	0
		Forests	-39.16	0.443	0.106	0.015	0.98	Open habita	-6177.22	2.44E-10	-0.032	0.11	11.19
		Roads	-39.03	0.438	0.093	0.013	1.11	Euclidean	-6120.79	0.104	-0.068	0.121	69.99
	Western coalfield	Grasslands	-74.58	0.32	0.045	0.06	0	Roads	-1.34E+04	0.113	-0.104	0.138	0
		Open habitats	-74.22	0.315	0.044	0.06	0.354	Grasslands	-1.31E+04	0.103	-0.1	0.138	287.51
		Forests	-73.23	0.3	0.045	0.06	1.346	Open habita	-1.30E+04	0.103	-0.1	0.139	321.24

Spatial genetic structure

Congruent results arose from spatial correlograms designed using Euclidian geographical distances and best-matching LCP and circuit-based geographical distances (Figure 4). Best matching LCP and circuit-derived resistances were those derived from landscape model 1. In all cases, a decrease of mean kinship estimates among individuals with geographical distance was observed, indicating a significant IBD. Strongly concordant results were found in the natterjack toad: the strength of SGS increased using ecologically-derived geographical distance compared

with Euclidean distance, with the highest signal of IBD using circuit-derived resistance (*Sp* = 0.055*** in eastern coalfield, *Sp* = 0.034*** in western coalfield, Figure 4E). Moreover, the spatial-scale of positive SGS was nearly identical for western and eastern coalfield areas, with mean natterjack toad individual kinship coefficient dropped under zero at similar scale: 5 km using Euclidean geographic distance, a resistance value of 250 for LCP-derived distance and a resistance value of 2 to 2.5 Ω for circuit-derived distances (Figure 4D, E, F). We performed the analyses using only Euclidean geographical distance for coastal natterjack toad populations because this metric proved to best fit with genetic divergence. A similar spatial scale of 5 km for positive SGS was found (data not shown). In the parsley frog, the SGS was stronger in western coalfield area using the LCP-derived resistance (Figure 4A, B, C). Identical threshold of positive SGS were found when compared to the natterjack toad: 5 km, a resistance value of 250 for LCP-derived resistance and a circuit-derived resistance of 2.5 Ω.

Mapping of connections among populations using these thresholds of positive SGS can be visualized in Figure 4. Euclidean geographical distance and LCP-derived distance gave similar connecting networks that occurred at very short spatial scale and only involved clusters of nearest-neighbors for both natterjack and parsley frog populations. The threshold of connection was more stringent using circuit-derived distances and suggested a striking isolation of population even for nearest neighbors (Figure 4).

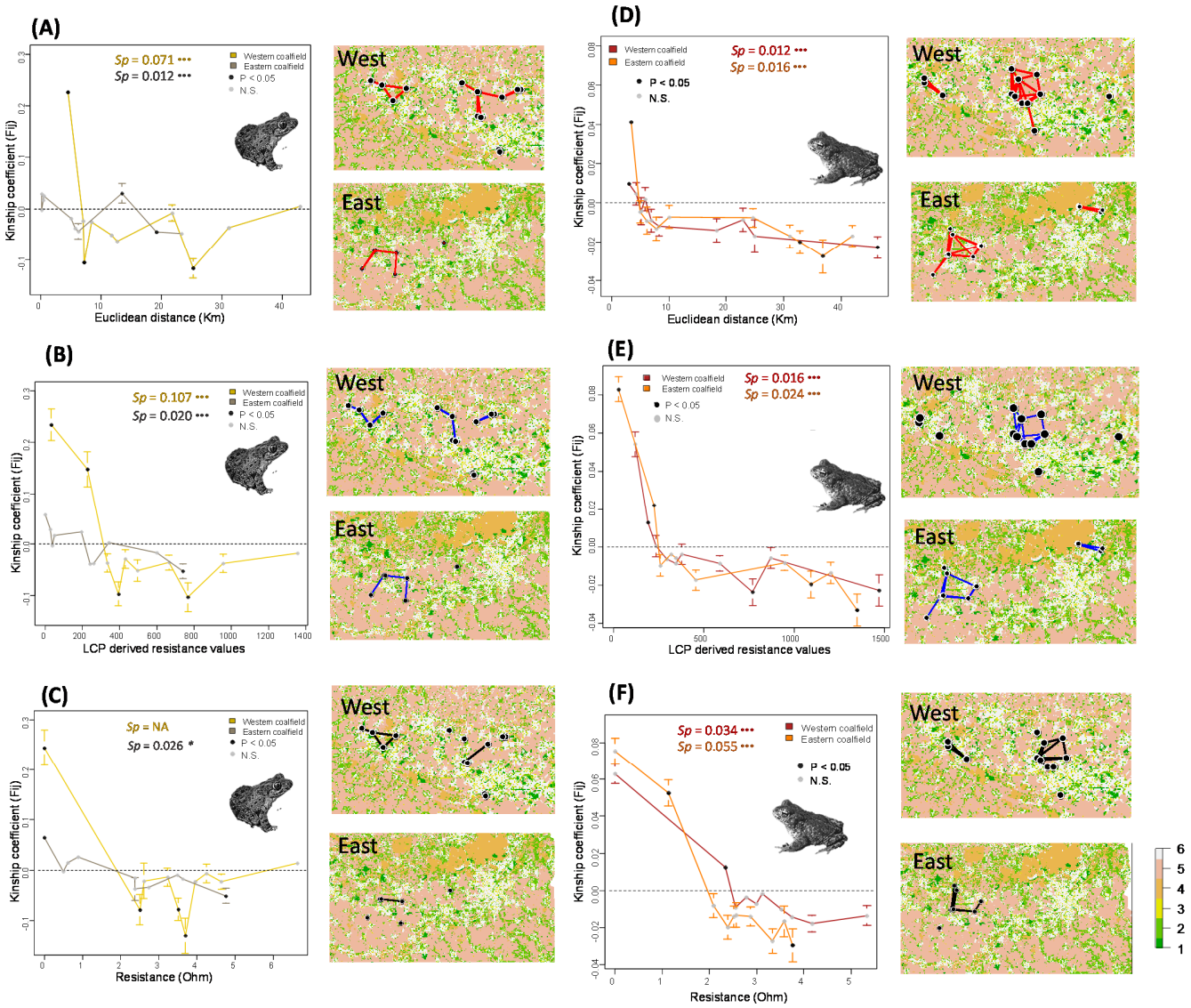


Figure 4: Spatial genetic structure in western and eastern coalfield areas. Autocorrelograms depict the relationship between average pairwise kinship coefficients (F_{ij}) among individuals and distance in the common parsley frog (*Pelodytes punctatus)* (A, B, C) and the natterjack toad (*Bufo* [*Epidalea*] *calamita*) (D, E, F). Black dotes indicated mean F_{ij} significantly different from random expectations. Standard errors were estimated by jackknifing. F_{ij} is plotted against (A, D) Euclidean geographical distance; (B, E) LCP-derived distance using the landscape model 1; (C, F) circuit-based resistance using the landscape model 1. Connecting network defined by the thresholds of positive spatial genetic structure depicted by autocorrelograms are presented on the maps depicting landscape resistance of model 1.

Discussion

Successful persistence of wild populations established in anthropogenically-impacted areas is a long-standing question in the context of worldwide human-induced changing environments (Fahrig 2003, 2017; Mimura *et al.* 2017). The level of intrapopulation gene diversity is the key to adaptive evolution in such highly modified environments (Miraldo *et al.* 2016). Two major microevolutionary processes drive the extent of neutral genetic diversity: gene flow and genetic drift (Wright 1931; Slatkin 1985; Baguette *et al.* 2013; Ellegren & Galtier 2016; Lowe *et al.* 2017). Therefore, understanding (i) the influence of asymmetry in local N_e and (ii) the influence of spatial arrangement and composition of land cover in shaping patterns of gene flow among fragmented populations is a prerequisite to design efficient conservation and management baselines. In this respect, multispecies comparative studies may hold great potential to assess the impact of the landscape matrix on levels of population connectivity and to provide cost-effective management plans (e.g. Goldberg & Waits 2010; Richardson 2012; Gutiérrez-Rodríguez *et al.* 2017). In this study, we investigated the influence of human-modified land cover on patterns of amphibian gene flow through the analysis of spatial genetic structure in relation to specific current landscape features and local N_e in two pioneering pond-breeding anurans. We hypothesized landscape features such as open fields/habitats (*i.e* early successional open canopy habitats), grassland or forests to play a role in determining efficient gene flow pathways for both species and greater genetic differentiation for the parsley frog. Our results supported some of these expectations, with the occurrence of (i) a stronger genetic differentiation among populations and a higher level of individual kinship in the parsley frog, (ii) strong genetic discontinuities depicted by high rate of local genetic differentiation in urbanized and reclaimed mine environments for both species, (iii) significant influences of landscape features in the former areas compared to more continuous natural habitats located along the coastline, and (iv) genetically differentiated lineages for which distinct landscape features impacted/disrupted the patterns of gene flow. This study is the first to perform a multispecies analysis of the effect of landscape connectivity in a post-industrial area. Evolutionary and conservation management implications of our results are discussed below.

Population genetic structure and effective population size

Limited dispersal capabilities, pronounced breeding site philopatry, specific habitat requirement for breeding and moderate N_e are thought to be major drivers of genetic divergence in amphibians (Beebee & Griffiths 2005; McCartney-Melstad & Shaffer 2015; Gutiérrez-Rodríguez *et al.* 2017). Concordantly, substantial levels of genetic differentiation was observed among populations in the parsley frog (mean F_{ST} = 0.268) and in the natterjack toad (mean F_{ST} = 0.110). This suggests small-scaled genetic structuring in line with broad genetic differentiation generally observed in various species of amphibians, for which mean F_{ST} estimates generally ranged from 0.022 for very fine-scaled metapopulation structure to 0.141 for larger geographical scales involving populations disjointed by more than 500 km (e.g. Rowe *et al.* 2000b; Vos *et al.* 2001; Arens *et al.* 2007; Knopp & Merilä 2009; Dudaniec *et al.* 2012; Wang 2012; Pisa *et al.* 2015). All things being equal, the strong levels of genetic differentiation observed in both species can be attributed to low genetic connectivity among populations, especially in inland areas. Indeed, whereas continuous variation of gene flow leading to stationary patterns of IBD can be found in less anthropogenically-impacted habitats located along the coastline, marked genetic discontinuities were depicted in inland habitats characterized by different slag heap massifs surrounded by a matrix of urbanized environment (Figure 2A, B). We further depicted distinct genetic lineages, especially for the parsley frog for which genetic distinctiveness even occurred at the scale of neighbouring spoil heaps with high level of individual kinship (see Figures 2 and 4). Altogether, these results suggested restricted dispersal at fine spatial scales (Slatkin 1985; McCartney-Melstad & Shaffer 2015).

In addition, the low level of genetic diversity in the parsley frog may enhance the population genetic differentiation observed in the parsley frog. Indeed, several loci were monomorphic whereas they showed high polymorphism (up to 20 alleles) in the southern part of the parsley frog geographical distribution (Jourdan-Pineau *et al.* 2009; Van De Vliet *et al.* 2009). Sequential past founder events, jointly with associated bottlenecks and small N_e, may explain higher genetic differentiation and low polymorphism because northern France is the expanding geographical margin of this species. As such, genetic drift effects enhanced by leading-edge expansions and human disturbances of land-cover may explain this pattern of decreasing gene diversity and strong genetic differentiation, given the regional scale of our study area. This core-to-edge patterns of

genetic structure, rooted in the central-marginal hypothesis, is observed in a large number of different taxa (e.g. Micheletti & Storfer 2015; reviewed in Waters *et al.* 2013).

Beyond past history colonization and/or a lack of historical gene flow spanning several generations, departures of migration-drift equilibrium may also involve critically low N_e and asymmetry in local N_e between nearest-neighboring populations, a factor often underestimated in conservation genetics (e.g. Serrouya *et al.* 2012). Indeed, assuming a general effect of landscape in limiting gene flow, N_e is another key factor determining the rate of loss of genetic diversity, especially for isolated and remnant populations (Wright 1931; Ellegren & Galtier 2016; reviewed in Lowe *et al.* 2017). With unbalanced effects of genetic drift and gene flow, we could expect a negative exponential relationship between genetic divergence and mean pairwise N_e for nearest-neighbor populations (e.g. Serrouya *et al.* 2012; Weckworth *et al.* 2013; Lourenço *et al.* 2017). This trend was only observed in the natterjack toad for coastal populations and western coalfield populations, suggesting drift as a predominant evolutionary driver of genetic structure among part of the smallest populations in these areas. This non-linear effect of N_e on patterns of genetic divergence is not trivial as it can confound the effect of landscape components on genetic divergence outcomes (Serrouya *et al.* 2012).

Nonetheless, this effect was not observed in the parsley frog and in the eastern coalfield area for the natterjack toad. Beyond a lack of statistical power because of low number of studied populations, this lack of effect of N_e on genetic structuring often involved adjacent populations with large N_e but exhibiting high level of genetic divergence, even when genetic divergence was corrected by the geographical distance. In that respect, we cannot exclude that past independent colonization events, along with major current barriers to dispersal, left genetic signatures of departures from migration-drift equilibrium independent of N_e effects.

Influence of landscape features on genetic divergence

To minimize the effect of past population history on patterns of genetic divergence, we carried out the analyses of landscape effects within the main depicted genetic lineages for which we can reasonably assume a spatial scale falling within the movement range of individuals, *i.e.* within a few km, the spatial scale of individual toad dispersal (Beebee & Denton 1996; Sinsch 1997; Rowe *et al.* 2000b; Stevens *et al.* 2006b). Furthermore, restricting analyses to adjacent pairs of populations can improve power and accuracy of landscape analyses (Jaquiéry *et al.* 2011).

Patterns of amphibian connectivity are known to be greatly influenced by land cover composition and environmental factors, such as roads and other paved surfaces, soil moisture, vegetation and canopy cover, riparian networks (Vos *et al.* 2001; Janin *et al.* 2009; Goldberg & Waits 2010; Richardson 2012; Youngquist *et al.* 2017). For instance, Gutiérrez-Rodríguez *et al.* (2017) carried out resistance optimization procedure on continuous landscape variables of land cover. They found a positive role of structural heterogeneity with scrubland and open oak woodlands facilitating connectivity among populations in pond breeding amphibians inhabiting highly fragmented Mediterranean landscapes. In the natterjack toad, Stevens *et al.* (2006a,b) showed that toadlet habitat preference was the best variable explaining population genetic structure and thus, dispersal among populations located in a fragmented landscape. In particular they highlighted a positive effect of forests on gene flow among populations. Conversely, Flavenot *et al.* (2015) found that areas of dense wood negatively correlated with allelic richness in natterjack toad populations, while it positively correlated in common toad (*Bufo bufo)* populations. They also showed that bare ground such as quarrying habitats host highly genetically diverse natterjack toad populations.

In this respect, we aimed at characterizing the effect of landscape features acknowledged to potentially favor or impede dispersal among structured amphibian populations in contrasted areas: semi-natural habitats along the coastline and post-industrial and urban habitats in inland areas. Our results clearly indicated a significant IBD signature in both parsley frogs and natterjack toads whether geographical distance accounted or not for landscape composition (Table 2 and Figure 4). Mantel tests showed significant correlations between Euclidean or ecologically-derived geographical distance and the genetic divergence among populations or the genetic relatedness among individuals. As expected, our results suggested a predominant effect of Euclidean geographical distance for natterjack populations located along the coastline. This area indeed constitutes relatively continuous swaths of natural habitats like sandy dunes that allows stationary patterns of IBD reflecting spatial patterns of post-glacial recolonization, as suggested by Faucher *et al*. (2017). In contrast, partial Mantel tests indicated that ecologically-derived geographical distance estimated by LCP and circuit-derived resistance clearly outperformed Euclidean geographical distance in explaining patterns of genetic connectivity among populations located in coalfield areas. Our results are in agreement with those of Rowe *et al.* (2000b) that documented reduced gene flow associated with urban areas in the natterjack toad. Overall, we observed that the landscape model 1, which assumes that open fields and grasslands may facilitate gene flow,

performed the best in these post-mining industrial areas. Furthermore, controlling for N_e improved the fit of landscape models tested, which means that genetic drift was further enhanced for smallest populations. Partial Mantel tests have often been used in population genetic studies to identify connectivity models that are significant beyond their correlation with Euclidean distance (Cushman & Landguth 2010). Some authors have criticized these tests for ascribing significance to connectivity metrics that are merely correlated with migration but not truly influential on their own (Balkenhol *et al.* 2009). Nonetheless, Cushman & Landguth (2010) showed that partial Mantel tests can be successfully applied when alternative models are tested against each other, as we did here to identify the most biologically relevant landscape model to explain the observed genetic divergence. Thus, our results were in line with previous studies dealing with landscape effects on American or European pond-breeding amphibians (e.g. Arens *et al.* 2007; Coster *et al.* 2015; Youngquist *et al.* 2017; Gutiérrez-Rodríguez *et al.* 2017).

Broadly speaking, urban areas, highways and crops acted as effective barriers that impede gene flow in the parsley frog and the natterjack toad. In contrast, open habitats, grasslands, and surprisingly, roads have good permeability to gene flow in the post-industrial areas we studied. These effects were particularly pronounced for the parsley frog, which suggests a higher sensitivity to landscape components together with a less ability for dispersal compared with the natterjack toad. The effect of forest habitats remained equivocal. Whereas Stevens *et al.* (2006a) found that natterjack toadlets actively select forested environments when dispersing, our results suggested that such habitats have no effects or decrease the connectivity among populations over large and different landscapes. Our studied area is the less wooded district in France. Forests are in expansion for two centuries but still cover less than 10% of the landscape (Figure 1). Hence, inconclusive results may arise from an under-representation of this landscape feature in northern France.

It should be noted that the use of Mantel tests have limitations because of the non-independence of the distance data structure and the inability to precisely quantify the proportion of variance explained by subtle spatial structures such as landscape variables (Legendre & Fortin 2010). In this way, generalized mixed effect models may circumvent these drawbacks by determining which landscape features may better explain the population genetic structure. Applying such kind of analysis, Youngquist *et al.* (2017) studied the effect of landscape on population genetic structure in the cricket frog and did not obtain statistical support for combined landscape model that took into account the resistance of different landscape features. However,

they showed a significant effect of highways acting as barriers on cricket frog gene flow. In the same way, Vos *et al.* (2001) and Arens *et al.* (2007) did not obtain statistical support to conclude on the effect of combined landscape features on moor frog patterns of gene flow, except for the roads that acted as barriers to individuals' dispersal.

Using generalized mixed effect models, we tested for potential corridor effects. Open habitats, grasslands, and surprisingly roads appeared to act as efficient corridors for gene flow in the natterjack toad. Slightly different results were obtained for the parsley frog: either crops or roads appeared the most likely corridors facilitating gene flow, depending on whether eastern of western coalfield areas were considered. Concordantly with Mantel tests and maps of local genetic differentiation, the Euclidean distance fitted the best to natterjack toad population genetic divergence for areas located along the coastline, with, however, a significant effect of roads when analyses were performed at the individual scale. This general counterintuitive effect of roads may be explained by the high attractiveness of natterjack toads for bare grounds (Stevens *et al.* 2006a). More generally, this effect could also be explained by the linear nature of such continuous open habitats characterized by low or no vegetation strata which can provide suitable paths for efficient dispersal. Nonetheless, roads and other paved surfaces are also acknowledged to induce very high mortality by vehicular traffic (Hels & Buchwald 2001; Goldberg & Waits 2010; reviewed in Beebee 2013). Road-kill effects may thus strongly mitigate the effects of roads on population connectivity, especially in coalfield areas characterized by a dense road network. The potential positive effect of crops on the parsley frog is also surprising because agricultural land have been identified as major barrier for dispersal for related amphibian species (Janin *et al.* 2009). Indeed, crops located in northern France are acknowledged to be wide open area with some potentially toxic pollutants and few hiding places leading to high predation risk. Furthermore, most of them are very disturbed for harvesting in the middle of summer and early autumn. However, crops encompass different categories of land uses and should be analyzed at a finer scale to disentangle the different effects of this kind of habitat. Finally, landscape resistance estimates encompass several ecological processes (physiological constraints, predation, locomotor capabilities, e.g. Stevens *et al.* 2004, 2006a) that define the connectivity among population. Given the lack of knowledge on parsley frog dispersal behavior, further investigations are needed to better identify species-specific preferences for different landscape features.

Methods and limitations of inferences

Landscape composition has an undeniable effect on pond breeding amphibians (Vos *et al.* 2001; Stevens *et al.* 2004; Arntzen *et al.* 2017; Youngquist *et al.* 2017; Gutiérrez-Rodríguez *et al.* 2017). However, the fact that no distinction can be done among the different landscape models could be explained by several methodological and ecological issues.

Firstly, each measure of ecological distance implies distinct hypotheses on the parameters conditioning the extent of gene flow. The circuit-derived resistances take into account the whole landscape mosaic located between pairwise localities and result in an estimate of dispersal probability assuming a random walk (McRae 2006). The LCP-based resistances only account for the easiest habitat patch to cross. At the landscape resolution of the study area (5 m), we accurately capture the landscape heterogeneity and it is unlikely to generate bias by underestimating the role of small patches of habitats acting as corridors. Amphibians generally present high orientation abilities with a complex set of olfactory, visual, acoustic, mechanical, and magnetic cues to better apprehend the surrounding habitats (Sinsch 1992; Dall'antonia & Sinsch 2001). In consequence, LCP-derived distances did not appear biologically irrelevant, given that amphibians may differentially select among different habitat types during dispersal and are able to optimize their movements. The high fragmentation of the coalfield areas leads to a landscape characterized by a mosaic of urban, agricultural and smaller fragmented habitats. The circuit-based resistance estimates may thus overestimate the ecological distance between populations given that amphibians may select efficient paths to avoid less permeable areas.

Secondly, distances derived from LCP approach and circuit theory are sensitive to the landscape composition and structure that can explain a contrasting signal of best landscape model between western and eastern coalfield areas when taking into account the relative resistance of the different landscape features. For instance, Jaquiéry *et al.* (2011) showed that barriers that strongly impede dispersal are the easiest to identify while power and accuracy decreased when the permeability of landscape elements increased. As such, it is not surprising that we did not observe a clear distinction between model 1 and model 5 that inverse the hierarchy among landscape features of low resistance values by setting forest as better corridor than grassland. Additionally, the effect of a given landscape feature may be easily detected when this feature is widely represented across the studied area. Hence, effects of larger and linear habitats like road or crops were more likely to be detected.

Finally, we must keep in mind that inferring landscape effects on individual dispersal by estimating the extent of gene flow from neutral genetic diversity is a tricky approach given that contemporary genetic divergences depend both from contemporary and historical patterns of dispersal events (Slatkin 1985; Frankham *et al.* 2002; Epps & Keyghobadi 2015). Indeed, changes in patterns of genetic structure can have a substantial time lag related to N_e. If the time since population separation has been insufficient for significant divergence due to random genetic drift, populations may not be genetically differentiated, even in the absence of present-day gene flow. Reversely, observed genetic differentiation may mirror past population history more than current population exchanges (Bossart & Pashley Prowell 1998; Epps & Keyghobadi 2015).

Evolutionary significant units, management and conservation implications

Gene flow is an evolutionary process that supports genetic connectivity and contributes to adaptive evolution to environmental changes. Gene flow are driven by landscape composition but also strongly depend on species dispersal capabilities. Population genetic divergence and individual estimates of kinship suggested lower dispersal capacities and/or higher philopatry in the parsley frog. The spatial scale over which we observed a positive SGS (Figure 4) were smaller than those observed in other amphibian species located in less human-disturbed area. For instance, Coster *et al.* (2015) found that populations of spotted salamander (*Ambystoma maculatum*) and wood frog (*Lithobates sylvaticus*) are still connected over a scale of 15 km in a forested area with few urban infrastructures. Thereby, connectivity among isolated adjacent populations should be preserved and improved, when possible, to prevent population from loss of genetic diversity. Besides, Stevens & Baguette (2008) highlighted the importance of connectivity in providing new migrants populations to prevent populations from the extinction risk link to the high environmental stochasticity in pioneering habitats. Thus, managing landscape to preserve safe corridors with low vegetation among the target habitats should be preconized. Yet, habitat quality is also a key parameter for population persistence (Stevens & Baguette 2008). With the end of mining activity without any specific management, the spoil heaps will become more and more wooded with no other suitable habitats to colonize in the surrounding area, populations could extinct. Hence, habitat management will be essential to maintain suitable habitat for pioneering species that need early successional habitat.

References

Allendorf FW, Luikart GH, Aitken SN (2012) *Conservation and the Genetics of Populations*. Wiley-Blackwell, Oxford (UK).

Allentoft ME, O'Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, **2**, 47–71.

Arens P, Sluis T Van Der, Van't Westende WPC *et al.* (2007) Genetic population differentiation and connectivity among fragmented Moor frog (*Rana arvalis*) populations in the Netherlands. *Landscape Ecology*, **22**, 1489–1500.

Arntzen JW, Abrahams C, Meilink WRM, Iosif R, Zuiderwijk A (2017) Amphibian decline, pond loss and reduced population connectivity under agricultural intensification over a 38 year period. *Biodiversity and Conservation*, **26**, 1411–1430.

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C (2013) Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, **88**, 310–326.

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

Banks SC, Cary GJ, Smith AL *et al.* (2013) How does ecological disturbance influence genetic diversity? *Trends in Ecology and Evolution*, **28**, 670–679.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1–48.

Beebee TJC (2013) Effects of road mortality and mitigation measures on amphibian populations. *Conservation Biology*, **27**, 657–668.

Beebee TJC, Denton JS (1996) *The natterjack toad conservation handbook (Species recovery program)*. English Nature, Northminster House, Peterborough PE1 1UA.

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

Bossart JL, Pashley Prowell D (1998) Genetic estimates of population structure and gene flow: Limitations, lessons and new directions. *Trends in Ecology & Evolution*, **13**, 202–206.

Buonomo R, Assis J, Fernandes F *et al.* (2017) Habitat continuity and stepping-stone oceanographic distances explain population genetic connectivity of the brown alga *Cystoseira amentacea*. *Molecular Ecology*, **26**, 766–780.

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

Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, **7**, 361–372.

Cleary KA, Waits LP, Finegan B (2017) Comparative landscape genetics of two frugivorous bats in a biological corridor undergoing agricultural intensification. *Molecular Ecology*, **26**, 4603–4617.

Cosentino BJ, Phillips CA, Schooley RL, Lowe WH, Douglas MR (2012) Linking extinction–

colonization dynamics to genetic structure in a salamander metapopulation. *Proceedings of the Royal Society of London, Series B*, **279**, 1575–1582.

Coster SS, Babbitt KJ, Cooper A, Kovach AI (2015) Limited influence of local and landscape factors on finescale gene flow in two pond-breeding amphibians. *Molecular Ecology*, **24**, 742–758.

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.

Dall'antonia P, Sinsch U (2001) In search of water: orientation behaviour of dehydrated natterjack toads, *Bufo calamita*. *Animal Behaviour*, **61**, 617–629.

Denton JS, Hitchings SP, Beebee TJC, Gent A (1997) A recovery program for the natterjack toad (*Bufo calamita*) in Britain. *Conservation Biology*, **11**, 1329–1338.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (*Ne*) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Dudaniec RY, Spear SF, Richardson JS, Storfer A (2012) Current and historical drivers of landscape genetic structure differ in core and peripheral salamander populations. *PloS ONE*, **7**, e36769.

Duforet-Frebourg N, Blum MGB (2014) Nonstationary patterns of isolation-by-distance: inferring measures of local geentic differentiation with bayesian kriging. *Evolution*, **68**, 1110–1123.

Dupont L, Torres-Leguizamon M, René-Corail P, Mathieu J (2017) Landscape features impact connectivity between soil populations: a comparative study of gene flow in earthworms. *Molecular Ecology*, **26**, 3128–3140.

Dyer RJ (2015) Is there such a thing as landscape genetics? *Molecular Ecology*, **24**, 3518–3528.

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

Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world : disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, **24**, 6021–6040.

van Etten J (2017) R Package gdistance : Distances and routes on geographical grids. *Journal of Statistical Software*, **76**, 1–30.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

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

Fahrig L (2017) Ecological responses to habitat fragmentation per se. *Annual Review of Ecology, and Systematics*, **48**, 1-23.

Faucher L, Godé C, Arnaud J-F (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo* (*Epidalea*)

calamita (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). *Journal of Heredity*, **107**, 660–665.

Faucher L, Hénocq L, Vanappelghem C *et al.* (2017) When new human-modified habitats favour the expansion of an amphibian pioneer species: evolutionary history of the natterjack toad (*Bufo calamita*) in a coal basin. *Molecular Ecology*, **26**, 4434–4451.

Favre-Bac L, Mony C, Ernoult A, Burel F, Arnaud J-F (2016) Ditch network sustains functional connectivity and influences patterns of gene flow in an intensive agricultural landscape. *Heredity*, **116**, 200–212.

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

Foley JA, Defries R, Asner GP *et al.* (2005) Global consequences of land use. *Science*, **309**, 570–574.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, New York, USA.

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

Frei M, Csencsics D, Brodbeck S *et al.* (2016) Combining landscape genetics, radio-tracking and long-term monitoring to derive management implications for Natterjack toads (*Epidalea calamita*) in agricultural landscapes. *Journal for Nature Conservation*, **32**, 22–34.

Godin J (2002) Degré de rareté, évolution de la distribution et particularités de l'herpétofaune de la région Nord - Pas de Calais. *Bulletin de la Société Herpétologique de France*, **104**, 16–35.

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.

Gubili C, Mariani S, Weckworth B V. *et al.* (2017) Environmental and anthropogenic drivers of connectivity patterns: a basis for prioritizing conservation efforts for threatened populations. *Evolutionary Applications*, **10**, 199–211.

Gutiérrez-Rodríguez J, Gonçalves J, Civantos E, Martínez-Solano I (2017) Comparative landscape genetics of pond-breeding amphibians in Mediterranean temporal wetlands: the positive role of structural heterogeneity in promoting gene flow. *Molecular Ecology*, **26**, 5407-5420.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hels T, Buchwald E (2001) The effect of road kills on amphibian populations. *Biological Conservation*, **99**, 331–340.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Janin A, Léna 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.

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.

Jourdan-Pineau H, David P, Crochet P-A (2012) Phenotypic plasticity allows the Mediterranean parsley frog *Pelodytes punctatus* to exploit two temporal niches under continuous gene flow. *Molecular Ecology*, **21**, 876–886.

Jourdan-Pineau H, Nicot A, Dupuy V, David P, Crochet P-A (2009) Development of eight microsatellite markers in the parsley frog (*Pelodytes punctatus*). *Molecular Ecology Resources*, **9**, 261–263.

Kalinowski ST (2005) HP-RARE 1.0 : a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

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

Knopp T, Merilä J (2009) Microsatellite variation and population structure of the moor frog (*Rana arvalis*) in scandinavia. *Molecular Ecology*, **18**, 2996–3005.

Koen EL, Bowman J, Walpole AA (2012) The effect of cost surface parameterization on landscape resistance estimates. *Molecular Ecology Resources*, **12**, 686–696.

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.

Lemoine G (2012) Interêt, gestion, protection et valorisation des terrils miniers dans le Nord–Pas-de-Calais. *Annales de la Société Géologique du Nord*, **19**, 1–8.

Loiselle BA, Sork VL, John N, Graham C (1995) Spatial genetic structure of a tropical understory shurb, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

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

Lowe WH, Kovach RP, Allendorf FW (2017) Population genetics and demography unitr ecology and evolution. *Trends in Ecology and Evolution*, **32**, 141–152.

McCartney-Melstad E, Shaffer HB (2015) Amphibian molecular ecology and how it has informed conservation. *Molecular Ecology*, **24**, 5084–5109.

McRae BH (2006) Isolation by resistance. *Evolution*, **60**, 1551–1561.

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19885–19890.

Micheletti SJ, Storfer A (2015) A test of the central-marginal hypothesis using population genetics and ecological niche modelling in an endemic salamander (*Ambystoma barbouri*). *Molecular Ecology*, **24**, 967–979.

Michels E, Cottenie K, Neys L *et al.* (2001) Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modelling of the effective geographical distance. *Molecular Ecology*, **10**, 1929–1938.

Mimura M, Yahara T, Faith DP *et al.* (2017) Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evolutionary Applications*, **10**, 121–139.

Miraldo A, Li S, Borregaard MK *et al.* (2016) An Anthropocene map of genetic diversity. *Science*, **353**, 1532–1535.

Phillipsen IC, Funk WC, Hoffman EA, Monsen KJ, Blouin MS (2011) Comparative analyses of effective population size within and among species: ranid frogs as a case study. *Evolution*, **65**, 2927–2945.

Pinheiro J, Bates D, DebRoy S, R Core Team (2017) Linear and nonlinear mixed effects models. *R package version 3.1-131*, https://CRAN.R-project.org/package=nlme.

Pisa G, Orioli V, Spilotros G *et al.* (2015) Detecting a hierarchical genetic population structure: the case study of the Fire Salamander (*Salamandra salamandra*) in Northern Italy. *Ecology and Evolution*, **5**, 743–758.

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

Richardson JL (2012) Divergent landscape effects on population connectivity in two co-occurring amphibian species. *Molecular Ecology*, **21**, 4437–4451.

Richardson JL, Brady SP, Wang IJ, Spear SF (2016) Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology*, **25**, 849–863.

Rogell B, Gyllenstrand N, Höglund J (2005) Six polymorphic microsatellite loci in the Natterjack toad, *Bufo calamita*. *Molecular Ecology Notes*, **5**, 639–640.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

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

Row JR, Blouin-Demers G, Lougheed SC (2010) Habitat distribution influences dispersal and fine-scale genetic population structure of eastern foxsnakes (*Mintonius gloydi*) across a fragmented landscape. *Molecular Ecology*, **19**, 5157–5151.

Rowe G, Beebee TJC, Burke T (1997) PCR primers for polymorphic microsatellite loci in the anuran amphibian *Bufo calamita*. *Molecular Ecology*, **6**, 401–402.

Rowe G, Beebee TJC, Burke T (2000a) A further four polymorphic microsatellite loci in the natterjack toad *Bufo calamita*. *Conservation Genetics*, **1**, 371–372.

Rowe G, Beebee TJC, Burke T (2000b) A microsatellite analysis of natterjack toad, *Bufo calamita*, metapopulations. *Oikos*, **88**, 641–651.

Séré M, Thévenon S, Belem AMG, De Meeûs T (2017) Comparison of different genetic distances to test isolation by distance between populations. *Heredity*, **119**, 55–63.

Serrouya R, Paetkau D, McLellan BN *et al.* (2012) Population size and major valleys explain microsatellite variation better than taxonomic units for caribou in western Canada. *Molecular*

Ecology, **21**, 2588–2601.

Sinsch U (1992) Sex-biassed site fidelity and orientation behaviour in reproductive natterjack toads (*Bufo calamita*). *Ethology Ecology & Evolution*, **4**, 15–32.

Sinsch U (1997) Postmetamorphic dispersal and recruitment of first breeders in a *Bufo calamita* metapopulation. *Oecologia*, **112**, 42–47.

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

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

Sokal RR, Jacquez GM, Wooten MC (1989) Spatial autocorrelation analysis of migration and selection. *Genetics*, **121**, 845–855.

Sokal RR, Wartenberg DE (1983) A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics*, **105**, 219–237.

Stevens VM, Baguette M (2008) Importance of habitat quality and landscape connectivity for the persistence of endangered natterjack toads. *Conservation Biology*, **22**, 1194–204.

Stevens VM, Leboulengé 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, 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, Verkenne C, Vandewoestijne S, Wesselingh RA, Baguette M (2006b) Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology*, **15**, 2333–2344.

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

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, **144**, 389–399.

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

Villemey A, Peterman WE, Richard M *et al.* (2016) Butterfly dispersal in farmland: a replicated landscape genetics study on the meadow brown butterfly (*Maniola jurtina*). *Landscape Ecology*, **31**, 1629–1641.

Van De Vliet MS, Diekmann OE, Serrão ETA, Beja P (2009) Highly polymorphic microsatellite loci for the Parsley frog (*Pelodytes punctatus*): characterization and testing for cross-species amplification. *Conservation Genetics*, **10**, 665–668.

Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJ (2001) Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity*, **86**, 598–608.

Wang IJ (2012) Environmental and topographic variables shape genetic structure and effective

population sizes in the endangered Yosemite toad. *Diversity and Distributions*, **18**, 1033–1041.

Wang J (2017) The computer program STRUCTURE for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources*, **17**, 981–990.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167–184.

Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution*, **28**, 78–85.

Weckworth B V., Musiani M, DeCesare NJ *et al.* (2013) Preferred habitat and effective population size drive landscape genetic patterns in an endangered species. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20131756.

Wilkinson JW, Griffiths RA (2013) Amphibian declines and conservation in Britain. In: *Amphibian biology. Status of conservation and decline of amphibians: Eastern hemisphere. Western Europe* (eds Heatwole H, Wilkinson JW), pp. 48–57. Pelagic Publishing, Exeter, UK.

Wright S (1931) Evolution in Mendelian Populations. *Genetics*, **16**, 97–159.

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

Youngquist MB, Inoue K, Berg DJ, Boone MD (2017) Effects of land use on population presence and genetic structure of an amphibian in an agricultural landscape. *Landscape Ecology*, **32**, 147–162

Supplementary files

Table S1: Mean estimates of nuclear genetic variation for 37 nuclear microsatellite loci over 49 populations of *Bufo [Epidalea] calamita* and for 22 loci over 16 populations of *Pelodytes punctatus* from northern France, with a minimum sampling size of seven individuals per population. Total number of alleles (A_T), mean number of alleles per populations (A_n), allelic richness (A_r), observed heterozygosity (H_O), expected heterozygosity (H_E) and mean F_{IS} and F_{ST} fixation indices, according to Weir & Cockerham (1984).* $P < 0.05$; *** $P < 0.001$.

Species	Locus	Reference*	Multiplex	Dye	F_{IS}	F_{ST}	A_T	A_R	A_n	H_O	H_E
P. punctatus	PPU10	1	1	FAM	0.015	0.282 ***	3	2.488	2.562	0.482	0.502
P. punctatus	PPU16	1	1	PET	0.019	0.351 ***	2	1.405	1.437	0.178	0.171
P. punctatus	PPU2	1	1	VIC	–	–	1	1	1	0	0
P. punctatus	PPU5	1	1	NED	0.169 *	0.230 ***	2	1.519	1.625	0.151	0.171
P. punctatus	PPU15	1	2	NED	–	–	1	1	1	0	0
P. punctatus	PPU17	1	2	VIC	–	–	1	1	1	0	0
P. punctatus	PPU11	1	2	PET	–	–	1	1	1	0	0
P. punctatus	ppu5	2	3	VIC	-0.019	0.355 ***	7	3.084	3.500	0.509	0.522
P. punctatus	ppu7	2	3	NED	-0.050	0.211 ***	16	4.635	5.500	0.758	0.710
P. punctatus	ppu13	2	3	FAM	0.010	0.269 ***	10	3.432	3.875	0.577	0.569
P. punctatus	ppu15	2	3		–	–	1	1	1	0	0
P. punctatus	ppu3	2	3	FAM	-0.064	0.240 ***	17	4.991	6.0625	0.719	0.689
P. punctatus	ppu9	2	3	PET	-0.037	0.192 ***	16	4.469	5.0625	0.736	0.710
P. punctatus	ppu10	2	4	NED	–	–	1	1	1	0	0
P. punctatus	ppu11	2	4	VIC	0.042	0.333 ***	11	3.476	3.9375	0.601	0.598
P. punctatus	ppu14	2	4	PET	0.003	0.147 ***	10	4.158	4.625	0.689	0.709
P. punctatus	ppu2	2	4	NED	0.064	0.264 ***	13	3.022	2.437	0.507	0.537
P. punctatus	ppu6	2	4	FAM	0.007	0.364 ***	12	3.462	3.875	0.584	0.591
P. punctatus	ppu8	2	4	FAM	-0.033	0.355 ***	5	2.700	2.937	0.524	0.521
P. punctatus	ppu1	2	5	FAM	–	–	1	1	1	0	0
P. punctatus	ppu12	2	5	PET	0.126 ***	0.237 ***	11	3.344	3.625	0.543	0.580
P. punctatus	ppu4	2	5	VIC	-0.002	0.242 ***	11	3.600	3.875	0.638	0.637
P. punctatus	Mean multilocus:				0.004	0.268 ***	6.955	2.581	2.815	0.373	0.374

B. calamita	Bcal7	3	1	NED	-0.022	0.155	5	2.245	2.551	0.334	0.339
B. calamita	Bcal1	3	1	FAM	-0.049	0.113	16	5.311	6.286	0.808	0.774
B. calamita	Bcal2	3	1	PET	0.092	0.131	6	2.672	2.939	0.413	0.459
B. calamita	Bcal11	3	2	FAM	0.003	0.134	10	2.832	3.306	0.438	0.436
B. calamita	Buca2	4	2	VIC	0.009	0.087	5	3.425	3.755	0.637	0.647
B. calamita	Buca3	4	2	PET	—	—	1	1.000	1.000	0.000	0.000
B. calamita	Buca5	4	2	NED	0.087	0.085	5	1.522	1.714	0.119	0.125
B. calamita	Bcal3	4	3	FAM	0.097	0.137	11	3.812	4.592	0.587	0.626
B. calamita	Bcal4	4	3	PET	-0.018	0.140	8	2.801	3.102	0.556	0.545
B. calamita	Bcal5	4	3	FAM	-0.039	0.185	5	2.683	3.082	0.441	0.427
B. calamita	Bcal6	4	3	VIC	-0.026	0.100	9	3.744	4.122	0.662	0.661
B. calamita	Bcal10	5	4	NED	-0.023	0.129	9	3.397	3.816	0.636	0.633
B. calamita	Bcal9	5	4	FAM	0.342	0.242	9	2.946	3.367	0.292	0.448
B. calamita	Buca6	4	4	FAM	0.108	0.068	6	2.211	2.510	0.288	0.320
B. calamita	Buca1	4	5	NED	-0.015	0.074	9	3.708	4.184	0.659	0.666
B. calamita	BC05	6	6	NED	-0.003	0.182	5	2.681	2.918	0.484	0.490
B. calamita	BC08	6	6	HEX	0.007	0.078	17	4.105	5.000	0.657	0.656
B. calamita	BC11	6	6	HEX	-0.040	0.089	10	4.993	5.959	0.771	0.759
B. calamita	BC19	6	6	FAM	-0.002	0.139	6	2.424	2.714	0.377	0.383
B. calamita	BC22	6	6	FAM	0.216	0.138	7	3.483	3.939	0.496	0.624
B. calamita	BC29	6	6	NED	0.007	0.107	8	3.351	3.694	0.590	0.591
B. calamita	BC46	6	6	FAM	-0.021	0.119	8	3.235	3.837	0.569	0.554
B. calamita	BC38	6	6	PET	—	—	1	1.000	1.000	0.000	0.000
B. calamita	BC01	6	7	FAM	-0.025	0.096	10	3.094	3.633	0.510	0.495
B. calamita	BC02	6	7	FAM	0.004	0.092	11	4.384	5.143	0.684	0.696
B. calamita	BC04	6	7	NED	-0.062	0.147	6	3.033	3.490	0.512	0.488
B. calamita	BC24	6	7	NED	-0.010	0.079	4	2.904	3.041	0.593	0.597
B. calamita	BC25	6	7	PET	-0.083	0.030	3	2.051	2.204	0.305	0.286

B. calamita	BC28	6	7	HEX	-0.184	0.157	3	2.733	2.878	0.610	0.523
B. calamita	BC34	6	7	HEX	0.011	0.071	4	2.205	2.265	0.465	0.449
B. calamita	BC35	6	7	PET	-0.074	0.134	3	1.455	1.571	0.118	0.114
B. calamita	BC15	6	8	FAM	0.029	0.117	7	1.773	2.020	0.152	0.156
B. calamita	BC18	6	8	FAM	-0.034	0.099	8	4.217	4.816	0.717	0.701
B. calamita	BC37	6	8	HEX	-0.016	0.088	9	4.346	4.959	0.732	0.723
B. calamita	BC39	6	8	HEX	-0.025	0.083	9	3.949	4.633	0.675	0.663
B. calamita	BC45	6	8	PET	-0.017	0.097	10	4.977	5.939	0.761	0.755
B. calamita	BC09	6	8	NED	-0.017	0.107	10	4.262	5.041	0.674	0.681
B. calamita	Mean multilocus				0.002	0.114	/				

References

1 Jourdan-Pineau H, Nicot A, Dupuy V, David P, Crochet P-A (2009) Development of eight microsatellite markers in the parsley frog (*Pelodytes punctatus*). Molecular Ecology Resources, 9, 261–263.

2 Van De Vliet MS, Diekmann OE, Serrão ETA, Beja P (2009) Highly polymorphic microsatellite loci for the Parsley frog (*Pelodytes punctatus*): Characterization and testing for cross-species amplification. Conservation Genetics, 10, 665–668.

3 Rowe G, Beebee TJC, Burke T. (1997) PCR primers for polymorphic microsatellite loci in the anuran amphibian *Bufo calamita*. Molecular Ecology, 6, 401–402.

4 Rogell B, Gyllenstrand N, Höglund J. (2005) Six polymorphic microsatellite loci in the Natterjack toad, *Bufo calamita*. Molecular Ecology Notes, 5, 639–640.

5 Rowe G, Beebee TJC, Burke T. (2000) A further four polymorphic microsatellite loci in the natterjack toad *Bufo calamita*. Conservation Genetics, 1, 371–372.

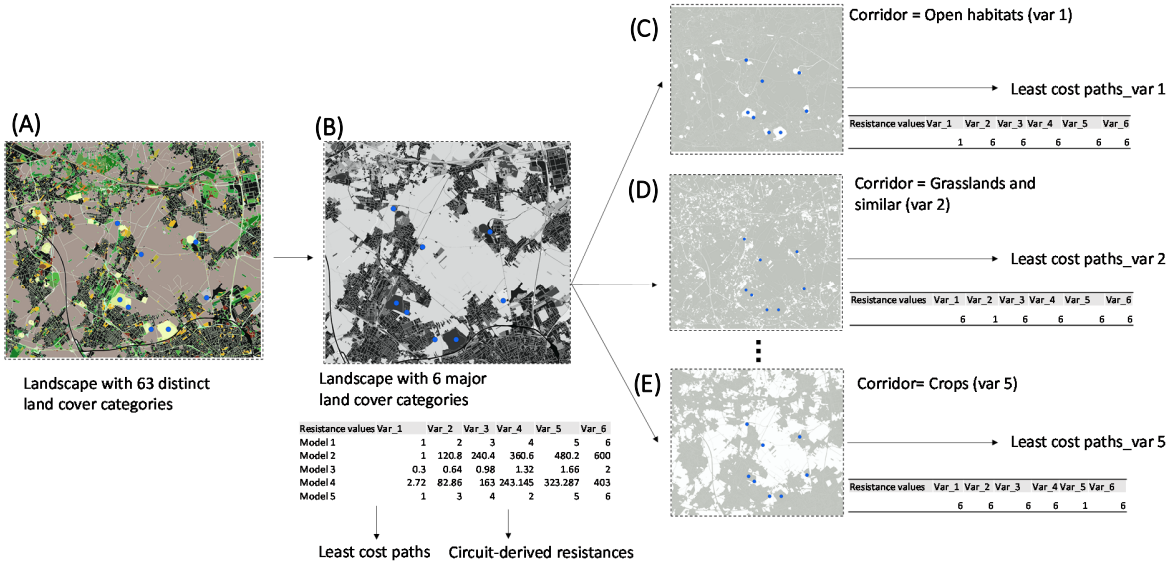
6 Faucher L, Godé C, Arnaud J-F. (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo* (*Epidalea*) *calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). Journal of Heredity, 107, 660–665.

Table S2. Estimates of nuclear genetic variation in 28 populations of *Pelodytes punctatus* and in 53 populations of *Bufo* [*Epidalea*] *calamita* sampled in northern France. Sample size (N), total number of alleles (A_T), the nuclear allelic richness (A_r), the observed heterozygosity (H_O), expected heterozygosity (H_E) and the mean intra-population fixation index (F_{IS}) according to Weir & Cockerham (1984). Also reported is the one-sample effective population size (N_e) estimates using the linkage disequilibrium method with jackknife 95% CIs estimated by jacknifing. Landscape analysis: specific populations used in landscape analysis (LCP: for least cost path estimates only, CS: for least cost path and circuit theory based ecological distance estimates). –: Not applicable (insufficient sampling size). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. X : longitude, Y : latitude (decimal degrees).

Species	Code	Pop	X	Y	N	A_T	A_r	H_O	H_E	F_{IS}	N_e	95%Cis Inf	95%Cis Sup	Landscape analysis
P. punctatus	Coast-1	CEN17-S2	1.784894	50.95188	2	26	–	0.4	0.4	–	–	–	–	
P. punctatus	Coast-2	CEN61-S1	1.780561	50.83092	5	32	–	0.386	0.379	–	–	–	–	
P. punctatus	Coast-3	CEN43-S4	1.603657	50.59935	4	43	–	0.571	0.551	–	–	–	–	
P. punctatus	Coast-4	CEN42-S1	1.622411	50.58334	9	59	3.87	0.631	0.6	-0.056	101.8	19.1	Infinite	
P. punctatus	Coast-5	CEN51-S2	1.575851	50.45408	3	32	–	0.444	0.418	–	–	–	–	
P. punctatus	Inl-1	GON52-S1	2.185514	50.24574	11	60	5.19	0.597	0.568	-0.054	19.5	6.7	Infinite	
P. punctatus	Inl-2	T24-S1	2.463206	50.5192	1	21	–	0.5	0.5	–	–	–	–	CS
P. punctatus	Inl-3	Bell-S1	2.698919	50.51707	10	65	4.02	0.638	0.656	0.03	35.4	15.6	Infinite	CS
P. punctatus	Inl-4	T13-S1	2.49169	50.51254	4	35	–	0.554	0.485	–	–	–	–	CS
P. punctatus	Inl-5	T12-S2	2.556442	50.50626	11	49	4.65	0.503	0.506	0.006	17.5	6.1	Infinite	CS
P. punctatus	Inl-6	T70b-S1	2.849757	50.50551	10	59	5.33	0.571	0.602	0.054	36	14	Infinite	CS
P. punctatus	Inl-7	T70b-S2	2.84193	50.50501	4	49	–	0.571	0.63	–	–	–	–	CS
P. punctatus	Inl-8	GON50-S1	2.737992	50.50185	8	52	3.38	0.589	0.599	0.018	18.7	7.1	Infinite	CS
P. punctatus	Inl-9	T73b-S2	2.801566	50.49368	24	41	2.38	0.519	0.413	-0.274	5.5	1.6	37.2	CS
P. punctatus	Inl-10	T10-S1	2.521788	50.48576	12	56	4.71	0.511	0.537	0.05	19.9	10.2	64.3	CS
P. punctatus	Inl-11	T58-S1	2.739983	50.46149	1	22	–	0.571	0.571	–	–	–	–	CS
P. punctatus	Inl-12	T58A	2.747008	50.46066	39	58	2.82	0.37	0.427	0.135	9.4	4.5	21.4	CS
P. punctatus	Inl-13	CPIE4-S1	2.795205	50.40722	4	44	–	0.628	0.587	–	–	–	–	CS
P. punctatus	Inl-14	T75-S2	2.79666	50.40376	14	55	4.92	0.546	0.546	0	12.2	3.5	40.7	CS
P. punctatus	Inl-15	T158-S1	3.410185	50.35857	9	43	–	0.504	0.483	-0.045	40.7	12.5	Infinite	CS
P. punctatus	Inl-16	GON1-S2	2.881007	50.24963	2	40	–	0.733	0.7	–	–	–	–	CS
P. punctatus	Inl-17	CG59-2-S2	3.238705	50.34541	20	57	3.16	0.527	0.536	0.018	Infinite	16.1	Infinite	CS
P. punctatus	Inl-18	CG59-2-S1	3.233922	50.34533	15	55	3.16	0.569	0.534	-0.067	53.7	21.2	Infinite	CS
P. punctatus	Inl-19	CPIE44-S2	3.236026	50.34499	9	47	2.96	0.557	0.539	-0.032	12.2	3.4	188.3	CS

P. punctatus	Inl-20	CPIE44-S1	3.238849	50.34565	17	55	3.12	0.408	0.535	0.254***	5.7	2.9	9.8	CS
P. punctatus	Inl-21	CPIE1-S1	3.289887	50.34252	12	68	4	0.704	0.679	-0.039	17.7	10	43.4	CS
P. punctatus	Inl-22	CPIE14-S1	3.204092	50.31675	6	69	—	0.598	0.731	—	—	—	—	CS
P. punctatus	Inl-23	Tmas-S1	3.286264	50.30792	1	22	—	0.571	0.571	—	—	—	—	CS
B. calamita	Coast-01	CEN2-S1	2.5398	51.0821	9	118	3.19	0.471	0.523	0.104	11.4	8.2	16.6	
B. calamita	Coast-02	CEN03-S2	2.504	51.0728	17	129	3.12	0.525	0.504	-0.043	53	30.7	149	
B. calamita	Coast-03	CEN03-S1	2.5044	51.0686	8	107	2.97	0.468	0.467	-0.001	Infinite	39.6	Infinite	
B. calamita	Coast-04	CEN59-S1	2.4501	50.976	2	63	—	—	—	—	—	—	—	
B. calamita	Coast-05	GS-Bc	2.2778	51.0199	29	133	2.9	0.473	0.521	0.094	46.9	33.3	73.8	
B. calamita	Coast-06	CEN09-S2	2.1699	50.9883	13	134	3.31	0.51	0.504	-0.013	49.2	25.3	276.2	
B. calamita	Coast-07	CEN09-S1	2.1744	50.9877	10	120	3.16	0.534	0.501	-0.071	Infinite	79.4	Infinite	
B. calamita	Coast-08	CEN63-Bc	1.954	50.983	9	122	3.07	0.509	0.507	-0.005	Infinite	Infinite	Infinite	
B. calamita	Coast-09	CEN17-S1	1.7849	50.9519	10	117	3.14	0.489	0.497	0.017	Infinite	67.3	Infinite	
B. calamita	Coast-10	CEN60-S1	1.6803	50.8945	4	83	—	—	—	—	—	—	—	
B. calamita	Coast-11	CEN28-S1	1.6292	50.7968	12	119	3.12	0.486	0.514	0.046	56.9	23.4	Infinite	
B. calamita	Coast-12	CEN34-S1	1.5788	50.649	11	111	2.91	0.423	0.443	0.047	21.6	12.4	53.2	CS
B. calamita	Coast-13	CEN43-S2	1.5876	50.5991	18	129	3.03	0.465	0.457	-0.02	12.5	9.8	16.2	CS
B. calamita	Coast-14	CEN43-S1	1.5987	50.5984	10	125	3.27	0.483	0.486	0.008	Infinite	90.4	Infinite	CS
B. calamita	Coast-15	CEN43-S3	1.5785	50.5968	8	121	3.33	0.443	0.471	0.065	560.4	52.5	Infinite	CS
B. calamita	Coast-16	CEN42-S1	1.6237	50.5829	21	109	2.64	0.45	0.412	-0.093	8.6	6.5	11.5	CS
B. calamita	Coast-17	CEN42-S2	1.6244	50.5823	16	102	2.5	0.409	0.387	-0.057	28	15.9	73.4	CS
B. calamita	Coast-18	CEN48-S1	1.5832	50.4869	20	145	3.3	0.499	0.485	-0.029	198.7	73.8	Infinite	CS
B. calamita	Coast-19	CEN51-S1	1.5758	50.456	13	138	3.35	0.479	0.486	0.015	65.6	30.2	Infinite	CS
B. calamita	Coast-20	CEN51-M13	1.5889	50.4398	11	124	3.32	0.463	0.455	-0.021	430.2	34.4	Infinite	CS
B. calamita	Coast-21	CEN51-P16	1.5758	50.435	22	157	3.15	0.516	0.498	-0.037	316.1	81.3	Infinite	CS
B. calamita	MA-01	GON52-S1	2.3279	50.5772	22	104	2.71	0.439	0.455	0.035	13.1	10	17.5	
B. calamita	MA-02	GON52-S2	2.3291	50.5788	15	102	3.25	0.493	0.474	-0.039	6.9	4.9	9.4	
B. calamita	MA-03	T1-S4	2.5198	50.4786	27	127	2.93	0.544	0.528	-0.031	20.2	16.2	25.8	CS
B. calamita	MA-04	T10-S1	2.5218	50.4858	2	67	—	—	—	—	—	—	—	CS
B. calamita	MA-05	T2-S1	2.5675	50.4591	25	107	3.11	0.47	0.455	-0.032	22.3	16.2	32.9	CS
B. calamita	MA-06	GON50-S1	2.738	50.5018	11	130	2.66	0.506	0.516	0.019	63.7	28.1	Infinite	CS
B. calamita	MA-07	T58-S1	2.74	50.4615	34	140	4	0.518	0.534	0.034	36.7	28.9	48.4	CS

B. calamita	MA-08	T58A-S2	2.7469	50.4576	25	140	4	0.533	0.533	-0.002	18.2	15.1	22.4	CS
B. calamita	MA-09	GON53-S1	2.7534	50.2618	10	122	3.14	0.513	0.512	-0.002	Infinite	76.2	Infinite	
B. calamita	MA-10	T48-S1	2.7574	50.4856	10	126	3.6	0.531	0.521	-0.021	22.2	15.4	36.4	CS
B. calamita	MA-11	T59-S1	2.7656	50.4461	33	129	3.686	0.486	0.485	-5.00E-04	18.3	15.1	22.3	CS
B. calamita	MA-12	T74-S1	2.7797	50.4461	36	125	3.571	0.507	0.49	-0.035	57.9	40.7	92.9	CS
B. calamita	MA-13	T75-S2	2.7967	50.4038	30	130	3.714	0.523	0.501	-0.044	50.9	36.5	79.3	CS
B. calamita	MA-14	CPIE7-S1	2.8022	50.492	52	134	3.39	0.504	0.5	-0.008	96.1	64.1	172.4	CS
B. calamita	MA-15	CPIE5-S1	2.811	50.4625	32	111	3.11	0.464	0.444	-0.048	49.4	33.5	83.7	CS
B. calamita	MA-16	T110-S1	2.9823	50.4583	6	95	–	–	–	–	–	–	–	LCP
B. calamita	MA-17	Monch-Bc	3.2041	50.3168	50	135	3.21	0.494	0.484	-0.026	29	24.5	34.7	CS
B. calamita	MA-18	CPIE44-S3	3.2332	50.3449	18	136	3.06	0.544	0.552	0.015	50.2	33.1	94.3	CS
B. calamita	MA-19	CPIE44-S1	3.2371	50.3447	13	128	3.38	0.497	0.533	0.069	103.4	35.4	Infinite	CS
B. calamita	MA-20	CPIE44-S2	3.2381	50.3453	21	142	3.35	0.554	0.553	-0.002	16.6	13.2	21.4	CS
B. calamita	MA-21	CPIE43-S1	3.2409	50.3827	38	142	3.27	0.545	0.529	-0.031	51.2	38.6	72.3	CS
B. calamita	MA-22	CPIE43-S2	3.2457	50.3751	17	127	3.15	0.506	0.508	0.003	75.8	39.1	474.4	CS
B. calamita	MA-23	CPIE1-S1	3.2899	50.3425	40	118	2.9	0.561	0.504	-0.116	39.3	29.3	56	CS
B. calamita	MA-24	CPIE3-S1	3.3054	50.3581	7	109	3.31	0.567	0.513	-0.117	Infinite	33.2	Infinite	CS
B. calamita	MA-25	CPIE3-S2	3.3059	50.3575	8	119	3.34	0.529	0.532	0.008	Infinite	64.3	Infinite	CS
B. calamita	MA-26	CPIE3-S3	3.306	50.3584	13	130	3.34	0.549	0.554	0.008	58.1	30.3	314.8	CS
B. calamita	MA-27	CPIE3-S4	3.3073	50.3585	7	117	3.24	0.527	0.539	0.026	Infinite	40.1	Infinite	CS
B. calamita	MA-28	CPIE2-S1	3.519	50.4135	9	105	2.85	0.517	0.501	-0.034	366.8	35.1	Infinite	LCP
B. calamita	MA-29	CPIE2-S2	3.5193	50.413	16	111	2.86	0.498	0.497	-0.001	54.7	27.8	333.5	LCP
B. calamita	MA-30	CPIE2-S3	3.5208	50.4133	14	108	3.11	0.516	0.5	-0.034	65.4	30.4	Infinite	LCP
B. calamita	MA-31	CPIE4-S1	3.567	50.4044	10	115	3.01	0.503	0.501	-0.005	Infinite	Infinite	Infinite	LCP
B. calamita	MA-32	CPIE4-S2	3.5706	50.4076	34	133	2.74	0.519	0.517	-0.003	83.7	52.3	179.3	LCP



Resistance values	Var_1	Var_2	Var_3	Var_4	Var_5	Var_6
Model 1	1	2	3	4	5	6
Model 2	1	120.8	240.4	360.6	480.2	600
Model 3	0.3	0.64	0.98	1.32	1.66	2
Model 4	2.72	82.86	163	243.145	323.287	403
Model 5	1	3	4	2	5	6

Figure S1: Flow-chart of the parameterization of resistance surfaces and the calculation of least-cost path distances and circuit theory derived distances. Each different color indicats a distinct resistance. A) shows the original land cover map description, B) shows the simplified map depicting the six landscape features analysed in this study: open habitats, grasslands, roads, forests, crops and urban, highways areas. C, D and E are examples of maps designed on the landscape models specific to each landscape features and further use to measure least cost path and test the likelihood of each specifi landscape to act as corridors. Blue dots indicated Natterjack toad populations. This example shows a small portion of the western coal basin analysed in this study.

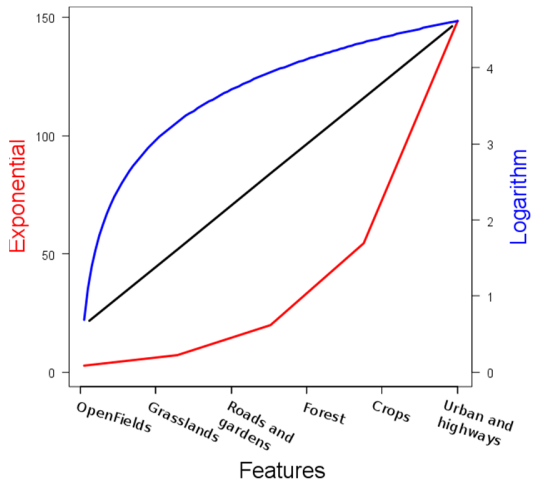


Figure S2: Resistance values given for the six landscape categories following an exponential or a logarithmic transformation.

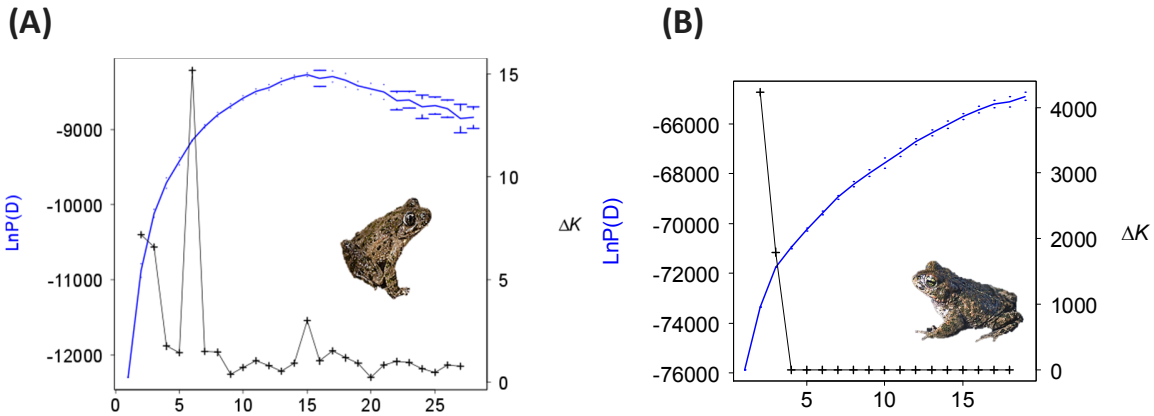
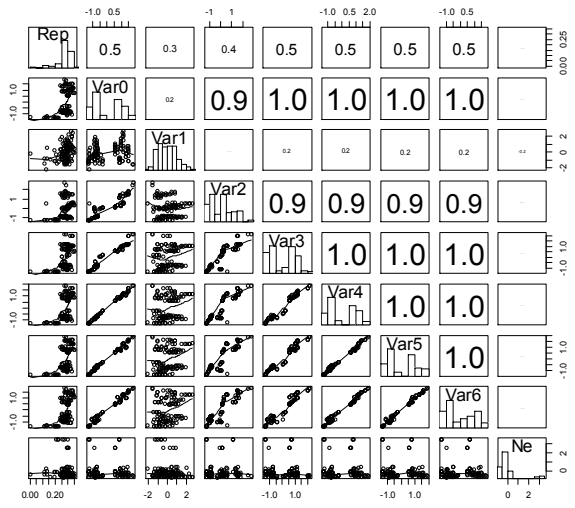
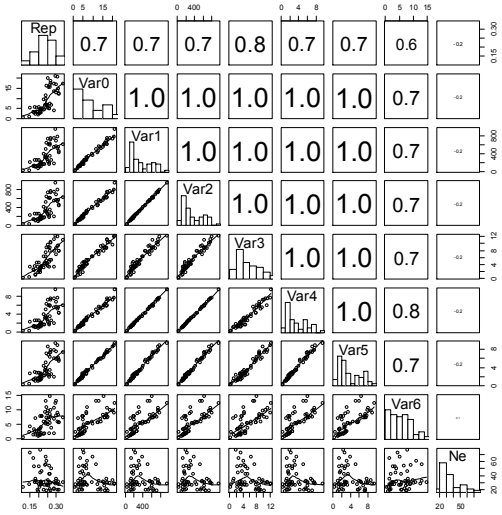


Figure S3: Ln(P(X|*K*)) and the *ad hoc* statistic Δ*K* (Evanno, *et al.* 2005) obtained over 30 replicated runs and plotted against the putative number of *K* clusters using: (A) the whole dataset of *Pelodytes punctatus* populations, (B) the whole dataset of *Bufo calamita* populations sampled in the studied area located in northern France.

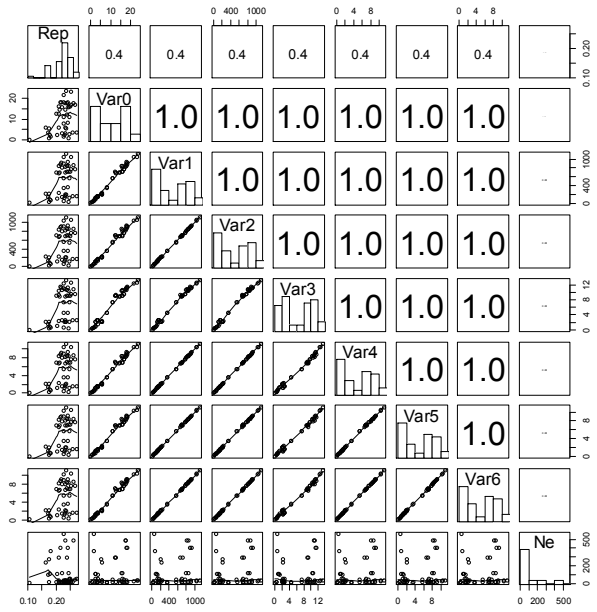
Eastern coalfield populations



Western coalfield populations



Coastline populations



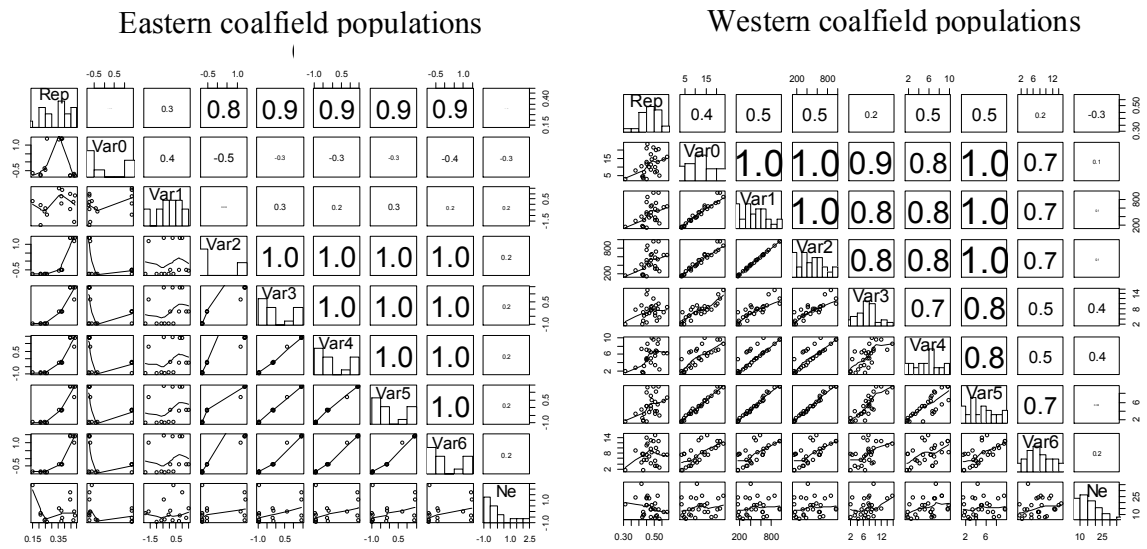


Figure S4: Graphs illustrating the distributions of centered reduced data (diagonal) and collinearity of variables (dot chart in the lower part of the matrix), and Pearson's correlation coefficients (in the upper part of the matrix) of explanatory variables used in linear models fitted by generalized least square regressions. Rep: genetic distance (D_{CE}) between population pairs, Var 0: geographical distance, Var 1 to Var 6: resistance estimated by LCP for landscape models specific to each landscape categories

CHAPITRE IV

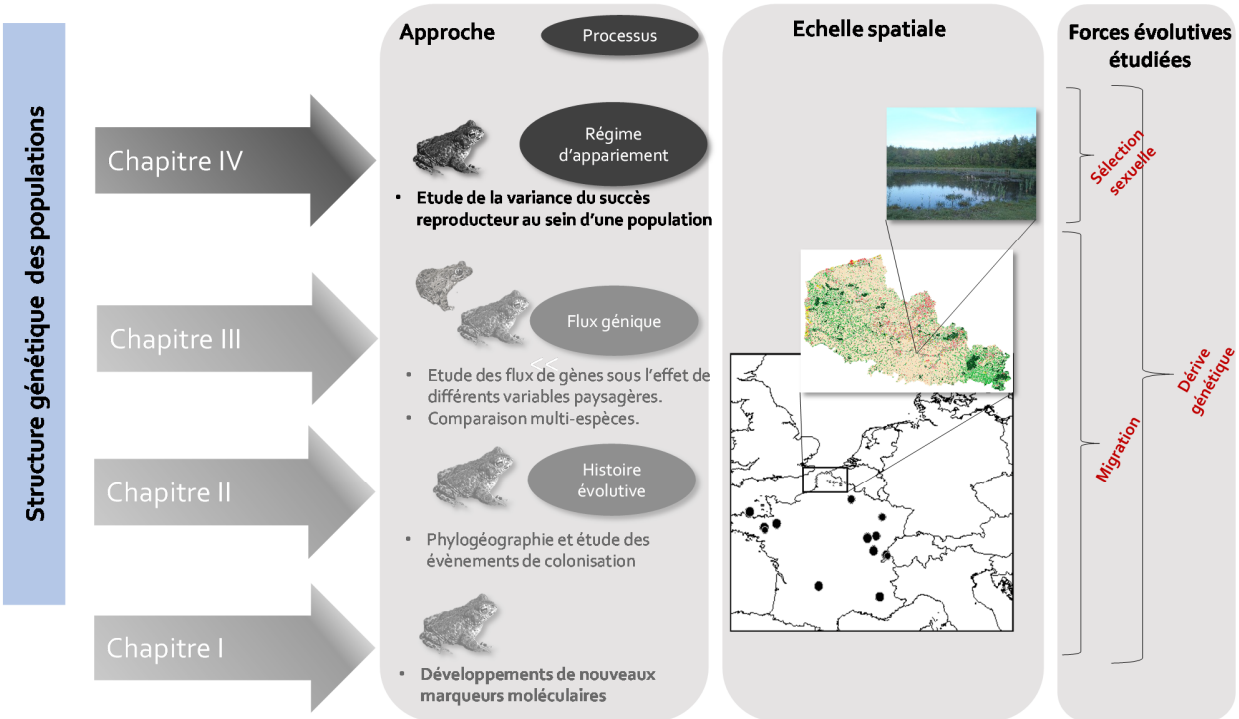
Analyse du succès reproducteur en population naturelle

"It is surprising that frogs and toads should not have acquired more strongly-marked sexual differences; for though cold-blooded, their passions are strong."
Charles Darwin, *The Descent of Man* (1871)

© Christophe Courteau

Photgraphie de crapauds calamites mâles en saison de reproduction

Une dernière partie de ce travail de thèse a porté sur l'étude du fonctionnement intra-population pour un site de reproduction chez le Crapaud calamite. En effet, au-delà de l'étude régionale de la structure génétique actuelle des populations, il est nécessaire de comprendre le fonctionnement démographique et génétique au sein d'une population bien définie géographiquement. Savoir notamment comment les évènements de reproduction s'établissent et quelles incidences ils ont sur la taille efficace d'une population sont essentiels en biologie de la conservation. Le Crapaud calamite constitue un excellent modèle d'étude dans ce cadre car l'espèce est facilement observable en période de reproduction, rendant possible un suivi complet des adultes et de leur progéniture. De plus, les évènements de reproduction ne semblent pas aléatoires mais sujets à une sélection sexuelle comme en attestent plusieurs études sur le sujet (Arak 1983, 1988; Tejedo 1988, 1992).



Chez le Crapaud calamite, comme chez beaucoup d'anoures se reproduisant autour de mares, durant la saison de reproduction, les mâles se regroupent pour former des chœurs sur les sites de reproduction tandis que les femelles réceptives ne viennent que ponctuellement pour s'y accoupler (Wells 1977). On parle de système en lek. Plus la saison de reproduction est longue, plus les visites des femelles sont éparses. Or, dans une telle population où l'un des sexes est sous-représenté par rapport à l'autre, la sélection sexuelle favorise la compétition intrasexuelle pour l'accès à la reproduction chez le sexe le plus représenté (dans notre modèle il s'agira des mâles) et un choix de partenaire chez le sexe limitant (Kokko *et al.* 2003). Le choix des femelles est un processus mis en évidence chez plusieurs taxons dont les anoures, et qui résulte de processus évolutifs complexes (Trivers 1985; Clutton-Brock 2007). Le bénéfice obtenu par la femelle au travers du choix de partenaire peut être direct (*i.e.* ressources et soins parentaux) ou indirect, à travers la production de descendants plus viables ou plus attrayants pour les autres femelles (effet dit « good genes »). Dans un système en lek, les traits mâles favorisant l'accès à la reproduction peuvent alors être rapidement avantagés et plus l'avantage qu'ils confèrent sera important plus le succès reproducteur sera biaisé en faveur des quelques mâles porteurs de ces traits (Clutton-Brock 2007). La pression de sélection sexuelle est d'autant plus forte que la variance du nombre d'accouplement entre les mâles qu'elle induit est forte (Arnold & Wade 1984). Dans les cas extrêmes, la reproduction n'est assurée que par une faible fraction des mâles et cela peut s'accompagner d'une forte perte de diversité génétique (Wright 1931).

L'objectif de ce chapitre de thèse était donc d'étudier l'impact d'un fonctionnement en lek, d'une part en mesurant le succès reproducteur des individus, et d'autre part en quantifiant l'intensité de la sélection sexuelle opérant sur les mâles et les femelles dans le cas de polygynie (accouplement d'un mâle avec plusieurs femelles) et, éventuellement, de polyandrie (accouplement d'une femelle avec plusieurs mâles) séquentielle. Cette dernière étude s'inscrit dans le double cadre de l'écologie comportementale et de la génétique et démographie des populations, ceci dans un contexte directement appliqué à la gestion et à la conservation d'espèces menacées. Nos hypothèses étaient que le comportement de regroupement en lek s'accompagne de stratégies de reproduction issues des pressions de sélection sexuelle donnant l'avantage à certains mâles plus que d'autres pour l'accès à la reproduction, par le biais d'une compétition entre mâle, et/ou par le biais d'un choix du partenaire par les femelles.

Pour tester cela, un suivi a été effectué sur toute une saison de reproduction en relevant les traits propres à chaque adulte présent, et les traits propre à la descendance que nous avons élevée en laboratoire. Ce suivi visait à estimer le succès reproducteur des adultes en prenant en compte plusieurs composantes de la valeur sélective : les tailles de pontes, les taux d'éclosion, les taux de mortalité, les poids des jeunes et leurs vitesses de développement jusqu'à la métamorphose. Ces différentes mesures nous ont permis d'estimer la variance du succès reproducteur chez les deux sexes au sein de la population et de mesurer unindice d'opportunité pour la sélection sexuelle. Grâce aux mesures prises à la fois sur les adultes et sur les descendants, nous avons pu tester si certains traits des adultes expliquaient leur succès reproducteur par le biais de la compétition entre mâles ou d'un choix femelle.

Le suivi de reproduction a duré 3 mois, d'Avril 2015 à Juin 2015. Tous les adultes ont été mesurés, génotypés sur 37 loci microsatellites et suivis sur l'ensemble de la période de reproduction grâce à des marquages par micro-puces sous-cutanées. L'ensemble des pontes a été échantillonné pour être génotypé et élevé en laboratoire. Les pontes ont été assignées à leurs deux parents par test d'assignation de parenté en se fondant sur le génotypage de 14 marqueurs microsatellites. 71 mâles et 54 femelles ont été suivis sur 44 nuits prospectées. 42% des mâles et 62% des femelles se sont reproduits au moins une fois. La distribution du succès reproducteur mâle ne diffèrait pas significativement d'une distribution de Poisson que l'on pourrait attendre sous l'hypothèse d'appariements aléatoires. Cela dit, la variance de ce succès reproducteur était suffisamment forte pour supposer qu'il puisse y avoir un effet de la sélection sexuelle sur la valeur sélective des individus. Certains traits des mâles peuvent ainsi expliquer en partie le nombre d'accouplements obtenus, ou encore la valeur sélective des pontes. Les traits femelles ont aussi un effet sur la valeur sélective des pontes. Une homogamie de taille a également été observée. Par ailleurs nos analyses indiquent que le niveau de consanguinité des femelles affecte en partie la valeur sélective des descendants. Les différences de valeurs sélectives chez les mâles et chez les descendants sont discutées, ainsi que l'observation de cette homogamie de taille, en prenant en compte les effets potentiels des processus de sélection sexuelle et de sélection naturelle sur ces patrons de reproduction.

Références

Arak A (1983) Sexual selection by male–male competition in natterjack toad choruses. *Nature*, **306**, 261–262.

Arak A (1988) Female mate selection in the natterjack toad : active choice or passive attraction ? *Behavioral Ecology and Sociobiology*, **22**, 317–327.

Arnold SJ, Wade MJ. (1984) On the measurement of natural and sexual selection: theory. *Evolution*, **38**, 709–719.

Clutton-Brock TH (2007) Sexual selection in males and females. *Science*, **318**, 1882–1885.

Kokko H, Brooks R, Jennions MD, Morley J (2003) The evolution of mate choice and mating biases. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 653–664.

Tejedo M (1988) Fighting for females in the toad *Bufo calamita* is affected by the operational sex ratio. *Animal Behaviour*, **36**, 1765–1769.

Tejedo M (1992) Large male mating advantage in natterjack toads, *Bufo calamita*: sexual selection or energetic constraints? *Animal Behaviour*, **44**, 557–569.

Trivers RLL (1985) *Social Evolution*. Benjamin Cummings, Menlo Park, CA.

Wells KD (1977) The social behaviour of anuran amphibians. *Animal Behaviour*, **25**, 666–693.

Wright S (1931) Evolution in Mendelian Populations. *Genetics*, **16**, 97–159.

STRUGGLE TO MATE: EVOLUTIONARY DETERMINANTS OF BREEDING SUCCESS IN A LEK-BREEDING SPECIES

Leslie Faucher[1], Julie Jaquiéry[2], Thomas Broquet[3,4], Florian Boizet[1], Baptiste Faure[5], Cécile Godé[1], & Jean-François Arnaud[1]

[1] Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France
[2] Université de Rennes 1, UMR CNRS 6553 – ECOBIO, 35042 Rennes Cedex, France
[3] CNRS, Team Diversity and Connectivity of Coastal Marine Landscapes Roscoff France
[4] UMR 7144 Sorbonne Universités, UPMC Univ, Paris 06 Roscoff France
[5] Biotope- Nord Littoral, F-62720 Rinxent, France

Keywords : amphibian, sexual selection, inbreeding, fitness, *Bufo* [*Epidalea*] *calamita*, assortative mating, parentage analysis

Abstract

Sexual selection is advocated as a key evolutionary driver for reproductive strategies and trait evolution increasing the individual fitness through enhanced access to mates. A major challenge of sexual selection theory is to determine whether reproductive strategies arise through male-male competition, female choice or both. Anurans are models of choice to study the evolutionary causes and consequences of sexual selection. In this respect, the Natterjack toad (*Bufo calamita*) exhibits a lek-like mating system offering a wide range of reproductive strategies with diverse evolutionary outcomes. We investigated the evolutionary determinants of the Natterjack toad breeding success by field-monitoring, parental-assignment method and offspring fitness assessment over a whole breeding season. We showed evidence of a strong opportunity for sexual selection interacting with balanced effects of natural selection on individual fitness. Indeed, variance among males in mating success was large: less inbred, largest and more attending males accessed to mates. Offspring quality was positively influenced by parent traits, involving body length and male mating success, and negatively influenced by female inbreeding level. Interestingly, we also found an assortative mating by size modulated by a negative relationship between female SVL and the number of successful access to mate for males, suggesting a tradeoff between mate quality and mate quantity. Our study highlights the puzzling effect of male mating strategies in a lek-breeding system and shed new lights in the evolution of reproductive patterns underlying the combined effect of sexual and natural selection on individual fitness.

Introduction

Natural selection is a major micro-evolutionary process that takes many forms (*i.e.* directional, balancing or negative selection) and acts at different intensities on individual fitness (Hurst 2009; Orr 2009). Sexual selection, a special case of natural selection, concerns individual sexual reproduction and the ability to obtain and successfully copulate with a mate. Darwin (1871) was the first to describe the role of sexual selection as a key driver for trait evolution, advocating that some heritable traits are favorable not because they increase fitness like survival or fecundity but solely because they enhance mating success. In species with separated sex, the disparity of investment in gametes by males and females have resulted in male-male competition and female choice which are thought to be important evolutionary forces driving the sexual dimorphism in animals (Wells 1977; Trivers 1985; Clutton-Brock 2007; Henshaw *et al.* 2016). The benefits of female choice are broadly divided into direct benefits (*e.g.* paternal care, nuptial gifts) involving increased female fecundity, lifespan or offspring survival, and indirect benefits, through increase in the genetic quality of offspring (Andersson & Simmons 2006). Offspring quality can take the form of more competitive or attractive sons (Fisher's effect), or general viability ("good genes", reviewed in Kokko *et al.* 2003; Hosken & House 2011). When variation among males in mating success is related to specific behavioral or morphological traits, the effects of sexual selection may either reinforce or counteract potentially antagonists effects of natural selection on that trait (Andersson & Iwasa 1996; van Doorn *et al.* 2009; Johnston *et al.* 2013). Sexual selection in particular can explain the occurrence of costly traits such as conspicuous colors, display movements and costly acoustic advertisements of males (Wells 1977; Andersson & Simmons 2006). Hence, reproductive systems that include behavioral strategies or morphological traits devoted to get matings play an important role in micro-evolutionary processes by driving neutral and adaptive genetic variation over time (*e.g.* Lande 1977; Andersson & Simmons 2006). An important question in sexual selection theory is whether sexually dimorphic characters and specific behaviors used in display by males have arisen through male-male competition, female choice or both.

Anurans are models of choice to study the evolutionary causes and consequences of sexual selection (Howard 1979; Woodward *et al.* 1988; Jaquiéry *et al.* 2010; Luo *et al.* 2016; Zhao *et al.* 2016). Reproductive males are found in breeding sites during the whole reproductive period and form choruses which attract both sexes (Wells 1977). Fertilization is external and males grasp females with their legs before fertilizing the eggs when the females spawn (*i.e.* a position referred to as "amplexus" hereafter). This allows easy record of matings, providing a

wealthy literature on their breeding systems (Wells 2010). In contrast to other well studied taxa such as birds or fishes, anuran exhibit no strongly marked morphological sexual differences. Instead, they display distinct behaviors like striking male calls, offering a wide range of sexual selection patterns modulated by different degrees of intra and inter-sexual selection. In this way, Wells (1977) defined explosive and prolonged breeders that mainly differ by the strength of male-male competition due to differences in spatial and temporal distribution of females at the breeding sites. In explosive breeders, female density is usually high enough to favor male searching behavior and scramble male-male competition. Thus, in species where males fight for access to females, there will be selection on any traits which confer an advantage for males in obtaining access to females. In most anuran species which engage in scramble competition, sexual selection acts on body size, as large males are usually able to dislodge smaller males in amplexus with females (*e.g.*, *Bufo pardalis* : Cherry 1992; *Hyla intermedia:* Botto & Castellano 2016; *Rhacophorus omeimontis:* Luo *et al.* 2016; see however Zhu *et al.* 2016 on *Philautus odontotarsus*). In such a case opportunity for female choice is restricted. By contrast, in prolonged breeders the asynchronous female arrival generates highly male-biased sex-ratio. Male calling on stationary position is then favored, as males are very unlikely to encounter a female by active searching. This lek-like breeding potentially raises the strength of female mate choice, especially when females approach and initiate mating with one of the many calling males (Sullivan 1982; Arak 1983b, 1988a; Höglund & Alatalo 1995a). However these two categories are not always exclusives and in some species males switch from calling to searching strategies as operational sex-ratio varies (OSR, defined as the ratio of potential mating males to the total number of breeding adults at any one time, see Emlen 1976; Danchin & Cézilly 2008)

The Natterjack toad (*Bufo* [*Epidalea*] *calamita*, Laurenti 1768) is an example of anuran with a lek-like mating system (Höglund & Alatalo 1995a; Beebee & Denton 1996). It is described as an explosive breeder in its southern geographical range and as a prolonged breeder in its northern range (Arak 1983a; Tejedo 1988; Denton & Beebee 1993a). Within the same population, a wide variety of mate-locating strategies can be observed, from calling at stationary position to active searching and clasping as density increases (Arak 1983a). However, reproductive strategy remains puzzling in this species. Large males better maintain exclusive zones where they call to attract females and can be displaced only by larger males (Arak 1983a; Tejedo 1992a). Nonetheless, males can also exhibit a sneaker or switcher behaviour (Tejedo 1988; Denton & Beebee 1993a). Sneaker refers to silent (small) satellite males that mate with females attracted by other calling males. Switcher, refers to males that alternate between calling behavior and silent, sneaker behavior. Furthermore, male strategies may shift across night

during the breeding season and appear dependent on the OSR. More than body size or calling effort, the attendance (the number of nights a males spends in the chorus) may also explain male mating success in this species (Arak 1988b; Tejedo 1992a; Denton & Beebee 1993a). However, the effect of inbreeding on male and female reproductive success remains unknown. Because the Natterjack toad seems to display very different reproductive strategies, this species offers a good opportunity to investigate the evolutionary processes at work.

We investigated this issue in a Natterjack population of northern France by field-monitoring, parental-assignment and offspring fitness assessment. We studied male reproductive variance in breeding success by measuring the number of successful matings and fertility and offspring fitness by accounting on female effects. We addressed the following questions: (1) Do individuals mate randomly, irrespective of male and female attributes? (2) Is there a strong opportunity for sexual selection (3) Which specific morphological or behavioral traits drive the mating success? (4) What would be the relative effects of the strength of sexual selection over natural selection on male and female breeding success, *i.e*. can we infer different reproductive strategies modulated by a trade-off between traits under sexual selection and traits more affected by natural selection?

This study provides empirical evidences of a strong variance among males for access to mates, a significant effect of male and female attributes on breeding success, and a potential trade-off of male mating strategies to maximize breeding success. Our findings shed new lights into the interactions of multiple evolutionary processes that modulate individual breeding success.

Materials and methods

Studied species

The Natterjack toad is a long-lived amphibian occurring in western and north-central Europe, protected by national legislation (Beebee & Denton 1996; Denton *et al.* 1997; Wilkinson & Griffiths 2013). The Natterjack toad is an effective pioneer species confined to natural or artificial primary habitats such as coastal sand dunes, inland heaths or in gravels or sand pits (Rowe *et al.* 2000b; Flavenot *et al.* 2015; Faucher *et al.* 2017). Spawning takes place in sunny shallow temporary ponds and is followed by a short larval development period to cope with the frequent desiccation of breeding ponds. In northern France, adults breed continuously from April to June. Reproductive males are observed in breeding ponds during the whole reproductive period and form choruses attracting both sexes whereas females may come to the

ponds exclusively for spawning and leave immediately afterward (Arak 1983a; Sinsch 1988; Denton & Beebee 1993b; Beebee & Denton 1996). Females remain nearly submerged until being close to a selected male and generally initiate amplexus by physical contact. Males rarely fight for already paired females but they attempt to maintain an exclusive zone of attraction for females by quickly grasping rival calling males (Arak 1983a).

Population monitoring

A population of approximately one hundred adults established in a breeding site of northern France (3°12'11.7144" E, 50°19'1.0560"N) was surveyed during the whole breeding season, from April 1st to June 25th 2015. This population was chosen because it was geographically isolated (> 6 km from the nearest population), which minimizes bias related to undetected individual emigration or immigration during the sampling period. Field survey was performed about every two nights (depending on suitable weather conditions to breed, see Figure S1), totalizing 44 surveyed nights. The breeding site included three ponds: one permanent pond, one temporary pond, and one puddle each separated by 40 to 110 meters. Each night, we monitored the lek from dusk to the end of breeders activity.

All adults were non-invasively sampled using buccal swabs for nuclear DNA genotyping (Broquet *et al.* 2006; Faucher *et al.* 2016), sexed, measured (snout to vent length [SVL] and tibia length, see Table S1), weighted and marked using transponders implanted subcutaneously (PIT-tags 7 x 1.35 mm, 134.2 kHz, Loligo® System, Denmark). Every prospected night, all individuals found in the breeding site were identified based on transponders, and new non-marked individuals were captured for the same round of DNA sampling, morphometric measurements and pit-tag marking. All this procedure was carried out in the field in order to minimize disturbance. Amplectant pairs - when a male grasps a female with his front legs - were recorded without separating the pair. For technical reasons, only amplectant adults were sampled and measured the first six nights prospected (from 1st April to 13th April). For this period, the attending males and females were only counted. OSR was calculated for each night and used as an index of potential mating competition among males (Emlen & Oring 1977; Jones *et al.* 2004).

All observed clutches were individually collected, and were further photographed in order to enumerate eggs when possible (*i.e.* when water clarity was sufficient). The date of laying was recorded based on the eggs and jelly appearance (first or second days of development). 120 eggs per clutch were sampled in three distinct sections of 40 eggs distributed in different parts of the clutch. Eggs were kept in 500ml of pond water in a plastic bucket and

took back to the laboratory until hatching. The rest of the clutch was released on the same place and marked to avoid resampling.

Offspring fitness

Survival and growth were measured on every collected clutches in controlled conditions in the laboratory. Clutches were reared in 400ml tanks at 20°C, 13h light per day. 200ml of water was changed every four days (with tap water dechlorinated after 24h). We also randomly swapped the position of the tanks to minimize uncontrolled effects. Hatching success was recorded as the proportion of hatched tadpoles per clutch. Time to hatching was recorded as the period from clutch sampling to hatching of more than 15 tadpoles (*i.e*. 12% of a clutch). Newly hatched tadpoles were released in their pond of origin, excepted for 30 tadpoles further reared for fitness assessment and 16 tadpoles frozen in 500µl of 70% ethanol for genotyping.

Tadpole stage at the beginning of development assessment was recorded as the mean Gosner stage (Gosner 1960) over the 30 sampled tadpoles. Indeed, within a clutch, the time lag between the different hatches induced stage heterogeneity. Tadpoles were daily fed *ad libitum* with crushed rabbit food (dehydrated alfalfa). Dead individuals were daily removed from the water. The 11th and 25th day after hatching, fifteen tadpoles per clutch were randomly chosen and individually weighted (Ohaus Scout Pro ±0.001, Greifensee, Switzerland). Survival was recorded at days 11, 25, 31, 45 and 50 (Table S1). Metamorphosis was considered to be completed when toadlet had four distinct legs. We noted the time period until the first tadpoles completed its metamorphose, until the 15th completed its metamorphose, and the number of metamorphosed individuals at days 45, 60 and 70. All measured traits are presented in Table S1. Due to logistical constraints, the rearing experiment was ended in late June. At this stage, 84% of the survival/remaining clutches had started to metamorphose. All tadpoles or toadlets were then released in their pond of origins.

DNA extraction and nuclear DNA genotyping

Whole adult genomic DNA was extracted from swabs using Macherey Nagel NucleoSpin® 96 trace kits (Düren, Germany) following the standard protocols outlined in the manufacturer's handbook. Tadpole DNA was extracted from entire individuals at stage 19-20 (Gosner 1960). All adults and 16 tadpoles per clutch were genotyped at 14 nuclear autosomal microsatellite loci described in Table S2 (Faucher *et al.* 2016). Adults were also genotyped at 23 supplementary nuclear microsatellite loci (Rowe *et al.* 1997; Rogell *et al.* 2005; Faucher *et al.* 2016 see Table S2). A pairwise comparison of adult multilocus genotypes was performed to

identify eventual duplicates due to sampling the six first nights without pit-tag marking. Multilocus genotypes considered as duplicates were characterized by a minimum of 85% of similitudes and differences were likely to arise from allele drop-out on one to 5 loci. . Nuclear genetic diversity was described in terms of expected gene diversity (H_E, Nei 1987), number of alleles and intra-population fixation index (F_{IS}), a measure of departure from panmixia, using *F*-STAT *v*2.9.3.2 (Goudet 1995). Significance of departures from Hardy–Weinberg (HW) equilibrium were tested for hypotheses of either excess or deficit in heterozygotes using a multisample score test (Rousset & Raymond 1995). Finally, a measure of individual inbreeding, the internal relatedness (*IR*, Amos *et al.* 2001), was estimated for each adult using the R package GENHET (Coulon 2009). All morphological and genetic measures are described in Table S1.

Parentage assignment

We inferred adult reproductive success using genetic-based parentage analysis. This method allows *a posteriori* parent and progeny assignments without ambiguities with a minimum disturbance for parent-pairs. Parentage analyses were carried out by full-pedigree likelihood methods using algorithms implemented in COLONY *v* 2.0.6.1 (Jones & Wang 2010). Multiple runs were carried out by setting, step by step, priors on sibship to guarantee detection of right sibship and parent pairs. We first set priors on the composition of sibship (out of the 16 offspring genotypes from each clutch) and allowed polygyny and polyandry by considering that we sampled contained 60% of mothers and 90% of fathers. This conservative choice was based on the fact that females were more cryptic and less likely to be sampled than calling males (Beebee 1979). Multiple rounds of analyses considering a range of 60% to 80% of females sampled and 90 to 95% of males sampled yielded identical results. As suggested by Jones & Wang (2010), we choose a medium length run with medium prior on the sibship size and sibship size scaling for a good compromise between running time and accuracy. Next, we verified the reliability of inferred parent pairs by cross validation with field observations, *i.e.* we verified the congruence in the date of presence of breeding adults inferred by parentage analysis and observation of amplexus in the field. Afterward, we set priors on parent identities of previously validated families (which represent 89% of clutches) to infer parent-pairs with unknown adults. Overall, all parent-pairs defined *a posteriori* based on genetic data were already observed in amplexus during field survey, corroborating our parentage assignments. Finally, these results were confirmed by additional analyses using algorithms implemented in CERVUS (Kalinowski *et al.* 2007, data not shown).

Measures of opportunity for selection

The breeding success of each adult was estimated from molecular parentage analyses. For each adult n of sex j, the breeding success was split into five different components: the mating success ($w1_{nj}$) measured as the number of successful mating, the reproductive success ($w2_{nj}$) measured as the sum of eggs produced, the hatching success ($w3_{nj}$) which is the estimated number of hatched eggs, the estimated number of surviving offspring ($w4_{nj}$), the estimated number of surviving tadpoles at day 45, and the estimated number of tadpoles having completed their metamorphose at day 60 ($w5_{nj}$). These estimates were based on the absolute number of eggs produced by each adults multiplied by the proportions of hatched tadpoles, survival tadpoles and tadpoles having completed their metamorphose. To investigate the relationship between mating success and reproductive success we estimated the standardized Bateman gradient (β'_{SS}, Bateman 1948) as the slope of least-squares regression of relative reproductive success (obtained by dividing reproductive success by mean reproductive success) on relative mating success (obtained by dividing mating success by mean mating success) (see Howard 1979; Jones 2009; Mobley & Jones 2013).

To test the randomness of male mating success, data on number of matings per male was examined for goodness-of-fit to a Poisson distribution (Wright 1931). We used likelihood ratio test implemented in the package R "vcd" (Meyer et al. 2016). The opportunities for selection on males and females (I_j) were further measured as described in Jones (2009): equation A1 separates the components due to the variance in mating success (opportunity for sexual selection, I_{js}) and reproductive success (I_{j2}), as well as hatching success (I_{j3}), offspring survival (I_{j4}), metamorphosis success (I_{jTotal}).

(A1) $I_{ji} = \frac{\sigma^2(w_{ji})}{(\overline{w}_{ji})^2}$ where $\overline{w}_{ji}$ and $\sigma^2(w_{ji})$ are mean and variance of the given component of breeding success *i*, in the sample of *N*=n individuals for each sex *j* (equation 9a and 9b in Arnold & Wade 1984).

Finally, and as suggested by Henshaw et al. (2016) we measured the maximum strength of premating sexual selection following the equation of Jones index, S' (equation A2, see Jones, 2009).

(A2) $S'_{j_{max}} = \beta_j'_{SS} \times \sqrt{I_{j_S}}$ I_{js} being the index of opportunity for sexual selection and $\beta_j'_{SS}$ being the standardized Bateman gradient previously described for each sex *j*.

Individual traits and reproductive success

We aimed at identifying the effect of male traits on male mating success and estimating the parental effects on offspring fitness. To accomplish the first objective, we analyzed male mating success both at a seasonal and at a nightly level. Indeed, previous studies reported that different male strategies occurred in a population and were dependent of nightly chorus composition which may greatly vary during the breeding season (Arak 1983a; Tejedo 1992a). Statistical analyses were conducted using the R Studio software v 1.0.44.

First, we investigated the potential effect of male body size, attendance and OSR on mating success by comparing traits across the different categories of breeders (no sires, or mated one, two or three times) using student t and Mann-Whitney *U* tests. The same statistical approaches were used to investigate trait differences between non-breeding and breeding females and to compare male and female SVL.

Second, we analyzed the effect of male traits on mating success using generalized linear models (GLM). At seasonal level, we regressed the seasonal mating success of a male against his SVL, chorus attendance (number of attending nights), mean OSR per night and *IR* estimated on multilocus genotypes. GLM using zero-inflated Poisson regression modelling (R package Pscl, Zeileis et al. 2015) was applied with mating success accounteing for 0, 1, 2 or 3 mates, as suggested in Zuur et al. (2009) for counted data (Model A). Because most males only sired one time (see results), a logistic regression (R package Lme4, Bates et al. 2015) was also applied considering only unsuccessful males and males having one partner (Model B). At nightly level (Model C), we considered the recording nights in which there were 2 or more males with at least one of them that mated (N = 6 nights and 64 males observed) and analyzed the competition effects night by night on the mating success. GLM was fit by maximum likelihood following Laplacian approximation with binomial distribution. As suggested in Botto & Castellano (2016a), male identity, mating events but also female mean size and the number of attending males were set as random factors (R package Lme4, Bates et al. 2015).

Third, the relationship between male mating success and female SVL was measured with a Spearman correlation. To verify the statistical significance of this correlation based on a very small sample size, we further performed 10,000 permutations of female SVL on male mating success and compared the observed Spearman index with the 10,000 Spearman index randomly obtained. Linear regression was performed on SVL of male and SVL of female for successful breeding pairs or for unsuccessful amplectant pairs. In the same way, we measured the differences between female SVL and male SVL for all possible breeding pairs in the population (*i.e* under random mating hypothesis, Δ_{Rand}) and compared the resulting distribution (that assumes random pairing with respect to body size) with the SVL difference observed for

successful pairs (Δ_{Succ}) and unsuccessful pairs (Δ_{Uns}). To control for night effect (*i.e*. potential assortative mating due to significant differences in size-related cohort composition, see Sinsch 1997), we also measured differences between all possible breeding pairs within each observed night (hypothesis of random mating within night, $\Delta_{Rand/Night}$). Pairwise distributions of Δ_{Rand}, $\Delta_{Rand/Night}$, Δ_{Succ}, Δ_{Uns} were compared using Mann-Whitney *U* tests.

Parent traits effects on offspring fitness

We investigated the effect of parent traits (SVL*, IR* and male attendance) on offspring fitness traits by controlling for potential experimental effects on offspring fitness. To analyze parent trait effects on hatching rate, survival rate and metamorphosis rate, a multivariate GLM with binomial error distribution and logit link fit by maximum likelihood (Laplace Approximation) was performed (Model D) (R package Lme4). Laying date, hatching date and egg stage at sampling, and tadpole stage at the beginning of fitness record were included as covariate (fixed effect). We did not set any random effect because the pseudo-replications of father ID represented a small sample size (data were available for only 2 on 5 males mated with two or more females) and model failed to converge.

In contrary to the former tadpole fitness traits, tadpole body mass was recorded on several tadpoles within each clutch and had a Gaussian distribution. Hence we were allowed to performed a linear mixed effect model fit by restricted maximum likelihood to analyze parent trait effects on tadpole body mass (Model E). Only male ID was set as random factor because no female was recorded two times in the offspring fitness dataset. Laying date, hatching date and egg stage at sampling, tadpole stage at the beginning of fitness record, parent SVL, *IR* and male attendance were set as covariates

Model selection

We ranked models based on the Akaike information criterion corrected for small sample size (AICc, Hurvich & Tsai 1989, reviewed in Bolker et al. 2008). This provides a metric for model selection by comparing the likelihood of a set of models while penalizing for their number of parameters. We determined the relative importance of each covariate based on the sum of Akaike weights (SW) across the entire model set using the package "MuMIn"(Bartoń 2016). In addition, we used a likelihood-ratio test implemented in the R package "Lmtest" (Zeileis & Hothorn 2002) to determine whether each explanatory variable of the model has a significant effect on goodness of fit of model. The model retained by drop AICc procedure was compared to the partial models excluding successive retained variables. For zero inflated model

and logistic regression, we computed the pseudo R^2, a measure of model goodness of fit defined by the ratio of log-likelihood of the full model and log-likelihood of intercept model (McFadden 1974). For mixed effect models, we calculated R^2 (m), the marginal R^2 that describes the variance explained by the fixed factors, and R^2 (c), the conditional R^2 describing the variance explained by both the fixed and random factors (Nakagawa & Schielzeth 2013). Finally, to evaluate the effect of parent traits on offspring fitness (Model D and E), the amount of variance in offspring fitness explained by each parent trait was calculated using linear regressions of male breeding values using R package "piecewiseSEM" (Lefcheck 2016). Regression were applied by calculating for each offspring-fitness trait y the residuals of the partial model:

$$y \sim \text{Egg stage} + \text{Sampling date} + 1|\text{Father} + 1|\text{Mother}$$

The father and mother ID were set as random factor when we analyzed tadpole body mass only. The residuals of all selected models were checked using diagnostic plots to see whether they met the assumptions of linear models.

Results

Chorus attendance, genotyping and parentage analysis

Overall, 71 males and 54 females were surveyed for morphological traits and genotyped during the whole breeding season. On 44 nights monitored, the number of observed calling males ranged from 0 to 53 males per night (mean 16.05 males, SD = 16.67). We observed females on 27 nights and breeding pairs on 17 nights. There were more males than females every nights (mean sex-ratio$_{♀/♂}$= 0.793 SD = 0.317). Eight nights were characterized by an absence of calling males (for details see Table S3, Table S4 and Figure S1). Male tenure ranged from 1 to 18 nights of attendance (mean = 7, SD = 4). All males caught in the pond were observed calling indicating that they were actively participating at the chorus. We only observed one toad ball composed of six males and one female. The female finally reproduced with one of the male. Nineteen females were captured at least twice and fourteen spawned after their second visit at the pond. Morphological measures were performed on almost all males and for 55 to 79% of females, depending on the trait measured (missing data mostly concerned body mass for amplectant individuals). A total of 44 clutches were collected. Out of 44 sampled clutches, 42 were successfully genotyped (2 clutches showed no sign of development and cannot be genotyped), and 41 were reared until 70 days. Five clutches stemmed from already sampled sibships and one tank contained two distincts sibships melted during sampling.

Offspring fitness data (presented below) were not analyzed for these five tanks totalizing 35 offspring monitoring and 36 offsprings genotyped

All adults were successfully genotyped at 37 microsatellite loci with an overall rate of 2% missing data. Genetic diversity ranged from 0 to 0.822 and the number of alleles ranged from 1 to 9. One locus (Bcalμ9) was excluded because of significant departures from HW equilibrium. For offspring genotypes, mean number of alleles per locus was of 4.79, SD=1.97 with 4% missing values (Table S2). Clutches were assigned without ambiguity based on multilocus genotypes at 14 microsatellite loci. 10 tadpoles (over the 670 genotyped) could not be assigned to due to genotyping failures. Thirty seven parent pairs were identified, and 33 clutches (91.7%) were assigned to a sampled father and 32 (86.5%) to a sampled mother. Genotypes of unsampled fathers and mothers which produced clutches early in April were inferred a posteriori based on offspring genotypes.

Reproductive success

Each mother laid only once except one female that arrived in April and laid a second clutch in June. Linear regression of the number of eggs enumerated on mother SVL (sample size = 13) allowed to estimate the clutch size from mothers' SVL when clutches were not enumerated but mother SVL was known ($r = 0.611$, $P < 0.05$). The mean number of eggs per clutch was 2744.6, SD= 964.3 (min= 1350, max= 4428). No within clutch multipaternity was observed. Overall, 42% of males (*i.e.,* 31 males) and 62% of females (*i.e.* 36 females) mated at least once. Forty-three males did not reproduced, whereas successful males had one to three mates (Figure 1). The distribution of male mating success, w_1 (mean = 0.5, SD=0.67) did not departed from a Poisson distribution (Likelihood ratio = 1.394539 df=2 0.498, $P = 0.49$, Figure S2). The resulting male reproductive success, w_2, varied from 1350 to 5627 eggs sired (mean = 1244.721, SD = 1660.313). Measures of offspring fitness was available for 33 of the 37 clutches genetically assigned to parents. From 74% to 100% (mean =0.95, SD = 0.06) offspring hatched and from 6% to 100% (mean = 0.76, SD = 0.28) of tadpoles per clutch survived until day 45. The first metamorphosis per clutch was observed from day 38 to day 65 (mean = 48, SD = 6.6). All offspring fitness measures are given Table S5. Twenty-two females (38%) did not lay any clutch. Twelve were caught in resting habitat and can be considered as non-sexually receptive and were therefore removed from the data set of breeders. The remaining 10 females were observed at least one night in the pond and can be considered as unsuccessful reproductive females.

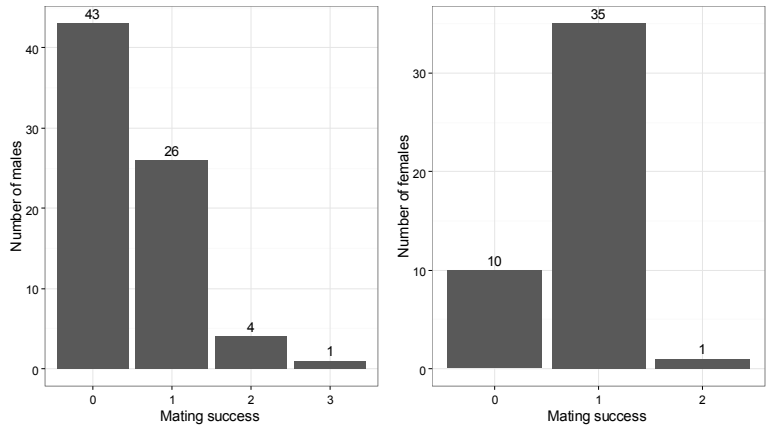


Figure 1: Distribution of mating success for males (A) and females (B) in the Natterjack toad (*B. calamita*).

Opportunity for selection

Indices of opportunity for selection were four to 14 times higher in male than in female (Table 1). For males, the variance in relative reproductive success was very similar to the variance of mating success. In contrast, variance in female reproductive success was 1.6 times higher than variance of female mating success. Finally, sexual selection (number of mates) appeared to contribute to 13.38 % and 16.80 % to total opportunity for selection when comparing the ratio I_S/I_{Total} in males and females respectively (Table 1). The Bateman gradient measured on males was β'_{SS} =0.941 (Figure 2). We could not estimate Bateman gradient for females because only one laid two clutches. The maximum strength of premating sexual selection on males measured by Jones index (S'_{max}) was 1.44.

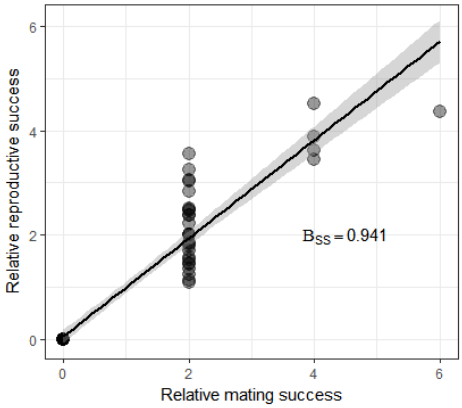


Figure 2: Bateman gradient estimates for *B. calamita* adult males. Male mating success and reproductive success are divided by their mean. Males with zero mating success and zero reproductive success are included in the analysis. Bateman gradient is given by y ~ 0.94137x +0.05863 and significantly differed from zero ($P < 0.001$).

Table 1: Variance in relative breeding success taking into account different components of breeding success in the Natterjack toad (*B. calamita*). Variances in relative mating success I_S (opportunity for sexual selection), in relative reproductive success (I_2), in relative hatching success (I_3), in relative offspring survival at day 45 (I_4) and in relative metamorphosis rate at day 60 (I_{total}).

	I_S	I_2	I_3	I_4	I_{total}
Males	1.781	2.117	2.433	11.469	13.302
Females	0.317	0.504	0.558	0.796	1.879

Adult traits and reproductive success

SVL did not significantly differ between sexes ($t = 0.904$, df = 93.538, $P = 0.368$). Male SVL ranged from 53.76 mm to 79.32 mm (mean = 66.5mm SD=5.6) and female SVL from 50.19 mm to 84.25 mm (mean = 67.5mm, SD = 6.5) (Table S4). No significant differences were observed for SVL and mass between successful breeding and non-breeding females or between successful breeding and unsuccessful breeding females (Kruskal-Wallis $\chi^2 = 3.635$, df = 2, $P = 0.162$ and Kruskal-Wallis $\chi^2 = 2.929$, df = 2, $P = 0.231$ respectively, Figure S3). No significant differences were detected among successful (one, two or three mates) males for morphological, *IR*, chorus attendance or mean OSR (Kruskal Wallis, $P > 0.05$ in each case, Figure S4). Nonetheless, successful males with one mate were significantly larger, heavier and less inbred than unsuccessful males (all at $P < 0.05$, Figure 3A, B and E). No significant differences were detected for other male traits (Figure 3 C, D, F, G).

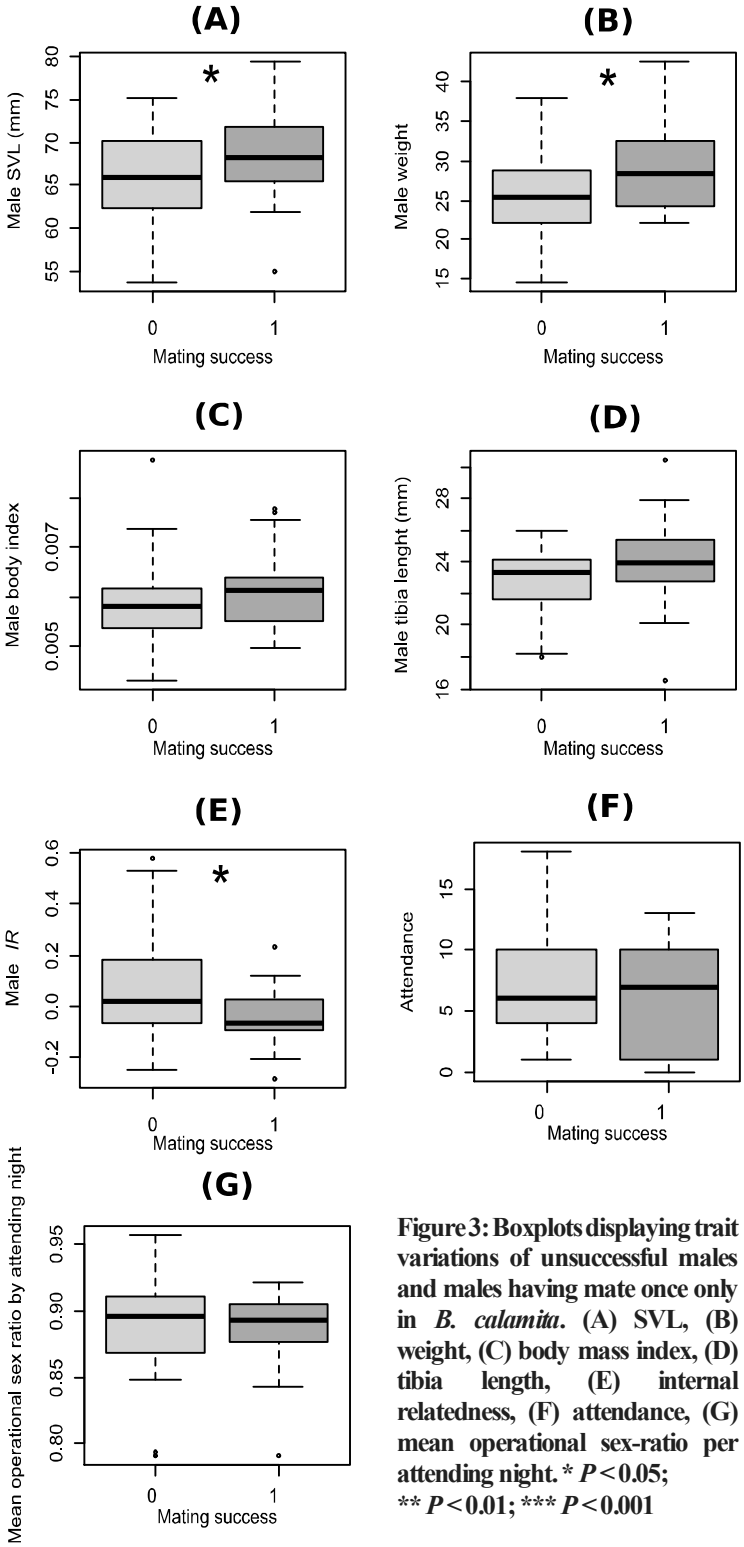


Figure 3: Boxplots displaying trait variations of unsuccessful males and males having mate once only in *B. calamita*. (A) SVL, (B) weight, (C) body mass index, (D) tibia length, (E) internal relatedness, (F) attendance, (G) mean operational sex-ratio per attending night. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Using GLM, we retained only the SVL as an indicator of individual body condition because all morphological traits were highly correlated. Distribution and collinearity of remaining variables are described in Figure S5. In model A, *IR* had a significant negative effect on the mating success ($P = 0.045$). Attendance also positively explained the mating success ($P = 0.002$, Table 2). Analysis of model B taking account only of males that mated once or not (*i.e*. discarding males with two or three
mates) indicated a significant positive effect of body size and a negative effect of *IR* on male mating success (Table 3). Mean OSR, and attendance had no significant effect (Table 3). Cross validation graphs for these generalized linear models are given Figures S6 and S7.

Table 2: Zero inflated Poisson model (Model A) on male mating success (0, 1, 2 or 3 successful mates) with standardized (centered-reduced) data. The standardized coefficient (βs), standard errors (Std.Error), sum of Akaike weights obtained for each explanatory variable and Akaike information criterion of model without the given explanatory variable (Droped AICc) are indicated. Likelihood ratio tests measure the significance of a variable in improving model goodness of fit. Pseudo R-squared is given for each global model as a measured of goodness of fit. AICc = 122.20. Pseudo R-square (Mc Fadden) as 1−log(Lc)/log(Lnull) = 0.179 (df = 10) for the full model

Explanatory variables	β_S	Std.Error	Droped AICc	Sum of Akaike weights	Likelihood ratio test
(Intercept)	-0.435	0.196			
Attendance	0.023	0.216	125.100	0.900	$\chi^2 = 12.626$, df=2, $P = 0.002$
IR	-0.025	0.221	128.500	0.570	$\chi^2 = 6.197$, df=2, $P = 0.045$
Mean OSR	-0.491	0.237	125.600	0.320	$\chi^2 = 0.934$, df=2, $P > 0.050$
SVL	-0.314	0.225	122.500	0.140	$\chi^2 = 5.702$, df=2, $P > 0.050$

Table 3: Logistic regression (Model B) on seasonal male mating success (0 or 1 successful mate) with standardized (centered-reduced) data. See Table 2 for additional information on data presented and model selection. AICc = 77.1, Pseudo R-square (Mc Fadden) as 1−log(Lc)/log(Lnull) = 0.161 (df = 5)

Explanatory variable	β_S	Std.Error	Droped AICc	Sum of Akaike weights	Likelihood ratio test
(Intercept)	-1.065	0.345			
Attendance	0.047	0.358	74.800	0.240	$\chi^2 = 0.017$, df=1, $P > 0.050$
Mean OSR	-0.144	0.337	74.900	0.260	$\chi^2 = 0.184$, df=1, $P > 0.050$
SVL	0.656	0.312	79.800	0.820	$\chi^2 = 5.076$, df=1, $P = 0.024$
IR	-1.117	0.466	83.000	0.950	$\chi^2 = 8.259$, df=1, $P = 0.004$

At a nightly level, Model C did not revealed any significant effect of male IR and of male SVL on mating access, perhaps due to low sample size ($N < 10$) that limited power (Table 4, Figure S8).

Table 4: Generalized mixed effect linear model (Model C) on nightly male mating success with standardized (centered-reduced) data. The standardized coefficient (βs), standard errors, sum of Akaike weights obtained for each explanatory variables and Akaike information criterion of model without the given explanatory variable are indicated. Likelihood ratio test results are provided as a measure of significance of variable in improving model goodness of fit. Root-mean-square deviation was equal to 0.610. Pseudo R-squared ($R^2GLMM_{(c)}$ and $R^2GLMM_{(m)}$) are also given for each global model as a measured of goodness of fit. AICc =117.8, $R^2GLMM_{(c)}$ =0.430, $R^2GLMM_{(m)}$ =0.050

Explanatory variable	β_S	Std.Error	Droped AICc	Sum of Akaike weights	Likelihood ratio test
(Intercept)	-3.071	0.889			
Attendance	0.47723	0.448	116.9	0.345	$\chi^2 = 1.263$, df=1, $P > 0.050$
Mean OSR	-0.353	0.354	116.6	0.317	$\chi^2 = 1.027$, df=1, $P > 0.050$
SVL	-0.029	0.374	116.0	0.304	$\chi^2 = 0.412$, df=1, $P > 0.050$
IR	-0.235	0.360	115.6	0.259	$\chi^2 = 0.006$, df=1, $P > 0.050$

With respect to female traits, we observed a significant negative correlation between male mating success and female SVL (Figure 4A, ρ = - 0.309, $P = 0.001$). Males that mated once reproduced with larger female partners than males that mated twice or thrice. A significant positive relationship between SVL of adults forming breeding pairs (linear regression $r^2 = 0.169$ $P < 0.05$ Figure 4B) was also observed (assortative mating), but not for the 19 unsuccessful amplectant pairs ($r^2 = 0.028$ $P = 0.235$).

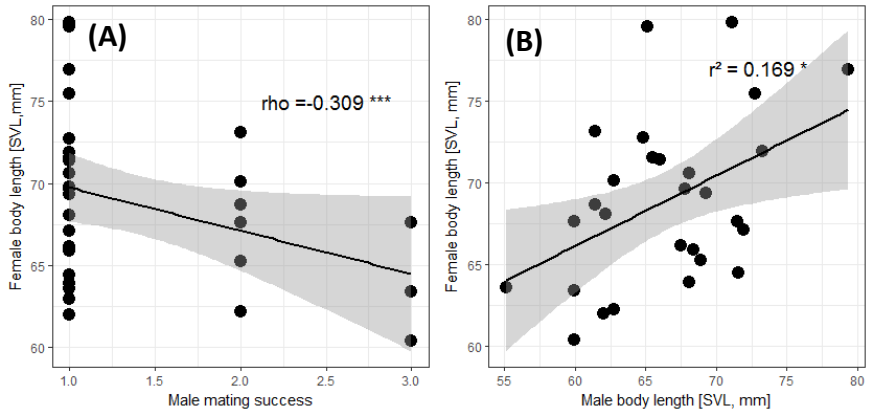


Figure 4: Relationships between mother body length and father mating success (A) and father and mother body lengths (B). 95% CI of the regression line is shown in grey, as well as Spearman correlation coefficient (A) and adjusted coefficient of determination (r^2) (B).

Furthermore, a significant difference was found between Δ_{Succ} ($N = 28$) and Δ_{Uns} ($N = 19$) (W = 389, P = 0.008, Figure 5). In the same way, Δ_{Succ} significantly differed from Δ_{Rand} measured on the whole population ($W = 21586$, $P = 0.02865$) while Δ_{Uns} did not ($W = 23424$, $P = 0.1069$, Figure 5). This corroborated that differences between SVLs of amplectant individuals are significantly lower for successful breeding pairs than expected under random pairing. However, this trend was not observed at nightly level: Δ_{Succ} did not differed from $\Delta_{Rand/Night}$ while Δ_{Uns} did ($W = 12811$, $P = 0.167$ and $W = 13331$, $P = 0.025$, Figure 5). More details on male and female SVL for successful breeding pairs or for unsuccessful amplexus can be found in Figure S10.

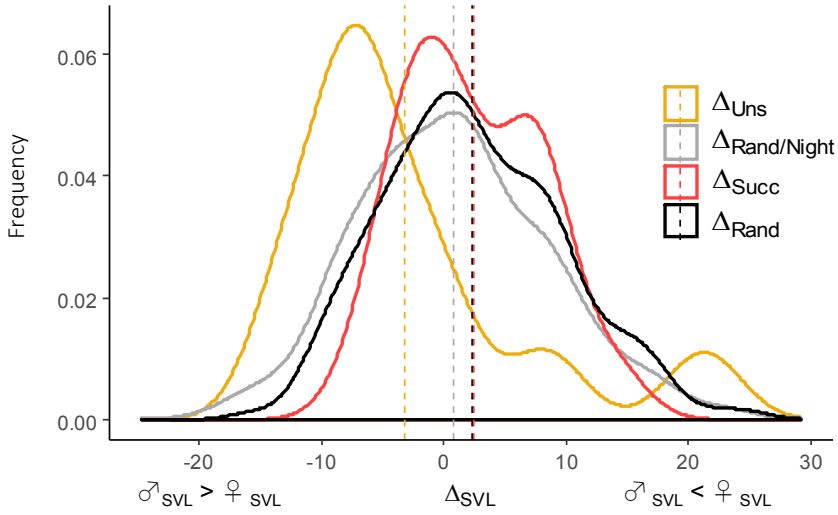


Figure 5: Density plot depicting the distribution of differences between female body length and male body length for different data set: all possible pairs in the population (Δ_{Rand}), all possible pairs within night of attending ($\Delta_{Rand/Night}$), amplectant unsuccesfull pairs (Δ_{Uns}) and observed successful breeding pairs (Δ_{Succ}) in a population of the Natterjack toad (*B. calamita*). Means were represented by a dashed line.

Offspring fitness and parent traits

We observed several significant effect of female traits on fecundity and offspring fitness. Mother IR significantly decreased clutch hatching success (Figure 6A). A large part of the variance in clutch size was also explained by mother SVL (Figure 6B). Mother SVL was significantly correlated with both metamorphosis success and mean tadpole body mass (Figure 6 C, D, E, F). In addition, after controlling for experimental effects, models of parent trait effects on offspring fitness indicated a significant positive effect of father SVL, father mating success and mean OSR per attending night on tadpole survival. We also observed a negative effect of father and mother *IR* on tadpole metamorphosis success ($P < 0.01$, Table 5). Finally we detected an important positive effect of male mating success on tadpole survival ($P < 10^{-3}$).

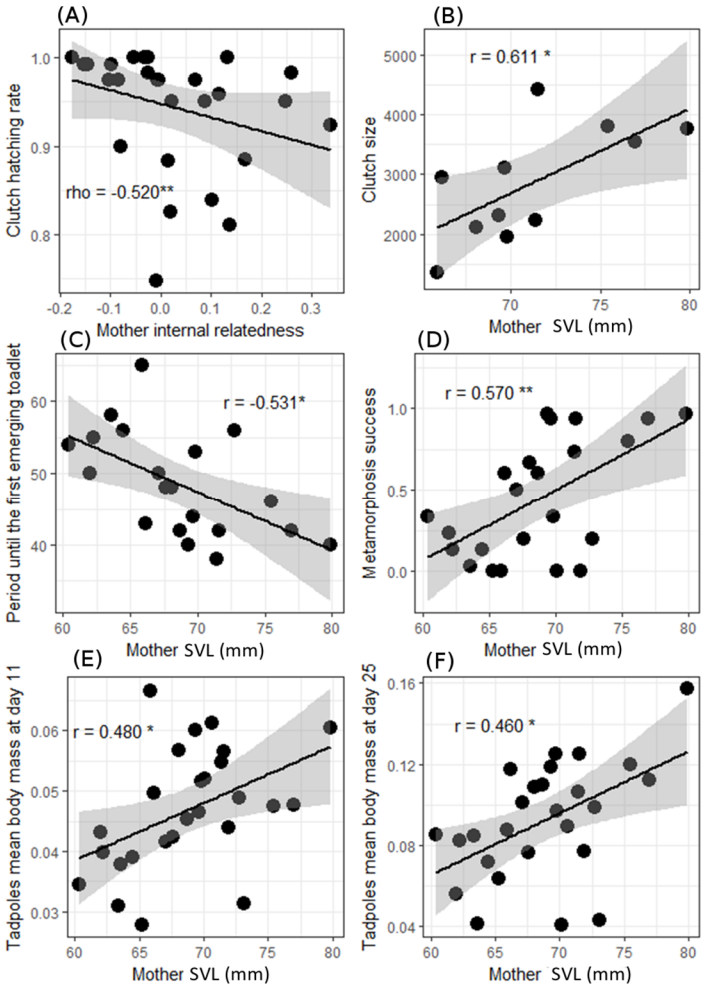


Figure 6: (A) Hatching success against mother internal relatedness (a measure of inbreeding). (B) Clutch size, (C) period until the first emerging toadlet and (D) metamorphosis success, (E and F) mean tadpole body mass at day 11 and day 25, plotted against mother body length (mm). 95% CI of the regression line are shown in grey. Also presented, adjusted coefficient of determination (r) or the Spearman' rank coefficient.

Table 5: Variance in offspring fitness explained by each parent trait in the Natterjack toad (*B. calamita*). The sign of the relationship (brackets) and the significance level estimated with generalized linear model and linear mixed effect model are also given. *NS*: non-significant, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

	Generalized linear model (Model D)			Linear mixed effects model (Model E)	
	Hatching success	Offspring survival at day 45	Metamorphosis success at day 60	Body mass at day 11	Body mass at day 25
Male mating success	1.859 % *NS*	16.515% (+) ***	0.445 % *NS*	0.022% *NS*	0.014% *NS*
Male attendance	3.575 % *NS*	5.671 % *NS*	2.051% *NS*	0.009% *NS*	0.015% *NS*
Mean OSR	3.473 % *NS*	8.977% (+) *	1.832% *NS*	0.002% *NS*	0.000% *NS*
Male SVL	15.709 % *NS*	3.486% (+) *	3.929 % *NS*	0.001% *NS*	0.034% *NS*
Male *IR*	2.048 % *NS*	7.621 % *NS*	7.210% (-) **	0.008% *NS*	0.014% *NS*
Female SVL	0.860 % *NS*	1.725% *NS*	0.079 % *NS*	0.048% *NS*	0.042% *NS*
Female *IR*	17.277 % *NS*	9.644 % *NS*	6.498% (-) ***	0.001% *NS*	0.050% *NS*

Discussion

The theory of evolution by sexual selection accounts for elaborated and costly behavioural and morphological traits, and is rooted on the ability of individuals to access to mate (Danchin & Cézilly 2008; Jennions & Kokko 2010; Johnston et al. 2013). Sexual selection thus underlies some of the most extreme traits in nature that would otherwise remain unexplained by natural selection and may affect macroevolutionnary processes such as speciation and extinction (Ritchie 2007; Mank 2009; van Doorn et al. 2009; Lisle & Rowe 2015, but see Kraaijeveld et al. 2011). Variance in breeding success of males and females affects the strength of sexual selection (reviewed in Henshaw et al. 2016). In particular, lek breeding populations are ripe to high opportunity for sexual selection on males, resulting from high variance in mating success (Wells 1977; Höglund & Alatalo 1995b). As observed in previous studies, the Natterjack toad is mostly monogamous (26 out of 31 reproducing males mated with a single female in our study) (e.g. Arak 1983; Tejedo 1992). However this lek-like mating system with no within-clutch multipaternity supposes harsh pre-copulatory sexual selection on traits correlated with mating success (Höglund & Alatalo 1995a; Jones 2009; Lüpold et al. 2017). Assuming variance in male quality, females may have a great benefit to choose the best father because they lay a single clutch in most cases (Andersson & Simmons 2006). We thus hypothesized that in such prolonged-breeder population, sexual selection occurs through male-male competition and through female choice or male choice as it is observed in a wide range of taxa (Wells 1977; Wilbur et al. 1978; Owens et al. 1994; Andersson & Iwasa

1996; Sullivan & Kwiatkowski 2007). We expected high opportunity for sexual selection in males and lower opportunity for sexual selection in females (Kluge 1981).

In our study, variance in female mating success ($I_{s♀}$) was similar to values reported in other studies on *B. calamita*. Denton & Beebee (1993a) found that 44 to 64% of females doing not reproduce in one year. We observed non breeding females after the first breeding peak, when water level was lower and density of tadpoles higher (personal observations) compared to the beginning of the reproductive season. Therefore, a delaying behavior could have occur (Denton & Beebee 1993a).

The distribution of male mating success did not depart from a random Poisson distribution. A large part of males did not reproduce and very few of them mated more than once. The highly male-biased OSR and the low number of reproductive females can explain the numerous unsuccessful males (Kvarnemo & Ahnesjö 1996; Jirotkul 1999; Pröhl 2002; Wang et al. 2009). We measured opportunity for sexual selection on males as previously measured in a wide range of anuran populations where sexual selection on male traits have already been demonstrated (median opportunity for selection = 1.8, see Table S6). Overall, the opportunity for sexual selection seems to be lower in species with male territory defense mating (Kluge 1981). In our study, variance in mating success was similar to variance in reproductive success suggesting that there are weak variance in mate quality (i.e, each mate give similar number of sired eggs). This trend is verified by the measure of Bateman gradient which is close to 1 and indicates that number of sired eggs increased proportionally with the number of matings. Variance in reproductive success is attributable partially to sexual selection and partially to natural selection on fecundity or fertility. S'_{max} is one step closer to the sexual selection process, because it only deals with selection generated by differential mating success. Herein the Jones index measured suggest high opportunity for precopulatory sexual selection on natterjack toad males.

Over the whole breeding season, we found a positive association between male attendance and mating success. Conversely, a negative effect of male *IR* on mating success was observed. Males that mated once were larger and less inbred than the other males. Nonetheless, the few males that sired two or three clutches were not larger or less inbred than expected under random sampling and, interestingly, a negative correlation between female SVL and partner mating success was observed. We also found an assortative mating by size. Altogether, these complex patterns of mating system were further influenced by female quality that determine offspring fitness. Below, we outline the different evolutionary processes at work that may explain our findings.

Individual features and breeding success: SVL

Behavioral and morphological secondary sexual traits, like male extravagant ornaments, are thought to evolve by sexual selection (Danchin & Cézilly 2008). Male body size is one of these traits, that is often under positive pre-copulatory sexual selection (Andersson & Iwasa 1996). In anurans, both male-male competition and female choice have been identified as mechanisms that contribute to selection for larger male body size (Wilbur et al. 1978; Arak 1983a; Morris 1989; Tejedo 1992a; Böll & Linsenmair 1998; Bowcock et al. 2013; Luo et al. 2016). Body size could be under directional selection driven by direct or indirect interactions on male-male competition, i.e struggle or endurance rivalry (Arak 1988a; Murphy 1998), or by female choice resulting from direct or indirect benefits to mate with males of specific body size (Morris 1989 but see Ursprung et al. 2011).

In the Natterjack toad, Arak (1983b) found that large body size was an advantage in direct male-male competition to defend territory by clasping rivals. Indeed, larger males are more likely to initiate attack. Larger males might be also advantaged indirectly by call better detected by females, increasing their probability to encounter females (Arak 1988a). We observed concordant behaviours: males rarely fought for already paired females but struggle to defend an exclusive zone of attraction for females. As observed in other anuran species, tenure is thought to be very costly for males (e.g., Wells 1978; Arak 1983b; Cherry 1992; Tejedo 1992a; Murphy 1994). Large body size provides an additional advantage to males in endurance rivalry through the greater energy reserves that can be allocated to the chorus tenure during the breeding season (Tejedo 1992a; Judge & Brooks 2001). We indeed found that males mated once were significantly larger than other males. However, in contrast to Tejedo (1992a), we did not show any significant relationship between male SVL and attendance. Therefore, the hypothesis of advantage in direct male-male struggle to access to females appears most likely.

Nonetheless, we could not reject the potential effect of female choice in driving large body size advantage. Other studies reported female choice for larger males in amphibians (e.g., Morris 1989; Chandler & Zamudio 2008 but see Morrison *et al*. 2001; Zhu et al. 2016). Adaptive mate choice in species lacking male resource control and/or paternal care might be maintained by selection because preferred males sire genetically superior offspring. We observed a significant effect of male SVL on offspring survival which may be a direct effect of assortative mating. The positive association between male SVL and offspring fitness was reported in several taxa including some anuran (Woodward et al. 1988; Luo et al. 2016). As opposed to what can be observed in other taxa (Lewis et al. 2004; Crean *et al*. 2016), there is no paternal care or nuptial gift in the Natterjack toad. Furthermore, mating is external, with no

evidence that nutrients are passed from males to females during mating, and there is no effect of male body size on hatching success. Some authors proposed that female preference for largest males is an indirect choice for older males or males with faster growth, that reflect their genetic quality (Brooks & Kemp 2001). However, heritability of body length is low in bufonids from temperate region. Studies showed that growth was highly dependent of environmental fluctuations and body size is acknowledged to be a poor indicator of individual age in amphibians (Sullivan 1987; Halliday & Verrell 1988). This stochastic variance in body size may decrease the strength of selection on male body size by female choice and may explain the variance observed in mate quality, *i.e*. some female choose largest but not highest quality males and have lower offspring survival due to variation in male inbreeding level and its consequences on offspring metamorphosis success. Thus, we could speculate that the observed difference in offspring survival may be best ascribed to a "good gene" effect independent of male body size (Moller et al. 1999). Such effect may have drove to female preference for larger males thanks to linkage disequilibrium or pleiotropy between genetic basis of adult body length and offspring survival (Kokko & Brooks 2003).

In this respect, we observed a clear size-assortative mating. Assortative mating by size could result from selection on optimized female/male body size ratio that maximizes fertilization success. In 90% of anuran species, females are larger than male, and assortative mating by size could be driven by female preference for larger body length (Harari et al. 1999; Montiglio et al. 2015 but see Fan et al. 2013). In our study, breeding pairs involved females that were on average 2.5 mm (SD = 5.48) larger than males. However, in the whole population, females were not significantly larger than males while size dimorphism was reported in the majority of anuran species (reviewed in Liao et al. 2013). This hypothesis of assortative mating resulting from female preference for larger males is thus unlikely to occur. Finally, in the common toad (*Bufo bufo*) Höglung & Robertson (1988) proposed that a possible direct benefit of mating with a larger male was that these males were better able to resist attacks by rivals and hence females paired with these males were less likely to be in the center of a toad ball. However toad balls are rare in the Natterjack toad (this study, Arak 1983b; Tejedo 1988). Therefore, the two most likely advantages for being a large male are: (1) an increased likelihood of obtaining a mate due to female preference, and (2) increased likelihood of mating because of a greater number of opportunities to obtain a mating thanks to a larger territory (or just being able to defend a territory).

Nonetheless, the males that mated more than once (13%, n=5) were not larger than expected under random mating. This contrasting result therefore questions the reliability of

sexual selection on body size. We could hypothesize (i) that these males may have benefited of nights with low OSR, increasing probability to mate by chance thanks to low male-male competition or (ii) that these males could be advantaged by another traits under sexual selection.

The best or the luckiest?

Interestingly, males that mated more than once had relatively small mating partners (Fig. 4), meaning that they also had lower fecundity. In amphibians, female body length and fecundity are often positively correlated (e.g., Davies & Halliday 1977; Wilbur et al. 1978; Tejedo 1992b; Yu & Lu 2010). Therefore, mating with a large female increases male fitness by increasing the number of sired offspring. Taking into account the fitness advantage to breed with large female, male preference for large female is likely to set up (Andersson & Simmons 2006). Conversely males with low competitiveness should bred with small females less coveted. Nonetheless, the benefit of large males to mate with large females may have its own cost. For instance, Lengagne et al. (2007) showed that embracing a large female costed more than embracing a small female for males Bufo bufo and limited additional opportunities to mate. Our results suggested two reproductive male types: one large, competitive and choosy breeder, one more opportunistic and less choosy breeder. The tradeoff between mate quality and mate quantity may explain why males that mated with largest females did not mate more than once.

Our data did not allow analysis of sexual selection within night. Nevertheless, this result may be crucial to further understand whether body size were under (sexual) selection (Morris 1989). Indeed male competitiveness is dependent on chorus composition within night and female choice is limited by its sampling capacities (Morris 1989; Kokko et al. 2003; Botto & Castellano 2016). Such stochastic fluctuations may decrease the strength of sexual selection and lead to random mating. In particular, the Natterjack toad females are generally only present at the pond the night on which they mate, and usually mate within 30–40 min of their arrival at the pond. Females are thus restricted to sample those males present on a given night, and they are restricted in the period over which they can sample potential mates. It may be costly for female to delay its breeding taking into account the high variation of pond water body, tadpole competitors and male quality within the season (Denton & Beebee 1993a). On the 36 females that spawned, 18 bred after their second visit in the pond and two thirds of whom mated with males that sired two or three females. The social environment characterized by male aggregation may lower selective pressure for the evolution of non-random mating in respect to body size by favoring males with opportunistic behaviour (e.g., Lu et al. 2009). The balanced

effect of each chorus-dependent strategy would result in such assortative mating and contrasting successful breeding patterns.

Individual features and breeding success: inbreeding level

Heterozygosity is often assumed to be positively correlated to individual fitness and overall, to population fitness (e.g., in amphibians: Rowe *et al*. 1999 but see Rogell et al. 2010; in bird and mammals: Amos et al. 2001, Huisman et al., 2016). Positive heterozygosity-fitness correlation arise as a consequence of inbreeding (Szulkin et al. 2010). In several taxa inbreeding is acknowledged to decrease fitness component such as offspring growth and development (reviewed in Keller & Waller 2002; Hedrick & Garcia-Dorado 2016). Effect of parent relatedness on offspring hatching, development or survival has been shown in several studies (e.g., Frère et al. 2010; Huisman et al. 2016). Here, male and female level of inbreeding significantly decreased the metamorphosis success. Moreover, we observed a significant negative effect of inbreeding on male mating success and clutch hatching rate. To the best of our knowledge, we are among the first to demonstrate that inbreeding depression could act at later stages (adults) by directly affecting adult probability to mate (see also Richter & Nunziata 2014). Inbreeding depression, which results from increasing genome-wide homozygosity could lead to population extinction (Hedrick & Garcia-Dorado 2016). Our data suggest that more inbred males had lower probability to mate than less inbred males. Such a pattern might imply lower competitiveness of inbred males (e.g., low "endurance rivalry") or maybe female avoidance of inbred mates. Inbreeding avoidance is thought to be a key driver in the evolution of mating systems through different evolutionary processes (Pusey & Wolf 1996). Inbreeding avoidance might evolve when the cost of mating with an inbred partner exceeds that of inbreeding avoidance (Pusey & Wolf 1996; Kokko & Ots 2006). In amphibian two main mechanisms may act: pre- or post-copulatory selection of non-genetically related males by female mate choice (e.g., Jehle et al. 2007) and, dispersal (reviewed in Waldman & McKinnon 1995). Kin recognition has been already proved in amphibian larvae, by demonstrating kin aggregation or repulsion in the wild (Halverson et al. 2006). Blaustein & Waldman (1992) suggested that it may facilitate outbreeding in adults. In the Natterjack toad no empirical data allow us to firmly support such a conclusion. Experimental data are needed to explore the evolutionary forces which could disadvantage reproduction of the most inbred males.

The rise and fall of sexual selection facing natural selection

As previously discussed, we observed variance in offspring fitness linked to parent traits potentially under direct sexual selection (male body), undirected sexual selection (male mating success and mean OSR) or under natural selection (internal relatedness). These patterns result in variance in mate quality and variance in breeding success that greatly increased through the different stage sof tadpole development (explain). The true (expected) quality of a mate is difficult to separate from stochastic processes that influence its realized reproductive success (Henshaw et al. 2016). Hence our result brings out the puzzling evolutionary mechanisms leading to the established mating system by unbalanced effect of sexual and natural selection on individual fitness. In addition, the effect of inbreeding depression on early and late stages questions the potential of sexual and natural selection to purge deleterious alleles (Hedrick 1994; Keller & Waller 2002). Survival and mating success of individuals with greater genetic variability should prolong the persistence of this geographically isolated population (e.g., Richter & Nunziata 2014)

Conclusion

We provided evidences for opportunity for sexual selection in the Natterjack toad. Evolutionary processes at work remain puzzling. We observed an advantage to largest males to access to mate and a positive association of male body size, male mating success and the intensity of competition (mean OSR per night of attending) with offspring quality that highlight the potential for female choice and male-male competition. However largest males mated with largest and more fecund females, but did mate only once whereas males that sired smaller females mated more than once suggesting a trade-off between mate quality and mate quantity. On the other hand, we showed a significant effect of inbreeding level on mating success and overall, on breeding success. Our study is one of the first demonstrating a significant effect of inbreeding level on male mating success. We also observed inbreeding depression in early stage of developments. Such decreased fitness of inbred individuals could purge the genetic load (Hedrick 1994). In species with first stages of development highly dependent to environmental factors, strength of selection could be lowered by random/unpredictable stochastic events, like frequent pond desiccation and fluctuating density of competitors (Griffiths 1997). As a result, selection against inbred males may represent a critical evolutionary mechanism acting to purge genetic load in species facing harsh stochastic events. Understanding the relative impact of stochastic factors as density of population, family structure and in situ breeding success (number of emerging toadlets) is a major issue to better infer the potential effect of demographic

and/or environmental fluctuation on population mating system and consequently population dynamics. A better understanding of this efficiency of potential selection against inbreeding would greatly contribute to a better awareness of the threat that inbreeding may constitute in small isolated populations. Comparative studies on different species or monitoring populations during several years are needed to better quantify the impact of such stochastic fluctuations on the strength of sexual selection.

Acknowledgements

We are very grateful to Virginie Callipel for describing and giving free full access to the study site, Cédric Vanappelghem and Stéphanie Rondel for helpful discussions. This work was funded by the AMPHIDIV project (Région Nord-Pas De Calais). This work is also a contribution to the CPER research project CLIMIBIO. The authors thank the French Ministry for Higher Education and Research, the Hauts de France Regional Council and the European Regional Development Fund for their financial support of this project.

Author contributions

LF, JJ, TB and JFA conceived and designed the study. LF, FB, BF and JFA conducted the field monitoring, LF and FB carried out the rearing of tadpoles, LF and CG genotyped the samples, LF, JJ and TB analyzed the data, LF and JFA wrote the manuscript. All authors provided editorial comments.

References

Amos W, Wilmer JW, Fullard K *et al.* (2001) The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London, Series B*, **268**, 2021–2027.

Andersson M, Iwasa Y. (1996) Sexual selection. *Trends in Ecology and Evolution*, **11**, 53–58.

Andersson M, Simmons LW. (2006) Sexual selection and mate choice. *Trends in Ecology and Evolution*, **21**, 296–302.

Arak A. (1983a) Male-male competition and mate choice in anuran amphibians. In: *Mate choice* (ed Bateson P), pp. 181–210. Cambridge University Press. Cambridge.

Arak A. (1983b) Sexual selection by male–male competition in natterjack toad choruses. *Nature*, **306**, 261–262.

Arak A. (1988a) Female mate selection in the natterjack toad : active choice or passive attraction ? *Behavioral Ecology and Sociobiology*, **22**, 317–327.

Arak A. (1988b) Callers and satellites in the natterjack toad: evolutionarily stable decision rules. *Animal Behaviour*, **36**, 416–432.

Arnold SJ, Wade MJ. (1984) On the measurement of natural and sexual selection: theory. *Evolution*, **38**, 709–719.

Bartoń K. (2016) MuMIn: Multi-Model Inference. R version 1.15.6.

Bateman AJ. (1948) Intra-sexual selection in *Drosophila*. *Heredity*, **2**, 349–368.

Bates D, Mächler M, Bolker B, Walker S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1–48.

Beebee TJC. (1979) A review of scientific information pertaining to the Natterjack toad *Bufo calamita* throughout its geographical range. *Biological Conservation*, **16**, 107–134.

Beebee TJC, Denton JS. (1996) *The natterjack toad conservation handbook (Species recovery program)*. English Nature.Northminster House, Peterborough PE1 1UA.

Blaustein AR, Waldman B. (1992) Kin recognition in anuran amphibians. *Animal Behaviour*, **44**, 207–221.

Bolker BM, Brooks ME, Clark CJ *et al.* (2008) Generalized linear mixed models : a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, **24**, 127–135.

Böll S, Linsenmair KE. (1998) Size-dependent male reproductive success and size-assortative mating in the midwife toad *Alytes obstetricans*. *Amphibia-Reptilia*, **19**, 75–89.

Botto V, Castellano S. (2016) Signal reliability and multivariate sexual selection in a lek-breeding amphibian. *Behavioral Ecology*, **27**, 1797–1807.

Bowcock H, Brown GP, Shine R. (2013) Sexual selection in cane toads Rhinella marina: a male's body size affects his success and his tactics. *Current Zoology*, **59**, 747–753.

Brooks R, Kemp DJ. (2001) Can older males deliver the good genes? *Trends in Ecology and Evolution*, **16**, 308–313.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L. (2006) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics*, **8**, 509–511.

Chandler CH, Zamudio KR. (2008) Reproductive success by large, closely related males facilitated by sperm storage in an aggregate breeding amphibian. *Molecular Ecology*, **17**, 1564–1576.

Cherry MI. (1992) Sexual selection in the leopard toad, *Bufo pardalis*. *Behaviour*, **120**, 164–176.

Clutton-Brock TH. (2007) Sexual selection in males and females. *Science*, **318**, 1882–1885.

Coulon A. (2009) GENHET: an easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources*, **10**, 167–169.

Crean AJ, Adler MI, Bonduriansky R. (2016) Seminal fluid and mate choice. *Trends in Ecology and Evolution*, **31**, 253–255.

Danchin E, Cézilly F. (2008) Sexual selection: another evolutionary process. In: *Behavioural Ecology* (eds Danchin E, Giraldeau L-A, Cézilly F), pp. 363–426. Oxford University Press.Oxford, UK.

Darwin C. (1871) *The descent of man, and selection in relation to sex*. Princeton, N.J. :Princeton University Press.London.

Davies NB, Halliday TR. (1977) Optimal mate selection in the toad *Bufo bufo*. *Nature*, **269**, 56–58.

Denton JS, Beebee TJC. (1993a) Reproductive strategies in a female-biased population of natterjack toads, *Bufo calamita*. *Animal Behaviour*, **46**, 1169–1175.

Denton JS, Beebee TJC. (1993b) Density-related features of natterjack toad (*Bufo calamita*) population in Britain. *Journal of Zoology*, **229**, 105–119.

Denton JS, Hitchings SP, Beebee TJC, Gent A. (1997) A recovery program for the natterjack toad (*Bufo calamita*) in Britain. *Conservation Biology*, **11**, 1329–1338.

Van Doorn GS, Edelaar P, Weissing FJ. (2009) On the origin of species by natural and sexual selection. *Science*, **326**, 1704–1707.

Emlen ST. (1976) Lek organization and mating strategies in the bullfrog. *Behavioral Ecology and Sociobiology*, **1**, 283–313.

Emlen ST, Oring LW. (1977) Ecology, sexual selection, and the evolution of mating systems. *Science*, **197**, 215–223.

Fan XL, Lin ZH, Ji X. (2013) Male size does not correlate with fertilization success in two bufonid toads that show size-assortative mating. *Current Zoology*, **59**, 740–746.

Faucher L, Godé C, Arnaud J-F. (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms for the natterjack toad, *Bufo* (*Epidalea*) *calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). *Journal of Heredity*, **107**, 660–665.

Faucher L, Hénocq L, Vanappelghem C *et al.* (2017) When new human-modified habitats favour the expansion of an amphibian pioneer species: evolutionary history of the natterjack toad (*Bufo calamita*) in a coal basin. *Molecular Ecology*, **26**, 4434–4451.

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

Frere CH, Krutzen M, Kopps AM *et al.* (2010) Inbreeding tolerance and fitness costs in wild bottlenose dolphins. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 2667–2673.

Gosner KL. (1960) A simplified table for staging anuran embryos larvae with notes on identification. *Herpetologica*, **16**, 183–190.

Goudet J. (1995) FSTAT (Version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Griffiths RA. (1997) Temporary ponds as amphibian habitats. *Aquatic conservation: marine and freshwater ecosystems*, **7**, 119–126.

Halliday TR, Verrell PA. (1988) Body size and age in amphibians and reptiles. *Journal of Herpetology*, **22**, 253–265.

Halverson MA, Skelly DK, Caccone A. (2006) Kin distribution of amphibian larvae in the wild. *Molecular Ecology*, **15**, 1139–1145.

Harari AR, Handler AM, Landolt PJ. (1999) Size-assortative mating, male choice and female choice in the curculionid beetle Diaprepes abbreviatus. *Animal behaviour*, **58**, 1191–1200.

Hedrick PW. (1994) Purging inbreeding depression and the probability of extinction : full-sib mating. *Heredity*, **73**, 363–372.

Hedrick PW, Garcia-Dorado A. (2016) Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology and Evolution*, **31**, 940–952.

Henshaw JM, Kahn AT, Fritzsche K. (2016) A rigorous comparison of sexual selection indexes via simulations of diverse mating systems. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E300–E308.

Höglund J, Alatalo R. (1995a) Amphibians. In: *Leks* (eds Höglund J, Alatalo R V.), pp. 30–32. Princeton University Press.Princeton (NJ).

Höglund J, Alatalo R. (1995b) *Leks* (J Höglund, R Alatalo, Eds,). Princeton University Press. Princeton (NJ).

Höglung J, Robertson GM. (1988) Chorusing behaviour, a density-dependent alternative mating strategy in male common toads (*Bufo bufo*). *Ethology*, **79**, 324–332.

Hosken DJ, House CM. (2011) Sexual selection. *Current Biology*, **21**, R62–R65.

Howard RD. (1979) Estimating reproductive success in natural populations. *The American Naturalist*, **114**, 221–231.

Huisman J, Kruuk LEB, Ellis PA, Clutton-Brock TH, Pemberton JM. (2016) Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 3585–3590.

Hurst LD. (2009) Genetics and the understanding of selection. *Nature Reviews Genetics*, **10**, 83–93.

Jaquiéry J, Broquet T, Aguilar C, Evanno G, Perrin N. (2010) Good genes drive female choice for mating partners in the lek-breeding European treefrog. *Evolution*, **64**, 108–115.

Jehle R, Sztatecsny M, Wolf JBW *et al.* (2007) Genetic dissimilarity predicts paternity in the smooth newt (*Lissotriton vulgaris*). *Biology Letters*, **3**, 526–528.

Jennions MD, Kokko H. (2010) Evolutionary Behavioral Ecology. In: *Sexual selection* (eds Westneat DF, Fox CW), pp. 343–364. Oxford University Press.New York, USA.

Jirotkul M. (1999) Operational sex ratio influences female preference and male – male competition in guppies. *Animal Behaviour*, **58**, 287–294.

Johnston SE, Gratten J, Berenos C *et al.* (2013) Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*, **502**, 93–95.

Jones AG. (2009) On the opportunity for sexual selection, the bateman gradient and the maximum intensity of sexual selection. *Evolution*, **63**, 1673–1684.

Jones AG, Arguello JR, Arnold SJ. (2004) Molecular parentage analysis in experimental newt populations: the response of mating system measures to variation in the operational sex ratio. *The American Naturalist*, **164**, 444–456.

Jones OR, Wang J. (2010) COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551–555.

Judge KA, Brooks RJ. (2001) Chorus participation by male bullfrogs, *Rana catesbeiana*: a test of the energetic constraint hypothesis. *Animal Behaviour*, **62**, 849–861.

Kalinowski ST, Taper ML, Marshall TC. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

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

Kluge AG. (1981) The life history, social organization, and parental behavior of *Hyla rosenbergi* Boulenger, a nest-building gladiator frog. *Publications of the Museum of Zoology of the University of Michigan*, **160**, 1–170.

Kokko H, Brooks R. (2003) Sexy to die for? Sexual selection and the risk of extinction. *Annales Zoologici Fennici*, **40**, 207–219.

Kokko H, Brooks R, Jennions MD, Morley J. (2003) The evolution of mate choice and mating biases. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 653–664.

Kokko H, Ots I. (2006) When not to avoid inbreeding. *Evolution*, **60**, 467–475.

Kraaijeveld K, Kraaijeveld-Smit FJL, Maan ME. (2011) Sexual selection and speciation: the comparative evidence revisited. *Biological Reviews*, **86**, 367–377.

Kvarnemo C, Ahnesjö I. (1996) The dynamics of operational sex ratios and competition for mates. *Trends in Ecology and Evolution*, **11**, 404–408.

Lande R. (1977) The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics*, **86**, 485–498.

Lefcheck JS. (2016) PiecewiseSEM: Piecewise structural equation modeling.

Lewis SM, Cratsley CK, Rooney JA. (2004) Nuptial gifts and sexual selection in *Photinus fireflies*. *Integrative and Comparative Biology*, **44**, 234–237.

Liao WB, Zeng Y, Yang JD. (2013) Sexual size dimorphism in anurans: roles of mating system and habitat types. *Frontiers in Zoology*, **10**, 65.

Lisle SP De, Rowe L. (2015) Independent evolution of the sexes promotes amphibian diversification. *Proceedings of the Royal Society of London, Series B*, **282**, 20142213.

Lu X, Ma X, Li Y, Fan L. (2009) Breeding behavior and mating system in relation to body size in *Rana chensinensis*, a temperate frog endemic to northern China. *Journal of Ethology*, **27**, 391–400.

Luo Z, Li C, Wang H *et al.* (2016) Male-male competition drives sexual selection and group spawning in the Omei treefrog, *Rhacophorus omeimontis*. *Behavioral Ecology and Sociobiology*, **70**, 593–605.

Lüpold S, Jin L, Liao WB. (2017) Population density and structure drive differential investment in pre- and postmating sexual traits in frogs. *Evolution*, **71**, 1686–1699.

Mank JE. (2009) Sexual selection and Darwin's mystery of mysteries. *Science*, **326**, 1639–1640.

McFadden D. (1974) Conditional logit analysis of qualitative choice behavior. In: *Frontiers in Econometrics*, pp. 105–142. Academic Press.New-York.

Meyer D, Zeileis A, Hornik K. (2016) vcd: Visualizing Categorical Data. *R package version 1.4-3*.

Mobley KB, Jones AG. (2013) Overcoming statistical bias to estimate genetic mating systems in open populations: a comparison of bateman's principles between the sexes in a sex-role-reversed pipefish. *Evolution*, **67**, 646–660.

Moller AP, Alatalo R V, Möller AP *et al.* (1999) Good-genes effects in sexual selection. *Proceedings of the Royal Society B: Biological Sciences*, **266**, 85–91.

Montiglio P-O, Wey TW, Chang AT, Fogarty S, Sih A. (2016) Multiple mating reveals complex patterns of assortative mating by personality and body size. *Journal of Animal Ecology*, **85**, 125–135.

Morris MR. (1989) Female choice of large males in the treefrog *Hyla chrysoscelis*: the importance of identifying the scale of choice. *Behavioral Ecology and Sociobiology*, **25**, 275–281.

Morrison C, Hero J, Smith WP. (2001) Mate selection in *Litoria chloris* and *Litoria xanthomera* : females prefer smaller males. *Austral Ecology*, **26**, 223–232.

Murphy CG. (1994) Determinants of chorus tenure in barking treefrogs (*Hyla gratiosa*). *Behavioral Ecology and Sociobiology*, **34**, 285–294.

Murphy CG. (1998) Interaction-independent sexual selection and the mechanisms of sexual selection. *Evolution*, **52**, 8–18.

Nakagawa S, Schielzeth H. (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Nei M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press.New-York.

Orr HA. (2009) Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, **10**, 531–539.

Owens IPF, Burke T, Thompson DBA. (1994) Extraordinary sex roles in the Eurasian dotterel: female mating arenas, female-female competition, and female mate choice. *The American Naturalist*, **144**, 76–100.

Pröhl H. (2002) Population differences in female resource abundance , adult sex ratio , and male mating success in *Dendrobates pumilio*. *Behavioral Ecology*, **13**, 175–181.

Pusey A, Wolf M. (1996) Inbreeding avoidance in animals. *Trends in Ecology and Evolution*, **11**, 201–206.

Richter SC, Nunziata SO. (2014) Survival to metamorphosis is positively related to genetic variability in a critically endangered amphibian species. *Animal Conservation*, **17**, 265–274.

Ritchie MG. (2007) Sexual selection and speciation. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 79–102.

Rogell B, Gyllenstrand N, Höglund J. (2005) Six polymorphic microsatellite loci in the Natterjack toad, *Bufo calamita*. *Molecular Ecology Notes*, **5**, 639–640.

Rogell B, Thörngren H, Laurila A, Höglund J. (2010) Fitness costs associated with low genetic variation are reduced in a harsher environment in amphibian island populations. *Conservation Genetics*, **11**, 489–496.

Rousset F, Raymond M. (1995) Testing heterozygote excess. *Genetics*, **140**, 1413–1419.

Rowe G, Beebee TJC, Burke T. (1997) PCR primers for polymorphic microsatellite loci in the anuran amphibian *Bufo calamita*. *Molecular Ecology*, **6**, 401–402.

Rowe G, Beebee TJC, Burke T. (1999) Microsatellite heterozygosity, fitness and demography in natterjack toads *Bufo calamita*. *Animal Conservation*, **2**, 85–92.

Rowe G, Beebee TJC, Burke T. (2000) A microsatellite analysis of natterjack toad, *Bufo calamita*, metapopulations. *Oikos*, **88**, 641–651.

Seltmann P, Cocucci A, Renison D, Cierjacks A, Hensen I. (2009) Mating system, outcrossing distance effects and pollen availability in the wind-pollinated treeline species *Polylepis australis* BITT. (Rosaceae). *Basic and Applied Ecology*, **10**, 52–60.

Sinsch U. (1988) Temporal spacing of breeding activity in the natterjack toad, *Bufo calamita*. *Oecologia*, **76**, 399–407.

Sinsch U. (1997) Postmetamorphic dispersal and recruitment of first breeders in a *Bufo calamita* metapopulation. *Oecologia*, **112**, 42–47.

Sullivan BK. (1982) Sexual selection in Woodhouse's toad (*Bufo woodhousei*) I. Chorus organization. *Animal Behaviour*, **30**, 680–686.

Sullivan BK. (1987) Sexual selection in Woodhouse's toad (*Bufo woodhousei*). III. Seasonal variation in male mating success. *Animal Behaviour*, **35**, 912–919.

Sullivan BK, Kwiatkowski MA. (2007) Courtship displays in anurans and lizards: theoretical and empirical contributions to our understanding of costs and selection on males due to female choice. *Functional Ecology*, **21**, 666–675.

Szulkin M, Bierne N, David P. (2010) Heterozygosity-fitness correlations: a time for reappraisal. *Evolution*, **64**, 1202–1217.

Tejedo M. (1988) Fighting for females in the toad *Bufo calamita* is affected by the operational sex ratio. *Animal Behaviour*, **36**, 1765–1769.

Tejedo M. (1992a) Large male mating advantage in natterjack toads, Bufo calamita: sexual selection or energetic constraints? *Animal Behaviour*, **44**, 557–569.

Tejedo M. (1992b) Effects of body size and timing of reproduction on reproductive success in female natterjack toads (*Bufo calamita*). *Journal of Zoology*, **228**, 545–555.

Trivers RLL. (1985) *Social Evolution*. Benjamin Cummings.Menlo Park, CA.

Ursprung E, Ringler M, Jehle R, Hödl W. (2011) Strong male/male competition allows for nonchoosy females: high levels of polygynandry in a territorial frog with paternal care. *Molecular Ecology*, **20**, 1759–1771.

Waldman B, McKinnon JS. (1995) Inbreeding and outbreeding in fishes, amphibians, reptiles. In: *The natural history of inbreeding and outbreeding- Theoritical and Empirical Perspectives* (ed Thornhill NW), pp. 250–268. The University of Chicago Press.Chicago, USA.

Wang Q, He XZ, Yang L, Hedderley D, Davis LK. (2009) Biased sex ratio and low population density increase male mating success in the bug *Nysius huttoni* (Heteroptera : Lygaeidae). *Naturwissenschaften*, **96**, 87–91.

Wells KD. (1977) The social behaviour of anuran amphibians. *Animal Behaviour*, **25**, 666–693.

Wells KD. (1978) Territoriality in the green frog (*Rana clamitans*): vocalizations and agonistic behaviour. *Animal Behaviour*, **26**, 1051–1063.

Wells KD. (2010) *The ecology and behavior of amphibians*. University of Chicago Press.Chicago, USA.

Wilbur HH, Rubinstein DI, Fairchild L. (1978) Sexual selection in toads- the roles of female choice and male body size. *Evolution*, **32**, 264–270.

Wilkinson JW, Griffiths RA. (2013) Amphibian declines and conservation in Britain. In: *Amphibian biology. Status of conservation and decline of amphibians: Eastern hemisphere. Western Europe* (eds Heatwole H, Wilkinson JW), pp. 48–57. Pelagic Publishing.Exeter, UK.

Woodward BD, Travis J, Mitchell SL. (1988) The effects of the mating system on progeny performance in *Hyla crucifer* (Anura: Hylidae). *Evolution*, **42**, 784–794.

Wright S. (1931) Evolution in Mendelian Populations. *Genetics*, **16**, 97–159.

Yu TL, Lu X. (2010) Sex recognition and mate choice lacking in male Asiatic toads (*Bufo gargarizans*). *Italian Journal of Zoology*, **77**, 476–480.

Zeileis A, Hothorn T. (2002) Diagnostic checking in regression relationships. *R News*, **2**, 7–10.

Zhao M, Li C, Zhang W *et al.* (2016) Male pursuit of higher reproductive success drives female polyandry in the Omei treefrog. *Animal Behaviour*, **111**, 101–110.

Zhu B, Wang J, Zhao L *et al.* (2016) Bigger is not always better: females prefer males of mean body size in Philautus odontotarsus. *PLoS ONE*, **11**, e0149879

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM. (2009) *Mixed effects models and extensions in ecology with R*. Springer New York.New York, USA

Supplementary files

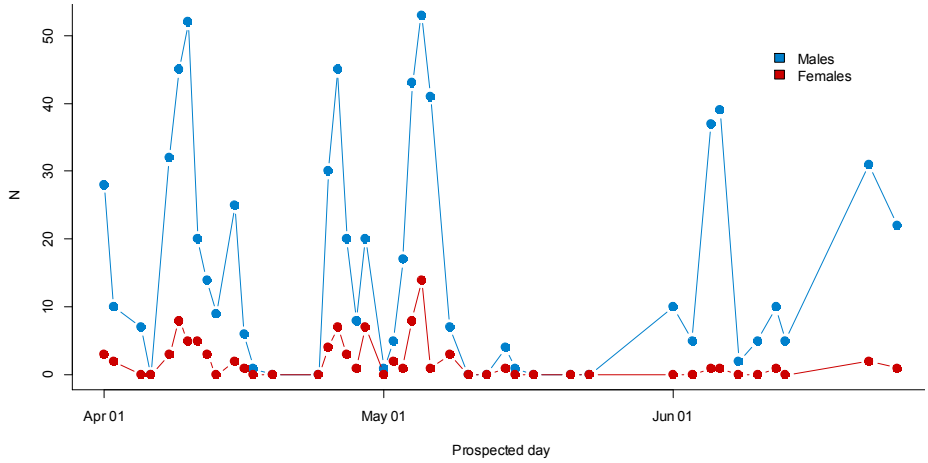


Figure S1: Adults attendance from 1st April to 25 June in a Natterjack toad (*B. calamita)* population in northern France.

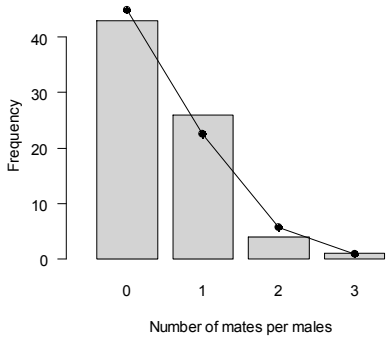


Figure S2: Males mating success distribution (grey barplots) and expected Poisson distribution (black dots) in a *B. calamita* population in northern France.

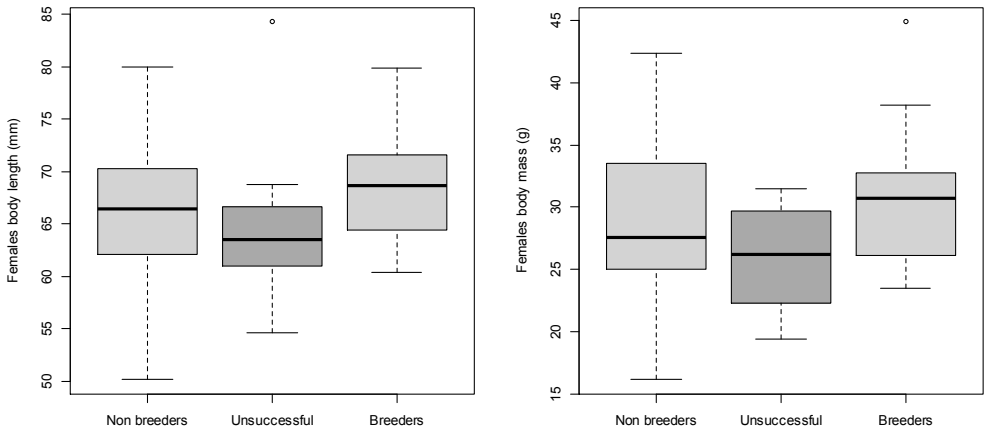


Figure S3: Female body length (Snout to vent length, SVL) and female body mass in each females categories of breeders in the Natterjack toad (*B. calamita*). Non-breeders were never saw in the pond whereas unsuccessful females were observed at least once in the pond but did not lay eggs. Successful female breeders produced at least one clutch during the breeding season.

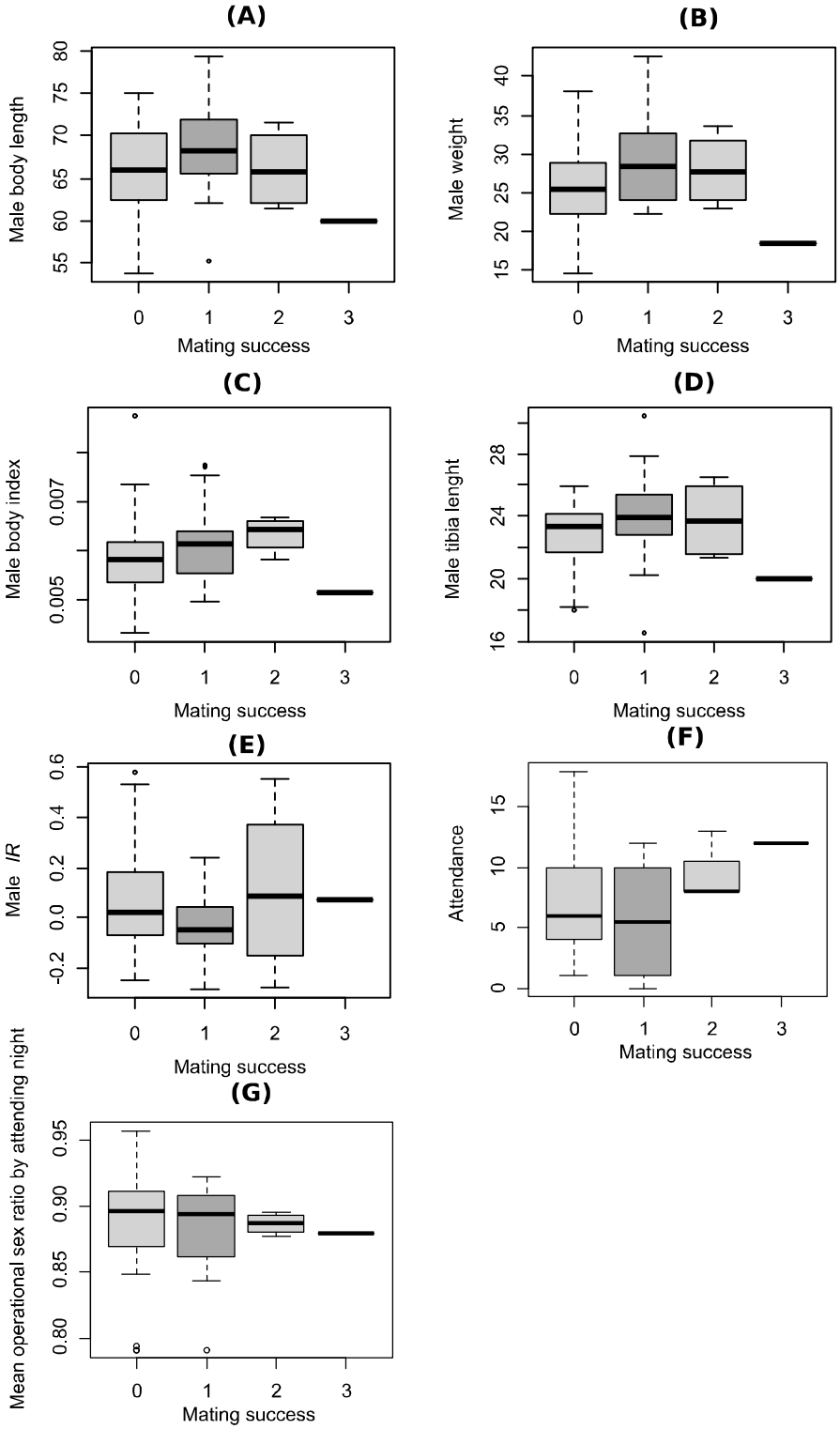


Figure S4: Boxplots displaying morphological and genetic traits as a function of male mating success (number of mates): (A) body length (snout to vent length), (B) body mass; (C) body mass index (body mass/ body length²), (D) tibia length, (E) internal relatedness, (F) attendance (G) mean operational sex-ratio per attending night. No significant differences were observed among the four categories using Kruskal-Wallis tests.

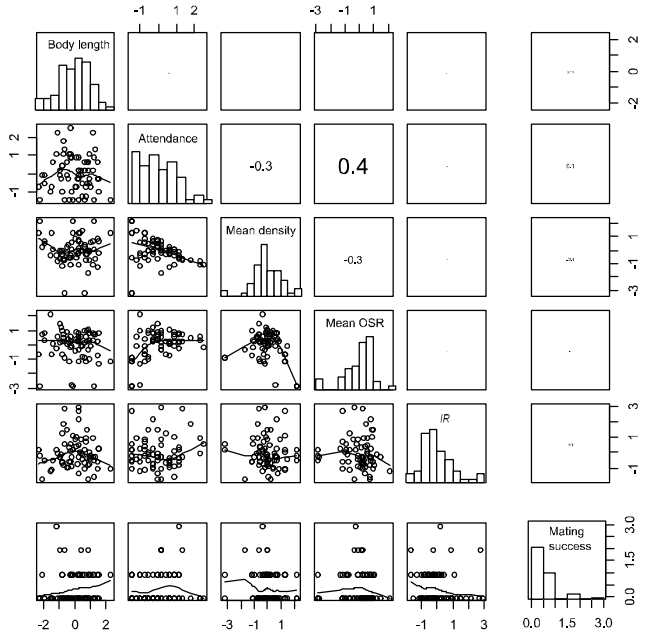


Figure S5: Graphs illustrating the distributions of centered reduced data (diagonal) and collinearity of variables between them (dot chart in the lower part of the matrix), and Pearson's correlation coefficient (in the upper part of the matrix) of explanatory variables included in zero inflated model and logistic regression on male mating success (Models A and B).

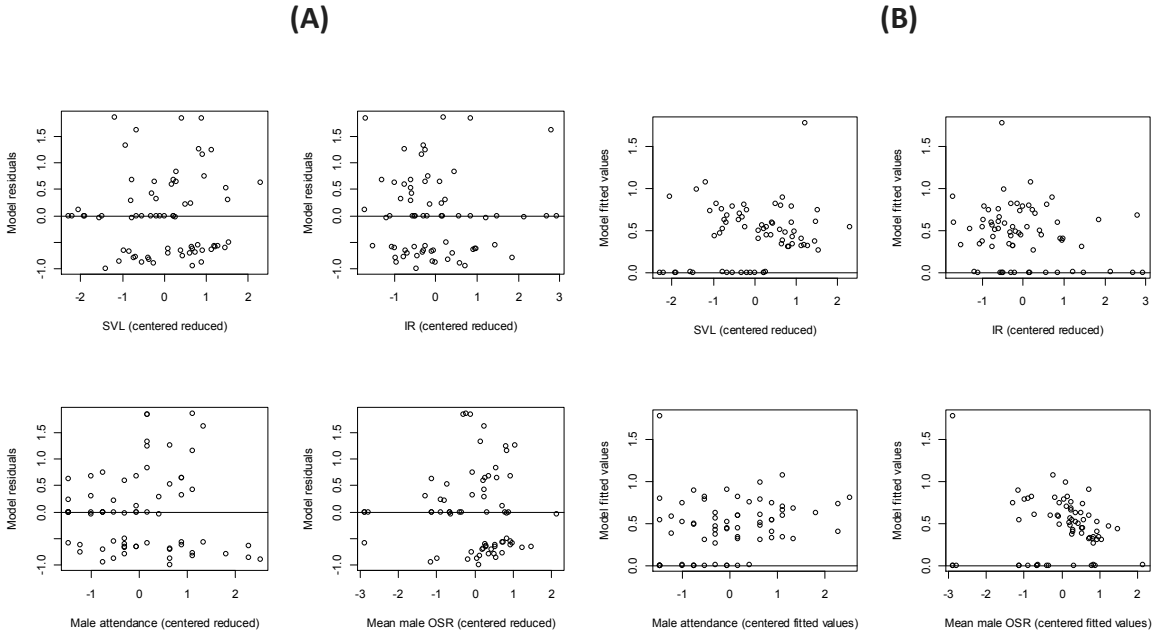


Figure S6: Cross validation graphs with (A) residuals of zero inflated model (model A) plotted against each explanatory variable and (B) fitted values plotted against each explanatory variable.

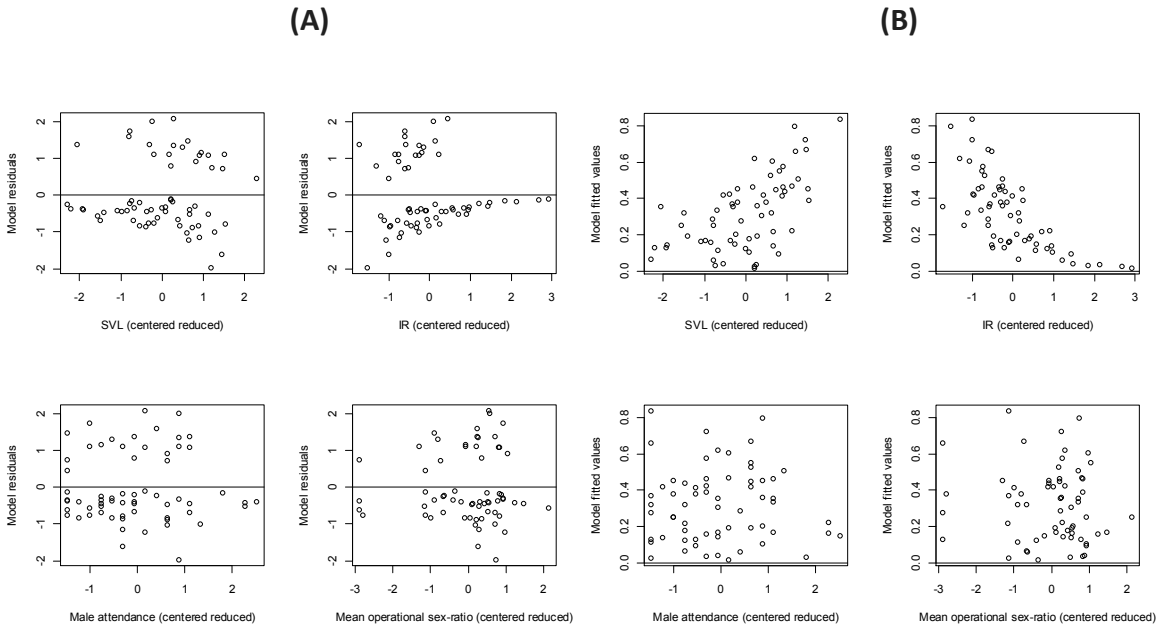


Figure S7: Cross validation graphs with (A) residuals of logistic regression (model B) plotted against each explanatory variable and (B) fitted values plotted against each explanatory variable.

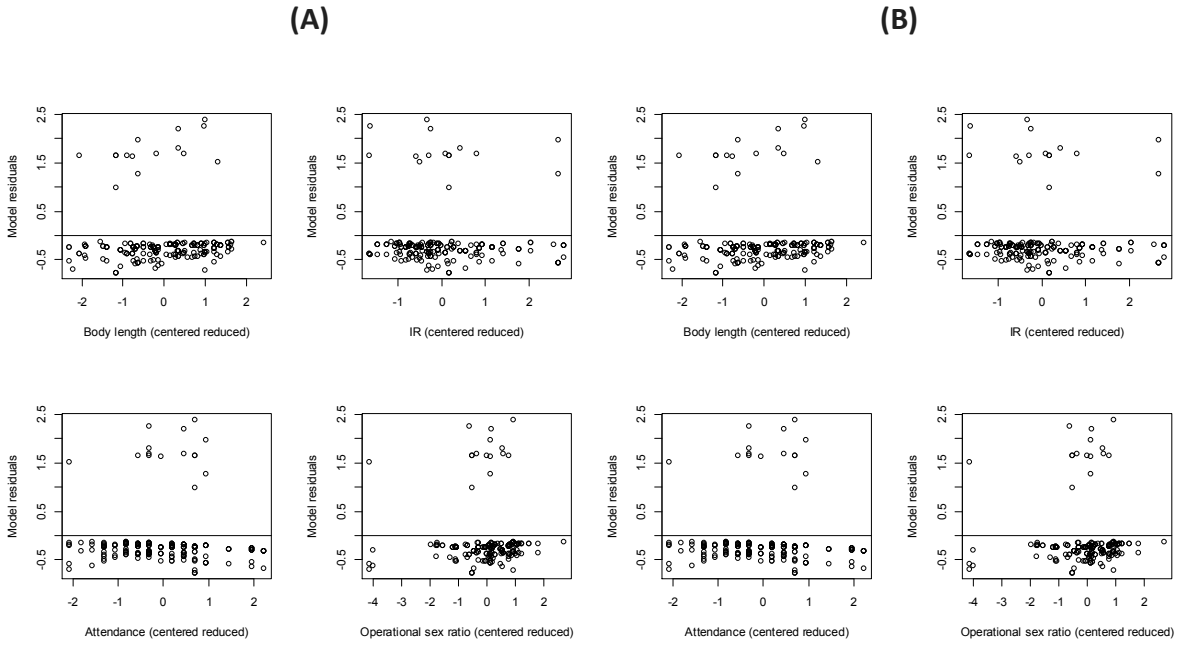


Figure S8: Cross validation graphs with (A) residuals of generalized linear mixed effect model (model C) plotted against each explanatory variable and (B) fitted values plotted against each explanatory variable.

Table S1: Description of traits recorded on adults (A) and offspring (B) in a *Bufo [Epidalea] calamita* population located in northern France

	Field	Variable name	Description
(A) ADULTS	Breeding success	Mating success	Number of mates
		Clutch size	Numbre of eggs sired
	Behavioural traits	Amplexus	Number of amplexus observed for this individual
		Partner_amplexus	identity of partner in amplexus
		Partner_clutch	identity of partner for reproduction
		Competitor	identity of the true male competitor that reproduced with the female in place of the male seen in amplexus
		Date_FirstCapture	Date of the measure of morphological traits
		Attendance	Number of nights where the individual was seen
		MeanDensity	Mean number of males observed the nights where the individual was present
		MeanOSR	Mean operational sex-ratio the nights where the individual was present
	Morphological traits	SVL	Snout to vent length (mm)
		Tibia	Tibia length (mm)
		Body mass	Body mass (g)
	Genetic traits	IR	Internal relatedness (a measure of inbreeding level)
(B) FFSPRIN	Field data	Laying.date	Date of clutch laying
		Pond	Lawing site
		OSR	Operational sex-ratio the night where the mother layed clutch
		ClutchSize	Clutch size
	Data for standardization	Neggssampled	Number of sampled eggs
	Hatching	Eclosion period	Hatching date
		NHatched	Number of hatched eggs
	Data for standardization	Stage_D0	Tadpole stage the first day of fitness survey
	Survival	N_D0	Number of surviving tadpoles at day 1
		N_D11	Number of surviving tadpoles at day 11
		N_D25	Number of surviving tadpoles at day 25
		N_D31	Number of surviving tadpoles at day 31
		N_D50	Number of surviving tadpoles at day 50
		Ntadpoles_D45	Number of surviving tadpoles at day 45
		Ntadpoles_D60	Number of surviving tadpoles at day 60
		Ntadpoles_D70	Number of surviving tadpoles at day 70
	Metamorphosis	X1stmetamorphose	Number of days until the first metamorphosis
		X15thmetamorphose	Number of days until the 15th metamorphosis
		Nmetam_D45	Number of metamorphosed individuals at day 45
		Nmetam_D60	Number of metamorphosed individuals at day 60
		Nmetam_D70	Number of metamorphosed individuals at day 70
		Last metamorphosis period	Number of days until the last metamorphosis

Growth	M_bodymass15D11	Mean tadpoles body mass at day 11 (measured on around 15 tadpoles per clutch)
	M_ bodymass15D25	Mean tadpoles body mass at day 25 (measured on around 15 tadpoles per clutch)
	V_ bodymass 15D11	Variance of tadpoles body mass t at day 11 (measured on around 15 tadpoles per clutch)
	V_ bodymass 15D25	Variance of tadpoles body mass at day 25 (measured on around 15 tadpoles per clutch)

Table S2: Nuclear genetic variation for 37 nuclear microsatellite loci over 113 adults and 667 tadpoles (from 37 sibships) in *Bufo [Epidalea] calamita*: expected heterozygosity (H_E), total number of alleles (A_T), and mean intra-population fixation indice (F_{IS}, according to Weir & Cockerham (1984). Departures from Hardy-Weinberg expectations were tested using the multiscore sample *U*-test and corrected using Bonferroni adjustment.* $P < 0.05$; **; $P < 0.01$; *** $P < 0.001$

				ADULTS			PROGENY		
	Publication*	Multiplex	Dye	H_E	A_T	F_{IS}	H_E	A_T	F_{IS}
BC05	1	1	NED	0.373	3	-0.020	0.372	4	-0.075 *
BC08	1	1	HEX	0.658	6	0.254 ***	0.640	5	-0.156 **
BC11	1	1	HEX	0.763	9	-0.032	0.760	9	-0.044
BC19	1	1	FAM	0.404	2	-0.030	0.395	2	-0.030
BC22	1	6	FAM	0.643	4	0.028	0.658	4	0.137 ***
BC29	1	6	NED	0.578	4	0.081 *	0.576	4	0.182 ***
BC38	1	6	PET	0	1	NA	0	1	NA
BC46	1	6	FAM	0.514	4	0.157 *	0.510	5	0.100
BC09	1	8	NED	0.493	5	-0.060	0.508	5	-0.017
BC15	1	8	FAM	0.277	4	-0.054	0.255	4	-0.091 **
BC18	1	8	FAM	0.819	6	-0.070 *	0.804	6	0.101 **
BC37	1	8	HEX	0.717	5	-0.055	0.717	4	0.014
BC39	1	8	HEX	0.777	6	-0.082	0.749	6	-0.012
BC45	1	8	PET	0.822	7	-0.043	0.833	7	0.070 ***
ALL				/	/	/	/	/	0.023 ***
Bcal7	2	1	NED	0.223	3	-0.068	/	/	/
Bcal1	2	1	FAM	0.790	6	-0.063	/	/	/
Bcal2	2	1	PET	0.494	3	0.296 **	/	/	/
Bcal11	2	2	FAM	0.465	3	0.016	/	/	/
Buca2	3	2	VIC	0.571	3	0.017	/	/	/
Buca3	3	2	PET	0	1	NA	/	/	/
Buca5	3	2	NED	0.163	3	-0.082	/	/	/
Bcal3	2	3	FAM	0.667	4	0.329 ***	/	/	/
Bcal4	2	3	PET	0.472	3	-0.059	/	/	/
Bcal5	2	3	FAM	0.616	3	-0.058	/	/	/
Bcal6	2	3	VIC	0.792	5	-0.001	/	/	/
Bcal10	2	4	NED	0.681	5	-0.011	/	/	/
Bcal9	2	4	FAM	0.550	4	0.401 ***	/	/	/
Buca1	3	5	NED	0.422	3	0.114	/	/	/
Buca6	3	4	FAM	0.396	4	0.009 **	/	/	/
BC01	1	7	FAM	0.622	5	-0.109 *	/	/	/
BC02	1	7	FAM	0.514	5	0	/	/	/
BC04	1	7	NED	0.527	5	0.077	/	/	/
BC24	1	7	NED	0.627	3	0.189 **	/	/	/
BC25	1	7	PET	0.415	3	0.233 *	/	/	/
BC28	1	7	HEX	0.400	3	0.027	/	/	/
BC34	1	7	HEX	0.422	3	-0.048	/	/	/
BC35	1	7	PET	0.101	2	-0.052	/	/	/
All						0.038 ***	/	/	/

*1 (Faucher *et al.* 2016) 2 (Rowe *et al.* 1997, 2000) 3 (Rogell *et al.* 2005)

Table S3: Air temperature of nights monitored and composition of the aggregation in *Bufo [Epidalea] calamita*: number of males and females, operational sex-ratio (OSR), number of reproduction event, number of amplectant pairs, mean internal relatedness (*IR*) of all individuals and of breeding individuals only and mean body length (snout to vent length, SVL) in each sex. Also indicated are the minimum, maximum, mean and standard deviation measured for each variable through the breeding season.

Date	Temperature	Attending males	Attending females	OSR	Mating episodes	Amplectant pairs	Mean *IR*	Mean *IR* for breeding pairs	Mean male SVL (mm)	Mean female SVL (mm)
2015-04-01	8	28	3	0.034	1	2	NaN	NaN	NaN	NaN
2015-04-02	2.5	10	0	0.000	0	0	NaN	NaN	NaN	NaN
2015-04-05	10	7	0	0.000	0	0	NaN	NaN	NaN	NaN
2015-04-06	4	0	0	0	0	0	0	0	0	0
2015-04-08	13	32	3	0.059	2	2	NaN	NaN	NaN	NaN
2015-04-09	15	45	7	0.118	6	6	NaN	NaN	NaN	NaN
2015-04-10	14	52	4	0.055	3	4	NaN	NaN	NaN	NaN
2015-04-11	10	NA	3	NA	2	3	NaN	NaN	NaN	NaN
2015-04-12	10	14	2	0.067	1	2	NaN	NaN	NaN	NaN
2015-04-13	8	9	0	0.000	0	0	NaN	NaN	NaN	NaN
2015-04-15	7	25	2	0.000	0	0	0.072	NaN	68.634	64.430
2015-04-16	10.6	6	1	0.143	1	1	-0.005	-0.267	70.192	NaN
2015-04-17	7	1	0	0.000	0	0	-0.267	NaN	71.470	NaN
2015-04-19	NA	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-04-24	NA	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-04-25	11	30	4	0.000	0	1	0.017	NaN	66.503	64.350
2015-04-26	11	45	12	0.022	1	3	0.025	-0.267	66.584	67.116
2015-04-27	11	20	3	0.048	1	3	0.065	0.084	65.611	64.657
2015-04-28	8	8	1	0.111	1	1	0.047	0.554	67.523	70.130
2015-04-29	9	20	9	0.000	0	0	0.026	NaN	66.884	66.709
2015-05-01	NA	1	0	0.000	0	0	-0.107	NaN	62.645	NaN
2015-05-02	13	5	2	0.000	0	0	0.154	NaN	65.916	62.100
2015-05-03	12	17	2	0.000	0	0	0.085	NaN	65.549	76.520
2015-05-04	15	43	9	0.000	0	1	0.036	NaN	66.103	61.129
2015-05-05	7	53	14	0.145	9	14	0.034	0.065	65.978	65.204
2015-05-06	8	41	1	0.024	1	1	0.046	-0.020	66.065	64.450
2015-05-08	14	7	3	0.000	0	0	0.119	NaN	68.600	68.388
2015-05-10	12	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-05-12	14	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-05-14	7	4	1	0.000	0	0	0.128	NaN	62.898	71.550
2015-05-15	8	1	0	0.000	0	0	0.072	NaN	65.160	NaN
2015-05-17	7	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-05-21	9	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-05-23	9	0	0	NaN	0	0	NaN	NaN	NaN	NaN
2015-06-01	11.6	10	0	0.000	0	0	0.214	NaN	65.296	NaN
2015-06-03	10.5	5	0	0.000	0	0	0.094	NaN	68.956	NaN
2015-06-05	17	37	2	0.000	0	1	0.009	NaN	66.410	72.165
2015-06-06	11	39	1	0.025	1	1	0.019	0.005	65.827	70.600
2015-06-08	13	2	0	0.000	0	0	-0.105	NaN	59.780	NaN
2015-06-10	17	5	0	0.000	0	0	0.021	NaN	64.309	NaN
2015-06-12	16	10	1	0.000	0	0	0.059	NaN	67.808	71.920
2015-06-13	11	5	0	0.000	0	0	0.092	NaN	67.630	NaN
2015-06-22	13	31	2	0.000	0	0	0.012	NaN	66.279	65.805
2015-06-25	18	22	2	0.000	0	0	-0.006	NaN	66.400	66.410
Mean	10.639	16.047	2.136	0.024	0.682	1.070	0.035	0.022	66.334	67.424
St. Dev	3.632	16.673	3.275	0.043	1.681	2.414	0.090	0.276	2.329	4.006
Min	2.500	0.000	0.000	0.000	0.000	0.000	-0.267	-0.267	59.780	61.129
Max	18.000	53.000	14.000	0.145	9.000	14.000	0.214	0.554	71.470	76.520

Table S4: Morphological measures (in millimeters) and mating success for 125 adults of the Natterjack toad *(B. calamita)* monitored over a breeding season in 2015 (mm1, mm2, mm2bis, mm3 and D1, D2, D3 are predicted parents defined *a posteriori* with parentage analysis). Body length was measured as the snout-vent length, attendance was the number of monitored night where the individual was observed (from April 15th to June 25th), mating success was the number of reproductive events and the number of eggs -sired or laid- were the number of eggs enumerated or estimated from mother body length (linear regression $r^2 = 0.385$, $P = 0.019$). —: missing data.

♂	Body length (mm)	Tibia length (mm)	Body mass (g)	Attendance	Mating success	Number of eggs sired	♀	Body length (mm)	Tibia length (mm)	Body mass (g)	Attendance	Number of clutches	Number of eggs laid
15MCA001	71.5	24.28	25.7	5	0	0	15MCA004	65.25	22.94	32.7	3	1	2014
15MCA002	68.89	26.49	30	8	2	—	15MCA008	63.61	21.01	25.7	3	1	1782
15MCA003	71.57	26.59	32.5	11	1	1901	15MCA045	70.6	20.77	30.5	4	1	2772
15MCA005	73.66	22.6	28.4	11	0	0	15MCA046	63.46	21.27	26.6	3	0	0
15MCA006	71.47	25.47	33.5	8	2	—	15MCA047	61.15	19.5	24.5	1	0	0
15MCA007	73.17	25.75	25.8	11	0	0	15MCA048	71.55	25.12	34.9	4	1	4428
15MCA009	64.84	22.63	23	12	1	3074	15MCA050	62.14	20.19	25.6	1	0	0
15MCA010	74.72	26.06	34.5	10	1	—	15MCA051	68.74	22.29	31.5	2	0	0
15MCA011	68	25.81	30	6	0	0	15MCA052	66.73	23	32.8	1	0	0
15MCA012	64.34	21.87	22.22	6	0	0	15MCA053	79.53	33.03	28.3	3	1	4036
15MCA013	75.15	23.95	29.9	6	0	0	15MCA054	70.13	22	32.2	4	1	2705
15MCA014	69.28	23.78	25.8	5	1	2308	15MCA055	62.14	20.29	22.7	1	0	0
15MCA015	67	23.49	28.9	4	0	0	15MCA056	63.38	19.3	27.8	3	1	1750
15MCA016	71.92	23.92	28.4	4	1	2276	15MCA071	71.92	21.57	38.2	2	1	2958
15MCA017	62.47	22.64	22.4	15	0	0	15MCA078	54.6	18.39	19.4	3	0	0
15MCA018	71	23	23.6	4	0	0	15MCA079	62.52	19.44	27	1	0	0
15MCA019	67	23.6	21.7	3	0	0	15MCA080	69.77	22.48	36.2	1	0	0
15MCA020	62.17	21.15	22.23	8	1	2109	15MCA081	60.38	20.7	23.5	4	1	1325
15MCA021	67.77	23.26	26.5	4	0	0	15MCA082	67.6	24	27.4	2	2	4694
15MCA022	70.25	25	30.1	17	0	0	15MCA083	62.21	17.84	24.7	3	1	1584
15MCA023	72.82	25.39	32	17	0	0	15MCA084	61.99	19.57	26.6	2	1	1553
15MCA024	60.52	21.69	21.7	6	0	0	15MCA085	79.92	27.3	42.4	1	0	0
15MCA025	67.46	23.8	25.4	2	1	2957	15MCA086	73.12	25.84	44.9	2	1	3128
15MCA026	70.25	23.22	29	10	0	0	15MCA087	59.11	19.69	22.3	2	0	0
15MCA027	58.63	20.29	19.7	1	0	0	15MCA089	63.88	20.92	30.98	2	1	1820
15MCA028	70.01	24.08	37.8	1	1	—	15MCA091	50.19	17.22	16.15	1	0	0
15MCA029	62.85	23.44	26	11	0	0	15MCA097	69.8	25	33.5	0	1	1954
15MCA030	64.46	24.13	24.5	7	0	0	15MCA098	67.1	23	25.6	0	1	2276
15MCA031	67.8	23.2	28.5	5	1	3102	15MCA099	71.405	20.915	30.9	1	1	2230
15MCA032	55.77	20.45	17.8	6	0	0	15MCA100	64.6	20.3	25.9	1	0	0
15MCA033	61.78	22.24	23.5	10	0	0	15MCA101	66.25	22.44	27.6	0	0	0
15MCA034	62.76	21.34	23	4	2	4289	15MCA102	69.64	20.98	32.8	0	1	3102
15MCA035	70.64	24	28.6	18	0	0	15MCA103	70.8	23.17	30.73	0	0	0
15MCA036	66.49	23	25.9	9	0	0	15MCA104	62.87	19.87	29.7	1	0	0
15MCA037	65.04	21.8	24.6	11	0	0	15MCA105	73.73	24.8	34.2	1	0	0
15MCA038	62.03	21.77	23.5	10	1	1553	15MCA106	68.67	23.29	31.5	0	1	2498

15MCA039	59.92	20.03	18.5	10	3	5421
15MCA040	63.56	22.2	25	11	0	0
15MCA041	69.9	25.03	30.7	4	0	0
15MCA042	65.47	24.26	26.8	4	1	4428
15MCA044	62.22	23.5	27.8	7	0	0
15MCA049	58.25	20.89	21.4	1	0	0
15MCA057	55.1	16.52	23.6	7	1	1782
15MCA058	64.77	18.2	30.9	11	0	0
15MCA059	68.7	23.93	31.1	1	0	0
15MCA060	65.16	20.22	24.7	8	1	4036
15MCA061	79.32	30.39	31.2	4	1	3534
15MCA062	68.07	24.61	23.6	11	1	1820
15MCA063	68.1	24.68	28.5	6	1	2772
15MCA065	74.7	23.74	24.1	8	0	0
15MCA066	70.2	25.97	28.5	8	0	0
15MCA067	61.39	21.82	25.2	1	2	5627
15MCA068	67.81	21.45	22.8	3	0	0
15MCA069	75	27.87	42.5	12	1	1404
15MCA070	62.645	20.955	23.3	12	0	0
15MCA072	55.94	21.31	18.2	6	0	0
15MCA073	53.76	17.97	16.4	4	0	0
15MCA074	72.78	22.88	34.2	8	1	3811
15MCA075	65.4	24.6	26.4	3	0	0
15MCA076	71.12	26.21	32.6	10	1	3762
15MCA077	71.66	23.91	34.2	6	0	0
15MCA090	63.58	24.46	23.2	7	0	0
15MCA092	66	24.5	38.1	1	0	0
15MCA093	69	NA	NA	2	0	0
15MCA094	54.2	21	14.6	1	0	0
15MCA095	73.28	NA	NA	1	1	2958
15MCA096	61.1	24	19.4	7	0	0
15MCA111	57.87	19.66	19	3	0	0
MonchBc019	66	NA	NA	0	1	2230
MonchBc021	68.39	NA	NA	0	1	1350
CPIE14S120b	NA	NA	NA	0	1	3121
D1	NA	NA	NA	0	1	1954
D2	NA	NA	NA	0	1	1690
D3	NA	NA	NA	0	1	–
Mean	66.512	23.211	26.411	6.405	0.500	1091
St.dev	5.582	2.339	5.373	4.416	0.667	1587
Min	53.760	16.520	14.600	0.000	0.000	0.000
Max	79.320	30.390	42.500	18.000	3.000	5627

15MCA107	66.15	22.97	34	0	1	2957
15MCA108	72.74	22.47	30	0	1	3074
15MCA109	64.45	21.81	25	2	1	1901
15MCA110	69.34	21.2	31	0	1	2308
15MCA112	69.36	22.3	NA	1	0	0
MonchBc001	NA	NA	NA	0	0	0
MonchBc002	NA	NA	NA	0	1	–
MonchBc008	79.84	NA	NA	0	1	3762
MonchBc010	75.46	NA	NA	0	1	3811
MonchBc014	76.95	24.38	NA	0	1	3534
MonchBc020	65.88	18.26	NA	0	1	1350
MonchBc022	NA	NA	NA	0	1	–
MonchBc026	68.06	19.29	NA	0	1	2109
MonchBc028	84.25	28.26	NA	0	0	0
MonchBc030	63.64	17.45	NA	0	0	0
MonchBc064	NA	NA	NA	1	1	–
MonchBc083	62.96	22.54	23.6	0	1	1690
MonchBc132	NA	NA	NA	1	0	0
mm1	NA	NA	NA	0	1	–
mm2	NA	NA	NA	0	1	1404
mm2bis	NA	NA	NA	0	1	3121
mm6	NA	NA	NA	0	1	–
Mean	67.542	21.923	29.306	1.172	0.638	1525
St.dev	6.458	2.937	5.651	1.286	0.520	1457
Min	50.190	17.220	16.150	0.000	0.000	0.000
Max	84.250	33.030	44.900	4.000	2.000	4694

Table S5: Offspring fitness traits recorded during tadpoles breeding under controlled conditions. Some clutches were accidentally sampled two times and identified *a posteriori* by parentage and sibship inference from multilocus genotype data. Such duplicated clutches were recorded as thanks A and B. Only thanks A were used for fitness analyses. Measures on tadpoles bred in thank were two sibships were *a posteriori* detected are not shown. Clutch size were estimated for non-enumerated clutches from mother body length after mother assignment ($r^2 = 0.385$, $P = 0.019$). Number of days until hatching was recorded as the number of days until more than half the total number of hatched eggs was hatched. Tadpoles body mass was measured on 15 tadpoles per thank (except for thank with less than 15 surviving individuals where all tadpoles were weighted). Metamorphosis was recorded when toadlet had four distinct legs.

Clutch	Laying. date	Thank	Sampling period	Pond	OSR	Clutch size	Clutch size estimated	Hatching success	Number of days until hatching	Gosner stage at day 0 of 30 tadpoles kept for growth monitoring	Days until the 1st metamorphose	Days until the 15th metamorphose	Days until the last metamorphosis	Days until end of growth monitoring	Metamorphoses success at day 45	Survival at day 45	Metamorphoses success at day 60	Survival at day 60	Mean tadpole body mass at day 11	Mean tadpole body mass at day 25	Variance in tadpole body mass at day 11	Variance in tadpole body mass at day 25
010415M1	31/03/2015	A	1	S2	NA	NA	NA	0.990	7	20	53	60	68	70	0.000	0.933	0.600	0.333	0.049	0.098	0.000	0.000
010415M2	31/03/2015	A	1	S2	NA	NA	3074	0.747	7	20	56	NA	62	70	0.000	0.433	0.200	0.233	0.049	0.099	0.000	0.000
010415M3	31/03/2015	A	1	S1	NA	1404	1404	0.961	7	20	49	53	67	70	0.000	1.000	0.700	0.300	0.049	0.099	0.000	0.000
010415M4	31/03/2015	A	1	S1	NA	1954	1954	0.811	7	20	53	65	70	70	0.000	1.000	0.333	0.667	0.052	0.097	0.000	0.000
010415M5	31/03/2015	A	1	S1	NA	3121	3121	0.979	7	20	50	62	70	70	0.000	0.967	0.367	0.633	0.047	0.083	0.000	0.000
050415M10	04/04/2015	A	5	S2	NA	NA	NA	0.923	7	19.5	49	59	67	69	0.000	1.000	0.517	0.483	0.058	0.118	0.000	0.000
090415M11	08/04/2015	A	9	S1	0.086	NA	2498	0.883	8	19.5	42	53	68	70	0.167	0.800	0.600	0.367	0.045	0.110	0.000	0.001
090415M12	08/04/2015	A	9	S1	0.086	NA	2276	0.975	8	19.5	50	60	62	70	0.000	1.000	0.500	0.500	0.042	0.101	0.000	0.000
100415M14	09/04/2015	A	10	S1	0.135	3811	3811	1.000	7	19.5	46	53	66	67	0.000	0.967	0.800	0.133	0.047	0.120	0.000	0.000
100415M15	09/04/2015	A	10	S1	0.135	NA	2488	0.992	8	20.5	39	46	46	70	0.467	0.267	0.500	0.133	0.048	0.109	0.000	0.004
100415M16	09/04/2015	A	10	S1	0.135	3102	3102	0.992	7	20.5	44	49	65	65	0.133	0.867	0.933	0.067	0.047	0.125	0.000	0.000
100415M17	09/04/2015	A	10	S1	0.135	2308	2308	0.975	7	20.5	40	45	58	70	0.500	0.500	0.967	0.033	0.060	0.119	0.000	0.000
100415M18	09/04/2015	A	10	S1	0.135	3762	3762	0.975	8	20.5	40	45	58	58	0.500	0.467	0.967	0.000	0.060	0.157	0.000	0.000
100415M19	09/04/2015	A	10	S1	0.135	3534	3534	0.975	7	19.5	42	49	59	70	0.100	0.900	0.933	0.067	0.048	0.112	0.000	0.000
110415M20	10/04/2015	A	11	S1	0.071	4428	4428	1.000	8	20.5	42	46	61	70	0.367	0.633	0.933	0.067	0.057	0.125	0.000	0.000

110415M21	10/04/2015	A	11	S1	0.071	NA	NA	0.958	8	20.5	45	52	59	70	0.067	0.933	0.967	0.033	0.061	0.129	0.000	0.000
110415M20	10/04/2015	B	11	S1	0.071	NA	NA	NA	8	20.5	43	48	69	70	NA	NA	NA	NA	NA	NA	NA	NA
110415M23	10/04/2015	A	11	S1	0.071	2230	2230	0.950	8	20.5	38	48	64	70	0.333	0.533	0.733	0.167	0.055	0.106	0.000	0.000
120415M24	11/04/2015	A	12	S1	0.130	1350	1350	0.975	7	20.5	65	NA	65	82	0.000	0.933	0.000	0.833	0.067	0.088	0.000	0.000
120415M26	11/04/2015	A	12	S1	0.130	2109	2109	0.950	7	20.5	48	54	68	70	0.000	1.000	0.667	0.333	0.057	0.109	0.000	0.000
130415M27	12/04/2015	A	13	S1	0.125	2957	2957	0.900	7	20.5	43	56	69	72	0.033	0.967	0.600	0.400	0.050	0.118	0.000	0.000
170415M28	16/04/2015	A	17	S1	0.143	NA	NA	0.825	8	20.5	NA	NA	NA	54	0.000	0.567	NA	NA	0.034	0.043	0.000	0.000
280415M29	26/04/2015	A	28	S2	0.196	NA	2347	1.000	6	20.3	48	NA	NA	52	0.000	1.000	0.200	0.800	0.042	0.077	0.000	0.000
280415M30	27/04/2015	A	28	S2	0.130	NA	1325	0.992	7	20	54	63	NA	69	0.000	0.967	0.333	0.633	0.034	0.085	0.000	0.000
290415M31	28/04/2015	A	29	S2	0.111	NA	2705	0.992	7	20	NA	NA	NA	68	0.000	0.600	0.000	0.533	0.052	0.041	0.000	0.000
060515M32	05/05/2015	A	36	S2	0.284	NA	1584	1.000	7	20	55	NA	NA	62	0.000	0.967	0.133	0.833	0.040	0.082	0.000	0.000
060515M34	05/05/2015	A	36	S2	0.284	NA	1750	1.000	7	20	NA	NA	NA	45	0.000	1.000	NA	NA	0.031	0.085	0.000	0.000
060515M35	05/05/2015	A	36	S2	0.284	NA	3128	0.885	7	19.2	NA	NA	NA	45	0.000	0.567	NA	NA	0.031	0.043	0.000	0.000
060515M35	05/05/2015	B	36	S2	0.284	NA	NA	NA	7	19.2	51	NA	NA	62	NA	NA	NA	NA	NA	NA	NA	NA
060515M37	05/05/2015	A	36	S2	0.284	NA	1553	0.838	7	19.5	50	NA	NA	62	0.000	0.467	0.233	0.233	0.043	0.056	0.000	0.000
060515M37	05/05/2015	B	36	S2	0.284	NA	NA	NA	7	19.9	NA	NA	NA	48	NA	NA	NA	NA	NA	NA	NA	NA
060515M39	05/05/2015	A	36	S3	0.284	NA	1782	0.950	7	19.7	58	NA	NA	62	0.000	0.100	0.033	0.067	0.038	0.041	0.000	0.000
060515M40	05/05/2015	A	36	S3	0.284	NA	2014	0.975	7	20	NA	NA	NA	45	0.000	0.933	0.000	0.033	0.028	0.064	0.000	0.000
060515M41	05/05/2015	A	36	S3	0.284	NA	2958	0.983	7	19.5	NA	NA	NA	37	0.000	0.065	0.000	0.065	0.044	0.077	0.000	0.000
070515M44	06/05/2015	A	37	S2	0.024	NA	1901	1.000	6	20	56	NA	NA	62	0.000	1.000	0.133	0.833	0.039	0.072	0.000	0.000
080615M45	06/06/2015	A	49	S2	NA	NA	2772	0.983	7	21	NA	NA	NA	31	NA	NA	NA	NA	0.061	0.089	0.000	0.000
080615M45	06/06/2015	B	49	S2	NA	NA	NA	NA	7	21	NA	NA	NA	31	NA	NA	NA	NA	NA	NA	NA	NA
250615M47	24/06/2015	A	74	S2	NA	NA	2347	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Mean					**0.166**	**2775**	**2486**	**0.950**	**7.189**	**20.076**	**48.179**	**53.300**	**63.857**	**62.108**	**0.083**	**0.760**	**0.479**	**0.338**	**0.047**	**0.093**	**0.000**	**0.000**
Standard error					**0.087**	**964**	**789**	**0.064**	**0.518**	**0.472**	**6.583**	**6.392**	**5.730**	**12.267**	**0.161**	**0.284**	**0.341**	**0.282**	**0.010**	**0.028**	**0.000**	**0.001**
Min					**0.024**	**1350**	**1325**	**0.747**	**6**	**19.2**	**38**	**45**	**46**	**31**	**0.000**	**0.065**	**0.000**	**0.000**	**0.028**	**0.041**	**0.000**	**0.000**
Max					**0.284**	**4428**	**4428**	**1.000**	**8**	**21**	**65**	**65**	**70**	**82**	**0.500**	**1.000**	**0.967**	**0.833**	**0.067**	**0.157**	**0.000**	**0.004**

Table S6: Opportunity for sexual selection on male mating success ($I_{s♂}$) in various amphibian species and their mating system

Species	$I_{Total♂}$	$I_{s♂}$	$I_{S♂}/I_{Total♂}$	References	Breeding period	Mating system
Rana sylvatica	4.6	4	0.87	Kluge 1981		Lek polygyny
Bufo bufo	5.9	5	0.85	Kluge 1981		Scramble competition polygamy
Bufo canorus 1976	14.2	13.6	0.96	Kluge 1981		Scramble competition polygyny
Bufo canorus 1977	2.1	2	0.93	Kluge 1981		Scramble competition polygyny
Bufo canorus 1978	1.8	1.7	0.93	Kluge 1981		Scramble competition polygyny
Bufo canorus 1979	1.9	1.8	0.94	Kluge 1981		Scramble competition polygyny
Bufo exsul 1977	5.8	5.2	0.9	Kluge 1981		Lek polygyny (no advertisement call)
Bufo exsul 1978	2.7	2.4	0.9	Kluge 1981		Lek polygyny (no advertisement call)
Bufo americanus	3.3	2.9	0.88	Kluge 1981		Scramble competition polygyny
Hyla versicolor	3.4	3	0.89	Kluge 1981		Territory-defense polygyny
Rana temporaria	1.7	1.4	0.81	Kluge 1981	Explosive breeders	Lek polygynandry
Rana castebeiana 1976	1.6	1.4	0.87	Kluge 1981		Scramble competition polygyny
Rana castebeiana 1977	1.7	1.5	0.92	Kluge 1981		Scramble competition polygyny
Rana castebeiana 1978	2	1.9	0.95	Kluge 1981		Scramble competition polygyny
Hyla rosenbergi 1977	1.5	1.4	0.99	Kluge 1981		Territory-defense polygynandry and clutch guarding
Hyla rosenbergi 1978	1.7	1.7	0.99	Kluge 1981		Territory-defense polygynandry and clutch guarding
Centrolenella colymbiphyllum	1.1	1.1	0.99	Kluge 1981		Territory-defense and clutch guarding
Centrolenella fleischmanni	0.5	0.5	0.99	Kluge 1981		Territory-defense polygynandry and clutch guarding
Centrolenella valerioi	0.7	0.7	0.98	Kluge 1981		Territory-defense polygynandry and clutch guarding
Dendrobates pumilio	-	0.33-1.48	-	Pröhl 2002		Lek polygyny
Rana catesbeiana	-	1.38	-	Arnold and Wade 1984		Lek polygyny
Bufo valliceps	-	2.28-4	-	Wagner and Sullivan 1992		Lek polygyny
Hyla gratiosa	-	2.26-3.01	-	Murphy 1994		Lek polygyny
Hyla arborea	-	0.8	-	Broquet *et al.*, 2009		Lek polygyny
Physalaemus pustulosus	2.9	2.5	0.89	Kluge 1981	Prolonged breeders	Lek polygamy (prolonged breeder)
Dendrobates pumilio	-	0.33-1.48	-	Pröhl 2002		Territory-defense polygyny and clutch guarding
Dendrobates pumilio	-	1.76	-	Pröhl and Hödl 1999		Territory-defense polygamy and clutch guarding
Epidalea (Bufo) calamita	13.302	1.781	0.169	This study		Lek polygyny

SYNTHÈSE

Résultats et perspectives

La structure spatiale du polymorphisme génétique est un phénomène complexe dont la mise en place dépend aussi bien de facteurs biogéographiques, historiques et contemporains que de caractéristiques intrinsèques de l'espèce considérée. En prenant en compte ces différents aspects, notre étude visait à évaluer l'impact conjoint du processus de colonisation, de la fragmentation de l'habitat et du régime d'appariement sur les populations établies sur les sites récents du bassin houiller du nord de la France. Les résultats, synthétisés dans le tableau 4, offrent un bon aperçu des facteurs ayant façonnés la structure génétique des populations de Crapauds calamites et Pélodytes ponctués localisées dans le bassin minier. Néanmoins, plusieurs pistes pourraient être envisagées pour (i) mieux connaitre les processus micro-évolutifs induits par l'établissement d'espèces sauvages sur des sites miniers formant un réseau d'habitats très localisés au cœur du bassin minier, (ii) mieux évaluer l'impact respectif des différents éléments paysagers comme les milieux végétalisés ouverts, les cultures ou les forêts sur la dispersion des espèces d'amphibiens pionnières dans des paysages fortement fragmentés, (iii) déterminer l'impact du régime d'appariement sur la taille efficace des populations de Crapaud calamite.

Tableau 4 : Rappel des hypothèses de travail et des résultats obtenus dans chacun des grands axes d'études abordés dans ce manuscrit de thèse. √: hypothèse non rejetée, X: hypothèse non validée.

	Attendus	Résultats
Historique de colonisation	Les populations du bassin minier sont situées sur des sites artificiels récents contrairement aux populations du littoral qui sont situées dans des sites naturels. Si les populations localisées sur les terrils sont issues d'une colonisation du bassin minier depuis les zones natives du littoral cela implique :	
	Niveaux de diversité génétique nucléaire et mitochondriale:	
	• Littoral > Bassin minier	X
	Différenciation génétique :	
	• Littoral < Bassin minier (si plusieurs évènements de fondation au sein du bassin minier et/ou flux de gènes restreint entre populations)	√
	• Littoral > Bassin minier (si un seul évènement de fondation au sein du bassin minier et/ou flux de gènes importants entre populations par translocation passive par exemple)	X
	Les populations du bassin minier se situent au cœur d'un paysage fortement fragmenté par des zones anthropogènes	
	Echelle géographique de la différenciation génétique spatiale:	
	• Littoral > Bassin minier	√
Effet du paysage sur les flux de gènes	Les flux de gènes chez les amphibiens semblent sensiblement impactés par la fragmentation du paysage, les éléments du paysage étant plus ou moins perméables aux déplacements des individus. Par ailleurs sur les deux espèces que nous étudions, le crapaud calamite est connu pour disperser sur des distances de l'ordre de 2kms. En revanche, rien n'est connu sur les capacités de dispersion du pélodyte.	
	Effet du paysage sur les flux de gènes	
	• Dans le bassin minier, la divergence génétique s'accorde plus avec des distances "écologiques" fondées sur les caractéristiques du paysage qu'elles ne le sont avec la distance euclidienne	√
	Echelle géographique de la différenciation génétique spatiale:	
	• Littoral > Bassin minier	√
	• Pélodyte ponctué < Crapaud calamite	X
	Le pélodyte ponctué est en limite d'aire de répartition	
	Différenciation génétique :	
	• Pélodyte ponctué > Crapaud calamite	√
Effet du régime d'appariement	Chez le crapaud calamite les mâles se regroupent en lek pendant la période de reproduction, ce qui favoriseune forte variance dans le nombre d'accouplements obtenus par les différents mâles	
	• Variance du succès reproducteur entre mâles	√
	• Opportunité pour la sélection sexuelle (certains traits peuvent être sous sélection sexuelle)	√
	Comportements liés à la sélection sexuelle	
	• Compétition des mâles : le nombre d'accouplements obtenus par les mâles est corrélé à un trait (taille, poids, temps de résidence…) pouvant entrer en jeu dans la compétition	√
	• Choix des femelles : le nombre d'accouplements obtenus par les mâles est corrélé à un trait et/ou est corrélé avec les traits liés à la valeur sélective des jeunes	√/X

Diversité génétique neutre et potentiel adaptatif des populations

Les sites miniers peuvent-ils être considérés comme des refuges pérennes de biodiversité au cœur du bassin minier ?

Les populations de Crapauds calamites aujourd'hui établies au cœur du bassin minier du nord de la France semblent être issues de différentes lignées. Les forts niveaux de diversité génétique nucléaire et les nombreuses lignées mitochondriales que l'on observe dans le bassin minier portent à croire que des introductions multiples aient eu lieu dans cette zone. Les niveaux de consanguinité ne sont pas non plus particulièrement élevés. Ces populations ne semblent donc pas directement vulnérables face à de possibles effets de dépression de consanguinité. Terrils et friches industrielles laissés à l'abandon agissent donc comme zones refuges palliant à la perte d'habitats naturels à l'intérieur des terres. Toutefois, les flux de gènes apparaissent beaucoup plus limités dans le bassin minier que le long du littoral. Au-delà de 5kms l'on ne détecte plus de patron d'isolement par la distance et les populations semblent évoluer indépendamment les unes des autres. Cette échelle spatiale est bien plus restreinte que celle observée le long du littoral qui est plutôt de l'ordre de 40 à 50 kms. Ce résultat, confirmé par notre étude multi-espèces prenant en compte la composition et la structure du paysage, suggère un fort impact de la fragmentation du paysage sur les flux de gènes entre populations d'amphibiens dans le bassin minier.

Dans notre approche, l'utilisation de marqueurs moléculaires neutres nous a permis d'inférer les flux de gènes et les processus démographiques agissant sur la structure génétique des populations. Les mesures de niveaux de diversité génétique neutre nous ont permis d'estimer la « viabilité » des populations en supposant que les forts niveaux de diversité génétique neutre s'accompagnaient d'un plus grand potentiel d'adaptation et de meilleures réponses immunitaires (Hansson & Westerberg 2002; Allentoft & O'Brien 2010). La diversité génétique neutre n'est toutefois pas toujours directement liée à la diversité génétique adaptative (DeWoody & DeWoody 2005; Holderegger *et al.* 2006). La diversité génétique adaptative au sein des populations devra donc être directement étudiée si l'on veut clairement déterminer le potentiel adaptatif de ces populations nouvellement fondées au sein d'un environnement fortement anthropisé. Il semble pertinent d'étudier le polymorphisme du complexe majeur d'histocompatibilité (CMH) comme indicateur de la diversité génétique adaptative. Cette famille multigénique, codant pour les protéines du système immunitaire, est généralement étudiée pour évaluer la capacité des populations à répondre aux attaques de pathogènes (May *et al.*

2011; Höglund *et al.* 2015). Il est généralement admis que plus la diversité génétique du CMH est forte plus l'individu est apte à répondre aux infections (Nowak *et al.* 1992). Elle a notamment déjà été étudiée chez le Crapaud calamite : May *et al.* (2011) ont montré dans des populations touchées par la chytridiomycose (*Batrachochytrium dendrobatidis*), un champignon parasite qui décime de nombreuses espèces d'amphibiens à travers le monde, que la diversité du CMH était plus faible que dans les populations saines. Ce contraste ne transparaissait toutefois pas au regard de la diversité génétique neutre mesurée à partir de marqueurs microsatellites. Les maladies infectieuses touchent de nombreuses populations d'amphibiens, et les sites miniers étant des lieux de passages pour nombre de visiteurs (humains) les amphibiens sont potentiellement plus exposés aux contaminations accidentelles. La diversité du CMH sera donc un paramètre important à prendre en compte pour la gestion des populations localisées au sein des terrils du bassin minier du nord de la France.

De même, il pourrait être intéressant de regarder si des phénomènes d'adaptation locale existent dans les populations établies sur les terrils. En effet, les terrils ont la particularité d'avoir des sols particulièrement riches en métaux lourds (Lemoine 2012; Pauwels *et al.* 2012). Des cas d'adaptations locales ont été relevés chez des plantes ou des champignons établis sur des sites miniers chargés en métaux lourds (Zalłçcka & Wierzbicka 2002; Adriaensen *et al.* 2005). Chez les amphibiens, plusieurs exemples ont montré que les changements physico-chimiques affectent le taux de développement et le taux de survie des larves. Réciproquement, des cas d'adaptation locale à la composition de l'eau ont été démontrés chez la Grenouille des près ou le Crapaud calamite par exemple (Gomez-Mestre & Tejedo 2004; Hangartner *et al.* 2012). Chez ce dernier, Gomez-Mestre & Tejedo (2004) ont montré une adaptation locale à l'eau saumâtre chez les pontes et les têtards par translocation d'œufs d'un environnement d'eau saumâtre à un environnement d'eau douce.

Les terrils, une « oasis » pour la biodiversité du nord de la France ?

Comme précisé en fin de chapitres II et III, le potentiel des sites miniers à héberger des populations pérennes au cœur du bassin houiller dépendra grandement de la gestion de la connectivité du paysage et de la gestion des sites en eux-mêmes. En effet la qualité de l'habitat est au moins aussi importante que la connectivité du paysage pour le maintien des populations chez les amphibiens (Stevens & Baguette 2008) comme chez les communautés d'espèces en général (Villemey 2015). L'habitat comprend des facteurs abiotiques mais également des facteurs biotiques. L'étude des dynamiques de communautés à travers les sites miniers serait l'occasion d'étudier à plus large échelle l'influence de la fragmentation de l'habitat sur les populations, en prenant en compte les interactions

interspécifiques (Vellend 2010). Bardsley & Beebee (2001) ont par exemple mis en évidence qu'en présence de Crapauds communs et de Grenouilles rousses, le taux de survie chez les têtards de Crapaud calamite était diminué par un effet de compétition. Les terrils qui forment des réseaux d'habitats en îles seraient un modèle de choix pour étudier la variation de la composition de la communauté d'espèces dans l'espace et dans le temps. Ce sont des sites faciles à prospecter, regroupant une grande diversité de taxons et qui se distinguent nettement du reste du paysage dans lequel ils s'inscrivent, donnant l'opportunité de suivre les processus micro-évolutifs agissant sur les populations. Un suivi des communautés d'espèces et de leurs interactions à travers ce réseau d'habitat pourrait être entrepris pour prendre en compte les interactions spécifiques dans la démographie des populations (Vellend 2010). De plus, ces habitats pionniers évoluent d'années en années vers des habitats plus fermés ce qui pourrait être l'occasion d'étudier l'effet des successions écologiques sur la composition de la communauté d'espèces (Sousa 1984).

Quelle trame verte pour les espèces d'amphibiens pionnières ?

Quels éléments du paysage sont à privilégier pour faciliter les flux de gènes entre les populations se trouvant dans le bassin minier ?

Comme dans de nombreux autres pays, en France la connectivité du paysage est au centre de programmes de conservation de la biodiversité, notamment avec la mise en place de la politique Trame verte et bleue qui vise « à intégrer la biodiversité dans les décisions d'aménagement du territoire, en s'attachant à la préservation et la remise en bon état des continuités écologiques». Cette démarche fait suite à une politique européenne visant à la création d'un réseau écologique européen nommé réseau écologique paneuropéen. Des données sur le rôle des différentes variables paysagères sur les flux géniques sont précieuses pour une meilleure gestion de ces réseaux de connectivité.

Le paysage très anthropisé du bassin minier limite beaucoup les flux de gènes entre populations au regard d'un paysage littoral moins impacté par l'homme et qui semble plus continu. Les zones urbaines et les autoroutes sont les premières barrières à la dispersion. Au contraire les champs, forêts, routes et prairies semblent relativement perméables à la dispersion mais, en l'état, nous ne pouvons pas déterminer quelle variable paysagère facilite le plus la dispersion. Les terres arables semblent également favoriser les flux de gènes entre populations de Pélodyte ponctué dans la région est du bassin minier tandis que les routes, les zones enherbées ouvertes (*p.e.* prairies) ressortent comme corridors expliquant le mieux les flux de gènes dans la majorité de cas.

La connectivité du paysage a été évaluée en utilisant trois approches couramment utilisées en écologie du paysage : la distance euclidienne, le chemin de moindre coût et la théorie des réseaux de circuits électriques. Dans un premier temps, des modèles de paysage définis *a priori,* ont été testé avant d'évaluer des effets corridor, sans *a priori,* sur chacune des variables. Différents points peuvent être discutés, d'abord concernant la définition des variables paysagères qui regroupent plusieurs habitats sous l'hypothèse qu'ils aient le même effet sur la dispersion de nos deux espèces d'amphibiens. Ensuite, la hiérarchisation des variables paysagères n'est pas encore bien définie. Notamment, nous ne pouvons pas encore définir quel type d'habitat, parmi les prairies ou les forêts sont les plus perméables à la dispersion. Cette information pourrait pourtant améliorer la compréhension de l'écologie de nos deux espèces. Rien n'est connu sur les déplacements du Pélodyte ponctué et les résultats divergent en ce qui concerne le Crapaud Calamite. Enfin, ces analyses font des hypothèses sur les sources de variabilités génétiques comme étant des populations discrètes et localisées dans le paysage. Hors, nos modèles d'études sont des espèces pionnières qui sont susceptibles de coloniser temporairement certains sites. La probable existence de populations non échantillonnées pourrait jouer un rôle dans la structure génétique actuelle des populations étudiées. D'autres méthodes pourraient donc être envisagées qui ne feraient pas d'hypothèse sur les provenances des flux de gènes.

Premièrement, au-delà de biais méthodologiques déjà discutés dans le chapitre III, la définition des catégories dans lesquelles les variables paysagères ont été classées pourraient être revues. Les résultats ont montré que les routes pouvaient faire office de corridors de dispersion. Or sur notre aire d'étude il existe de nombreuses routes nationales au trafic dense et celles que l'on trouve aux environs des terrils sont souvent des routes d'entrée de village où le trafic peut-être relativement faible, notamment de nuit. Si les routes peuvent présenter des surfaces facilitant les déplacement, le risque de mortalité peu grandement varier selon le trafic routier (Stevens *et al.* 2004; Beebee 2013). Une dissociation serait donc nécessaire pour distinguer des routes de forts trafics automobiles des routes peu fréquentées. Cette distinction sera essentielle pour mieux comprendre quels facteurs garantissent la connectivité entre populations. On peut imaginer que ce sont les fossés ou bandes enherbées qui accompagnent les routes qui permettent une bonne dispersion des individus plus que les routes en elles-mêmes. Dans le même ordre d'idée, les restes de voies ferrées des anciens cavaliers sur lesquels étaient transportés les matériaux des mines, ainsi que les chemins agricoles, pourraient être dissociés du reste des variables paysagères classifiées comme habitats ouverts. Les anciens

cavaliers ont tendance depuis leur abandon à se refermer. Des programmes ont toutefois permis d'en maintenir certains ouverts. Mieux cibler l'effet de ces habitats sur la connectivité des populations pourrait permettre de clairement établir s'ils font ou ont fait office de corridors majeurs pour la dispersion des individus entre terrils (Godin 2002).

Deuxièmement, pour affiner l'estimation de l'effet des différentes variables paysagères sur la connectivité des populations, nous pourrions envisager d'analyser le paysage sur des transects reliants les populations les unes aux autres, transects au sein desquels nous estimerions la connectivité du paysage. Cette approche par transect permettraient de modéliser de façon très schématique l'orientation dont font preuve les amphibiens lorsqu'ils doivent rejoindre d'autres habitats (Sinsch 2006). Dans cette approche la connectivité des habitats pourra être estimée par la méthode de la théorie des réseaux de circuits électriques de façon à prendre en compte la possibilité qu'il existe plusieurs chemins possibles pour aller d'un point à l'autre, ce que le chemin de moindre de coût ne fait pas puisqu'il ne retient que le chemin optimum.

Un large panel d'approches additionnelles avec ou sans *a priori* sur les résistances du paysage permettent d'estimer l'effet de la composition et/ou de la structure du paysage sur la dispersion des individus. Le corridor de moindre coût (Dutta *et al.* 2016), le transect de moindre coût (Van Strien *et al.* 2012), la zone potentiel de mouvement (Janin *et al.* 2009) ou encore la détermination des résistances en utilisant un algorithme génétique pour trouver la paramétrisation optimum du paysage afin d'expliquer au mieux les divergences génétiques observées entre populations (Khimoun *et al.* 2017; Gutiérrez-Rodríguez *et al.* 2017). Chacune de ces méthodes présente ses avantages et ses limites. Les approches de corridor de moindre coût ou de transect de moindre coût sont des variantes du chemin de moindre coût qui permettent de prendre en compte le chemin le moins couteux mais aussi le paysage avoisinant qui conditionne également la probabilité que l'individu a de suivre ce chemin (Van Strien *et al.* 2012; Dutta *et al.* 2016). Mais, en se fondant sur des chemins de moindre coût, ces méthodes n'estiment la connectivité qu'à travers un chemin optimum unique pour disperser, oubliant les chemins additionnels pouvant être utilisés par les individus et qui augmentent la connectivité entre populations (McRae *et al.* 2008). Un autre biais qui peut être introduit par le chemin de moindre coût est la sélection de corridors très sinueux, peu couteux mais impliquant des détours qui sont biologiquement peu réalistes (Dyck & Baguette 2005). (Peterman (2014) propose quant à lui une approche qui permet d'estimer les résistances *a posteriori* en utilisant un algorithme génétique, cet algorithme vise à déterminer la « paramétrisation » du paysage qui prédit le mieux la structure

génétique en travaillant soit sur des mesures de chemin de moindre coût, soit sur des résistances globales mesurées avec la théorie des réseaux de circuits électriques. Mais compte-tenu des fortes perturbations humaines modifiant l'environnement dans notre zone d'étude, le paysage actuel est sans doute bien différent de celui qui a structuré les populations que nous observons aujourd'hui et les résistances estimées *a posteriori* pourraient être biaisées par ce décalage temporel. Cette mise en garde est notamment mentionnée par McRae *et al.* (2008) et Epps & Keyghobadi (2015).

Enfin, comme cela a été déjà réalisé dans d'autres études (Flavenot *et al.* 2015), une analyse de relation potentielle entre les niveaux de diversité génétique et la composition du paysage sur un rayon de 1, 2 ou 3 kilomètres autour de chaque population échantillonnée, distances pour lesquelles les individus sont susceptibles de disperser, pourrait être entreprise (Figure 2). Cela permettrait d'évaluer quelle est la variable paysagère qui explique le mieux la connectivité générale de chaque population en prenant les niveaux de diversité génétique intra-population comme indicateur de flux génique. Cette méthode est appliquée dans l'étude de Flavenot *et al.* (2015) afin d'évaluer l'impact des différentes variables paysagères sur la diversité génétique au sein de populations de Crapauds Calamites et de Crapauds communs situés dans des zones de carrières. Outre le fait que cette méthode ne nécessite pas de faire des hypothèses *a priori* sur les effets des variables paysagères sur la dispersion des individus, ceci permet d'étudier les populations une à une sans faire d'hypothèse sur la distribution discrète des populations dans le paysage. En effet, les populations de Crapauds Calamites et de Pélodytes ponctués forment souvent des agrégations discrètes localisées autour de sites de reproduction. Ces deux espèces d'amphibiens sont toutefois des espèces pionnières et ont une bonne capacité à coloniser des sites instables et éphémères. Plusieurs populations localisées entre la zone est et la zone ouest du bassin minier n'ont pas pu être échantillonnées en 2013 et 2014 mais ont été recensées comme étant occupées entre 2010 et 2016 (Figure S1, Chapitre I). En ne considérant que les populations échantillonnées dans notre jeu de données nous oublions peut-être les effets de plus petites populations non échantillonnées qui pourraient être sources de variabilité génétique si certains des individus dispersent entre des sites éphémères non répertoriés et nos sites échantillonnés (*cf* Meirmans 2015). Même si cette approche corrélative a le défaut de ne prendre en compte que la composition du paysage et non sa structure, elle pourrait apporter un bon complément à nos résultats.

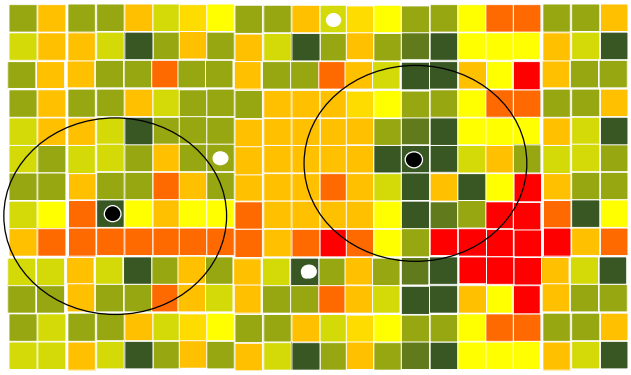


Figure 9: Schéma illustrant la matrice paysagère pour laquelle chaque variable paysagère est représentée par une couleur différente. Les populations échantillonnées sont représentées en noir, les points blancs représentent des populations plus labiles et/ou non échantillonnées mais qui sont susceptibles d'échanger des individus avec les populations échantillonnées. La composition du paysage est évaluée sur le périmètre décrit par les cercles autour de chacune des populations.

Rappelons que les estimations des flux de gènes se fondant sur des mesures de différenciation génétique peuvent être grandement impactées par les facteurs démographiques. La taille efficace des populations, les phénomènes d'expansion de populations, ou encore les goulots d'étranglement au même titre que les flux géniques affectent la différenciation génétique des populations. Or, la taille efficace est un paramètre assez difficile à estimer compte tenu des nombreux biais pouvant être introduits par les générations chevauchantes, les évènements de migration où les changements démographiques récents (*p.e.*Waples & England 2011; Ryman *et al.* 2013; Waples *et al.* 2014). Au cours de ma thèse, j'ai estimé les tailles efficaces à partir d'une méthode fondée sur des mesures de déséquilibre de liaison (Waples 2006). Sur de petits échantillons, ces estimations fondées sur un seul échantillon temporel de la population peuvent-être assez peu robustes (England *et al.* 2006). D'autres méthodes pourraient être utilisées pour recouper nos résultats, comme par exemple d'autres estimateurs fondés sur un seul échantillon temporel comme cela a déjà été fait sur notre jeu de données, ou des estimateurs fondés sur deux échantillons temporels (Wang 2009a). Dans l'ensemble, l'obtention d'estimations statistiques robustes des tailles efficaces nécessite de gros efforts d'échantillonnage au sein des populations. L'analyse de la taille efficace par un suivi de population comme celui que nous avons mené peut entre-autre améliorer nos connaissances sur les dynamiques au sein d'une population et pourquoi pas permettre d'estimer les tailles efficaces à partir des effectifs dénombrés sur une population en se fondant sur le ratio N_e/N moyen obtenu (Palstra & Fraser 2012). Dans le cadre d'analyses de la connectivité du paysage, ces estimations pourront permettre une meilleure évaluation des effets des flux de gènes entre populations.

Régime d'appariement, taille efficace et stochasticité

Le régime d'appariement observé sur une population de Crapaud calamite est-il susceptible d'induire une forte perte de diversité génétique au cours des générations ? Dans la population de Crapaud Calamite étudiée, tous les individus ne se reproduisent pas chaque année. En effet sur notre suivi d'une période de reproduction seulement 41% des mâles et 62% des femelles se sont reproduits. Nos résultats laissent à penser que dans ce système polygyne, la compétition entre mâles favorise l'accès aux femelles pour les mâles les plus grands, qui restent le plus longtemps sur le site de reproduction au cours de la saison, et qui ont des niveaux de consanguinité plus faibles. De plus, au sein de cette population atelier il ressort que parmi les individus s'étant reproduits, les plus consanguins semblent avoir des jeunes de moins bonne valeur sélective que les autres. L'étude du régime d'appariement a été l'occasion de discuter des processus sous-jacents à la variance du succès reproducteur. Les conséquences de cette variance du succès reproducteur sur les niveaux de diversité génétique, *i.e*. la taille efficace (N_e) de la population n'ont toutefois pas été abordées.

La taille efficace est pourtant un paramètre essentiel en biologie de la conservation puisqu'elle permet d'évaluer les niveaux de consanguinité et l'intensité de la dérive génétique (Frankham *et al.* 2002; Bergstrom & Dugatkin 2012; Allendorf *et al.* 2012). Broquet *et al.* (2009), à travers un suivi de population similaire au notre, ont pu estimer la perte de diversité génétique en prenant en compte: la proportion de mâles et la proportion de femelles se reproduisant sur une année ainsi que l'âge à la maturité et le taux de survie des adultes. L'estimation de la taille efficace est ainsi décrite par l'équation suivante :

$$\frac{1}{N_e} = \frac{1}{4T} = \left(\sum_{i=m,f} \frac{A_i\left(1 + I_{A,i}\right) + I_{P,i} + I_{B,i} + C_i}{N_i} \right)$$

où, T est le temps de génération, A_i et $I_{A,i}$ sont respectivement la moyenne et la variance relative de la durée de vie, $I_{P,i}$ et $I_{B,i}$ sont les variances du succès d'accouplement relatif et de la fécondité par accouplement que nous avons estimées dans le chapitre IV pour chacun des sexes i, N_i est l'effectif pour chacun des sexes, C_i représente la covariance entre les composantes P_i et B_i qui sont le nombre de partenaires et la fécondité par accouplement (Arnold & Wade 1984). Broquet *et al.* (2009) ont ainsi montré que chez la rainette verte, le régime d'appariement en lui-même induisait une faible perte de diversité au court du temps en comparaison aux pertes liées à la dérive génétique induites par la perte d'habitat.

Enfin, chez les anoures dont la reproduction dépend de point d'eau « partagés », la variance du succès reproducteur dépend en grande partie de facteurs démographiques et environnementaux (*p.e.* Tejedo 1988). Le succès reproducteur peut en effet être fortement impacté par les fluctuations démographiques et environnementales qui existent au sein d'une année entre population et au sein d'une même population d'une année à l'autre. Pour pouvoir estimer de façon robuste l'impact du régime d'appariement sur les tailles efficaces des populations de Crapaud calamite, des analyses comparatives réalisées sur plusieurs populations et/ou plusieurs années seraient ainsi nécessaires.

CONCLUSION

Les terrils offrent un bel exemple de reconversion post-industrielle où la préservation d'un patrimoine culturel se concilie à la protection de la biodiversité. La pérennité des communautés d'espèces aujourd'hui établies dans ces sites est tributaire du maintien en l'état de cet environnement pionnier. En Suède une étude a estimé la durée de la re-colonisation végétale à 50 ans sur des crassiers (Prach 2003). Les communautés d'espèces qui pourront suivre ne sont pas pour autant à dénigrer mais le maintien des diverses espèces pionnières dans cette région nécessitera alors que l'on veille à la disponibilité d'autres sites pionniers. De manière plus générale, la distribution éparse de ces habitats atypiques dans un paysage lourdement impacté par les activités humaines défini un modèle en île très singulier qui peut permettre d'appréhender, d'une part les effets de la dynamique des milieux, interne à chaque terril, et d'autre part la dynamique du « réseau » prenant en compte les interactions entre les différents sites. Le bassin minier est donc une belle opportunité pour l'étude des dynamiques écologiques dans un modèle où les unités de conservation de la biodiversité sont intégrées dans des zones d'activités humaines.

« *Danglard, vous croyez toujours que je ne fous rien sous prétexte que je ne fous rien. La réalité n'est jamais si simple et vous le savez mieux que quiconque*. »
Le commissaire Adamsberg, Fred Vargas – Coule la Seine, (2002)

Références

Adriaensen K, Vrålstad T, Noben JP, Vangronsveld J, Colpaert J V. (2005) Copper-adapted *Suillus luteus*, a symbiotic solution for pines colonizing Cu mine spoils. *Applied and Environmental Microbiology*, **71**, 7279–7284.

Allendorf FW, Luikart GH, Aitken SN (2012) *Conservation and the Genetics of Populations*. Wiley-Blackwell, Oxford (UK).

Allentoft ME, O'Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, **2**, 47–71.

Arnold SJ, Wade MJ. (1984) On the measurement of natural and sexual selection: theory. *Evolution*, **38**, 709–719.

Bardsley L, Beebee TJC (2001) Strength and mechanisms of competition between common and endangered anurans. *Ecological Applications*, **11**, 453–463.

Beebee TJC (2013) Effects of road mortality and mitigation measures on amphibian populations. *Conservation Biology*, **27**, 657–668.

Bergstrom CT, Dugatkin LA (2012) *Evolution* (B Twitchell, Ed,). W. W. Norton & Company, Inc, New-York & Londres.

Broquet T, Jaquiéry J, Perrin N (2009) Opportunity for sexual selection and effective population size in the lek-breeding European treefrog (*Hyla arborea*). *Evolution*, **63**, 674–683.

DeWoody YD, DeWoody JA (2005) On the estimation of genome-wide heterozygosity using molecular markers. *Journal of Heredity*, **96**, 85–88.

Dutta T, Sharma S, McRae BH, Roy PS, DeFries R (2016) Connecting the dots: mapping habitat connectivity for tigers in central India. *Regional Environmental Change*, **16**, 53–67.

Dyck H Van, Baguette M (2005) Dispersal behaviour in fragmented landscapes : routine or special movements ? , **6**.

England PR, Cornuet J-M, Berthier P, Tallmon D a., Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias in small samples. *Conservation Genetics*, **7**, 303–308.

Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world : disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, **24**, 6021–6040.

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

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

Godin J (2002) Degré de rareté, évolution de la distribution et particularités de l'herpétofaune de la région Nord - Pas de Calais. *Bulletin de la Société Herpétologique de France*, **104**, 16–35.

Gomez-Mestre I, Tejedo M (2004) Local adaptation of an anuran amphibian to osmotically stressful environments. *Evolution*, **57**, 1889–1899.

Gutiérrez-Rodríguez J, Gonçalves J, Civantos E, Martínez-Solano I (2017) Comparative landscape genetics of pond-breeding amphibians in Mediterranean temporal wetlands: The positive role of structural heterogeneity in promoting gene flow. *Molecular Ecology*, **26**, in press (DOI : 10.1111/mec.14272).

Hangartner S, Laurila A, Räsänen K (2012) Adaptive divergence in moor frog (*Rana arvalis*) populations along an acidification gradient: inferences from *Q*(st)-*F*(st) correlations. *Evolution*, **66**, 867–881.

Hansson B, Westerberg L (2002) On the correlation between heterozygocity and fitness in natural populations. *Molecular Ecology*, **11**, 2467–2474.

Höglund J, Wengström Å, Rogell B, Meyer-Lucht Y (2015) Low MHC variation in isolated island populations of the Natterjack toad (*Bufo calamita*). *Conservation Genetics*, **16**, 1007–1010.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21**, 797–807.

Janin A, Léna 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.

Khimoun A, Peterman WE, Eraud C *et al.* (2017) Landscape genetic analyses reveal fine-scale effects of forest

fragmentation in an insular tropical bird. *Molecular Ecology*, **26**, 4906–4919.

Lemoine G (2012) Interêt, gestion, protection et valorisation des terrils miniers dans le Nord–Pas-de-Calais. *Annales de la Société Géologique du Nord*, **19**, 1–8.

May S, Zeisset I, Beebee TJC (2011) Larval fitness and immunogenetic diversity in chytrid-infected and uninfected natterjack toad (*Bufo calamita*) populations. *Conservation Genetics*, **12**, 805–811.

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, **89**, 2712–2724.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223–3231.

Nowak MA, Tarczy-Hornoch K, Austyn JM (1992) The optimal number of major histocompatibility complex molecules in an individual. *Proceedings of the National Academy of Sciences of the USA*, **89**, 10896–10899.

Palstra FP, Fraser DJ (2012) Effective/census population size ratio estimation: a compendium and appraisal. *Ecology and Evolution*, **2**, 2357–2365.

Pauwels M, Vekemans X, Godé C *et al.* (2012) Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **193**, 916–928.

Peterman WE (2014) ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *bioRxiv*, doi: https://doi.org/10.1101/007575.

Prach K (2003) Spontaneous succession in Central-European man-made habitats: what information can be used in restoration practice? *Applied Vegetation Science*, **6**, 125–129.

Ryman N, Allendorf FW, Jorde PE, Laikre L, Hössjer O (2014) Samples from subdivided populations yield biased estimates of effective size that overestimate the rate of loss of genetic variation. *Molecular Ecology Resources*, **14**, 87–99.

Sinsch U (2006) Orientation and navigation in Amphibia. *Marine and Freshwater Behaviour and Physiology*, **39**, 65–71.

Sousa WP (1984) The role of disturbance in natural communities. *Annual Review of Ecology and Systematics*, **15**, 353–391.

Stevens VM, Baguette M (2008) Importance of habitat quality and landscape connectivity for the persistence of endangered natterjack toads. *Conservation Biology*, **22**, 1194–204.

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.

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

Tejedo M (1988) Fighting for females in the toad *Bufo calamita* is affected by the operational sex ratio. *Animal Behaviour*, **36**, 1765–1769.

Vellend M (2010) Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, **85**, 183–206.

Villemey A (2015) Trame verte et papillons de jour en contexte agricole : influence du paysage sur la dispersion, la diversité génétique et la composition des communautés. *PhD Thesis, Université d'Orléans*.

Wang J (2009) A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, **18**, 2148–64.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167–184.

Waples RS, Antao T, Luikart G (2014) Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, **197**, 769–780.

Waples RS, England PR (2011) Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, **189**, 633–644.

Załłçcka R, Wierzbicka M (2002) The adaptation of *Dianthus carthusianorum* L. (Caryophyllaceae) to growth on a zinc-lead heap in southern Poland. *Plant and Soil*, **246**, 249-257.

ANNEXE

Publications et communications nationales ou internationales

Publications :

Faucher, L., Hénocq, L., Vanappelghem, C., Rondel, S., Quevillart, R., Gallina, S., Godé, C., Jaquiéry, J. & Arnaud, J.-F. (2017) When new human-modified habitats favor the expansion of an amphibian pioneer species: evolutionary history of the natterjack toad (*Bufo calamita*) in a coal basin. *Molecular Ecology*, **26**, 4434-4451.

Faucher, L., Godé, C. & Arnaud, J.-F. (2016) Development of nuclear microsatellite loci and mitochondrial single nucleotide polymorphisms in the natterjack toad, *Bufo* (*Epidalea) calamita* (Bufonidae), using next generation sequencing and Competitive Allele Specific PCR (KASPar). *Journal of Heredity*, **107**, 660-665.

Faucher, L., Hénocq, L., Vanappelghem, C., Rondel, S., Quevillart, R., Godé, C. & Arnaud, J.-F. (2016) Genetic structure of native and newly founded populations of the Natterjack Toad (*Bufo calamita*) in northern France. *Herpetological Journal*, **26**, 2-2.

Almeida, C. E., **Faucher, L.**, Lavina, M., Costa, J. & Harry, M. (2016) Molecular individual-based approach on *Triatoma brasiliensis*: inferences on triatomine foci, *Trypanosoma cruzi* natural infection prevalence, parasite diversity and feeding sources. *PLoS Neglected Tropical Diseases*, **10**, e0004447

Communications nationales ou internationales :

Faucher, L., Hénocq, L., Vanappelghem, C., Rondel, S., Quevillart, R., Godin, J., Gallina, S., Godé, C., Jaquiery, J. & Arnaud, J.-F. (2016) Lorsque les milieux fortement anthropisés favorisent l'expansion d'espèces d'amphibiens pionnières : histoire évolutive du Crapaud Calamite dans le nord de la France. *44^{e} congrès de la Société Herpétologique de France / 2^{e} congrès franco-belge d'Herpétologie « L'Herpétofaune des milieux anthropiques ». 30 sept.- 02 oct. Namur, Belgique*. (Communication orale invitée).

Faucher, L., Vanappelghem, C., Rondel, S., Quevillart, R., Godé, C., Gallina, S. & Arnaud, J.-F. (2015) When human-disturbed landscapes lead to expansion of protected species: genetic structure of native and newly founded populations in the natterjack toad (*Epidalea calamita*). *49th Population Genetics Group Meeting 15-18 december, University of Edinburgh, UK* (Communication orale).

Faucher, L., Gallina, S. & Arnaud, J.-F. (2015) Un exemple d'inférences bayésiennes en génétique des populations, le cas du crapaud calamite (*Epidalea calamita*) dans le nord de la France. *Journées SUCCES – France Grilles, Institut de Physique du Globe de Paris, 5-6 novembre, France*. (Communication orale invitée).

Faucher, L. & Arnaud, J.-F. (2015) Genetic structure of native and newly founded populations of the natterjack toad (*Bufo calamita*) in Western Europe. *Joint Irish and British Herpetological Scientific Meeting, Amphibian and Reptile Biology and Conservation, 28-29th August, Trinity College, Dublin, Ireland*. (Communication orale invitée).

Faucher, L., Hénocq, L., Vanappelghem, C., Rondel, S., Quevillart, R., Godé, C. & Arnaud, J.-F. (2014) Genetic structure of native and newly founded populations of the natterjack toad (*Bufo calamita*) in northern France. *Joint "British Ecological Society" and "Société Française d'Ecologie" Annual Meeting, 09-12 December, Lille, France* (Communication affichée).

Hénocq, L., **Faucher, L.**, Vanappelghem, C., Rondel, S., Quevillart, R., Godé, C. & Arnaud, J.-F. (2014) Structure génétique de populations natives et nouvellement fondées chez le crapaud calamite (*Bufo calamita*) en Région Nord-Pas de Calais. *XXXVIème Réunion du Groupe de Biologie et Génétique des Populations, Université d'Orsay* (Communication affichée).

Faucher, L., Hénocq, L., Vanappelghem, C., Rondel, S., Quevillart, R., Godé, C. & Arnaud, J.-F. (2014) Structure génétique de populations en limite d'aire de répartition : exemple du Pélodytes ponctué (*Pelodytes punctatus*). *XXXVIème Réunion du Groupe de Biologie et Génétique des Populations, Université d'Orsay* (Communication affichée).

© J. Arbez

"Il existe un monde d'espace, d'eau libre, de bêtes naïves où brille encore la jeunesse du monde et il dépend de nous, et de nous seuls, qu'il survive…"

Samivel (1952)

Résumé :

La perte et la fragmentation des habitats générées par les activités humaines érodent la diversité génétique intra-spécifique, entrainant l'extinction de populations chez de nombreuses espèces. Paradoxalement, des habitats artificiels comme les terrils du nord de la France favorisent l'installation de populations sauvages. Néanmoins leur localisation au sein d'un paysage très anthropisé interroge sur la pérennité des populations qu'ils hébergent. Cette étude visait à étudier la diversité génétique neutre des populations de *Bufo calamita* et *Pelodytes punctatus*, deux espèces d'amphibiens établies dans le bassin houiller et dans des habitats littoraux plus sauvages. Des approches de génétique des populations ont permis de décrire les effets de différents processus micro-évolutifs sur les niveaux de diversité génétique, depuis le processus biogéographique de colonisation des terrils jusqu'au régime d'appariement dans une population, en passant par une analyse multi-espèces de la connectivité paysagère. Les populations de *B. calamita* du bassin houiller présentent de forts niveaux de diversité génétique pouvant résulter d'introductions d'individus de diverses localités. Toutefois, dans le bassin houiller, une forte différenciation génétique s'observe chez les deux espèces. Cela s'explique au moins en partie par la présence de barrières aux flux de gènes entre populations qui, à long terme, pourrait compromettre le maintien des populations. Enfin, le succès reproducteur inégal des mâles de *B. calamita,* qui pourrait induire des baisses de niveau diversité génétique intra-population, semble associé à une compétition entre mâles et implique plusieurs stratégies d'appariements.

Abstract:

Human activities induce habitat loss and fragmentation that have an erosive effect on the level of intraspecific genetic diversity, decreasing the individual fitness and jeopardizing populations' adaptive capability. Conversely, new human-made areas, such as spoil heaps of northern France, can provide suitable habitats for pioneering species. Spoil heaps being part of a highly human-fragmented landscape, the likelihood of population persistence is questioned given the scarcity of suitable habitats and the occurrence of potential barriers to dispersal. We studied the intraspecific genetic diversity of two anurans, *Pelodytes punctatus* and *Bufo calamita*, located in coalfield areas and semi-natural coastal habitats. We focused on the effects of micro-evolutionary processes of genetic drift and gene flow in shaping genetic structure. We studied (i) the biogeographical history of colonization of coalfield areas in *B. calamita*, (ii) the landscape connectivity using a multispecies approach, and (iii) the evolutionary determinants of variance in breeding success in *B. calamita*. In coalfield areas, *B. calamita* populations showed high levels of genetic diversity suggesting several independent colonization events. Nonetheless, marked local genetic discontinuities were observed within coalfield areas for both species, suggesting occurrence of environmental barriers impeding gene flow that may compromise population viability. Within a *B. calamita* population, we observed a polygynous mating system involving a possible decrease in genetic diversity. Our results suggested that variance in male mating success was linked to male-male competition and may imply distinct mating strategies.