



HAL
open science

Maladies émergentes en santé animale: diversité génétique et dispersion d'une espèce de moucheron vectrice des virus de la fièvre catarrhale ovine et de Schmallerberg en région paléarctique: *Culicoides obsoletus*, Diptera Ceratopogonidae

Antoine Mignotte

► **To cite this version:**

Antoine Mignotte. Maladies émergentes en santé animale: diversité génétique et dispersion d'une espèce de moucheron vectrice des virus de la fièvre catarrhale ovine et de Schmallerberg en région paléarctique: *Culicoides obsoletus*, Diptera Ceratopogonidae. Sciences agricoles. Université Montpellier, 2020. Français. NNT: 2020MONTG069 . tel-03340452

HAL Id: tel-03340452

<https://theses.hal.science/tel-03340452>

Submitted on 10 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En écologie de la santé

École doctorale GAIA - Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Unité Mixte de recherche UMR 117 ASTRE - Animal, Santé, Territoires, Risques et Ecosystèmes

Maladies émergentes en santé animale : diversité génétique et dispersion d'une espèce de moucheron vectrice des virus de la fièvre catarrhale ovine et de Schmallenberg en région paléarctique, *Culicoides obsoletus* (Diptera : Ceratopogonidae)

Présentée par Antoine MIGNOTTE
Le 14 décembre 2020

Sous la direction de Claire GARROS
et Karine HUBER

Devant le jury composé de

Stéphanie MANEL, DR, EPHE Paris

Stéphanie JACQUET, Chargé de recherche, CNRS Lyon

Patrick MARDULYN, Pr, Université Libre de Bruxelles

Renaud VITALIS, DR, INRAE Montpellier

Jérôme DEPAQUIT, Pr, Université de Reims Champagne Ardenne

Claire GARROS, Chargé de recherche, CIRAD Montpellier

Membre du jury

Membre du jury

Rapporteur

Président du jury

Rapporteur

Directrice de Thèse



UNIVERSITÉ
DE MONTPELLIER

UNIVERSITÉ DE MONTPELLIER

École doctorale : GAIA

THÈSE

Pour obtenir le grade de

Docteur de l'Université Montpellier II

Discipline : Écologie de la santé

Présentée et soutenue par

Antoine MIGNOTTE

Maladies émergentes en santé animale : diversité génétique et dispersion d'une espèce de moucheron vectrice des virus de la fièvre catarrhale ovine et de Schmallenberg en région paléarctique, *Culicoides obsoletus* (Diptera : Ceratopogonidae)

Thèse dirigée par Claire GARROS, CIRAD et Karine HUBER, INRAE

Soutenue le 14/12/2020

Devant le jury composé de

Stéphanie MANEL, DR, EPHE Paris

Stéphanie JACQUET, Chargé de recherche, CNRS Lyon

Patrick MARDULYN, Pr, Université Libre de Bruxelles

Renaud VITALIS, DR, INRAE Montpellier

Jérôme DEPAQUIT, Pr, Université de Reims Champagne Ardenne

Claire GARROS, Chargé de recherche, CIRAD Montpellier

Membre du jury

Membre du jury

Rapporteur

Président du jury

Rapporteur

Directrice de Thèse

Résumé

Depuis 2006, des épizooties importantes de fièvre catarrhale ovine et de virus (FCO) de Schmallerberg (SB) sont observées dans la région ouest paléarctique, ces deux virus étant transmis par des moucheron hémato-phages autochtones du genre *Culicoides* (Diptera : Ceratopogonidae). Ces virus affectent les bovins, les ovins, les caprins et les ruminants sauvages. La dispersion des *Culicoides* correspond à ce que les écologues appellent dispersion stratifiée ou « *stratified dispersal* ». Elle résulte de la combinaison de dispersion active de proche en proche, sur de courtes distances, et de sauts sur de longues distances via le vent de manière passive ou semi-passive, permettant des sauts de dispersion de plusieurs centaines de kilomètres. *Culicoides obsoletus* est une espèce vectrice importante pour les virus de la FCO et de SB et abondante en région paléarctique. Si la dispersion des *Culicoides* par le vent au-dessus des masses d'eaux est bien établie, il est aussi important d'évaluer la dispersion au-dessus des terres et de déterminer l'influence des facteurs biotiques et abiotiques influant sur celle-ci.

L'objectif de ce travail de thèse est de déterminer, par l'observation des patrons de distribution des fréquences alléliques, comment la dispersion a pu façonner l'arrangement spatial de la diversité génétique et comment elle est influencée par les paysages et par les caractéristiques des habitats qui peuvent favoriser ou limiter les mouvements des *Culicoides*. Pour cela, deux plans d'échantillonnage ont été utilisés : un échantillonnage populationnel à l'échelle de l'Europe (de la Scandinavie aux îles Canaries, à la Turquie), et un échantillonnage individuel à l'échelle française. L'utilisation complémentaire de marqueurs mitochondriaux et nucléaires à ses différentes échelles spatiales nous ont permis (i) de clarifier la distribution et de réviser le schéma taxonomique du Complexe *Obsoletus/Scoticus*, (ii) de mettre en évidence des flux de gènes importants permettant d'homogénéiser la structure génétique des populations à l'échelle d'un pays et d'un continent, (iii) de quantifier l'impact d'un large éventail de facteurs environnementaux et météorologiques sur la dispersion des *Culicoides* et de mettre en évidence l'importance de la densité d'hôtes, (iv) et de proposer un scénario présentant l'histoire évolutive récente des populations européennes de *C. obsoletus*. L'hypothèse retenue est une recolonisation post-glaciaire récente depuis le refuge glaciaire des Balkans et un refuge plus au nord-est. Ces populations auraient recolonisé ensemble l'ouest de l'Europe, engendrant une admixture des populations.

Nos résultats démontrent une capacité de dispersion de *C. obsoletus* à l'intérieur des terres très importante, qui doit être prise en compte pour les travaux à venir sur la compétence des vecteurs et la modélisation épidémiologique de la transmission des virus transmis.

Mots-clés : *Culicoides*, dispersion, génétique des populations, microsattellites, gènes mitochondriaux.

Discipline : Écologie de la santé

Cirad, UMR ASTRE - Animal, Santé, Territoires, Risques et Ecosystèmes
Campus International de Baillarguet, 34398 Montpellier Cedex 5, France

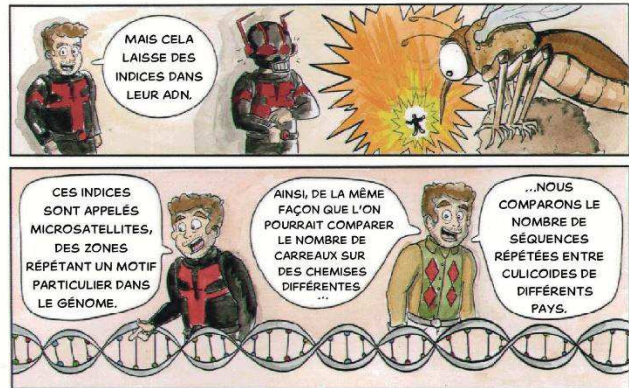
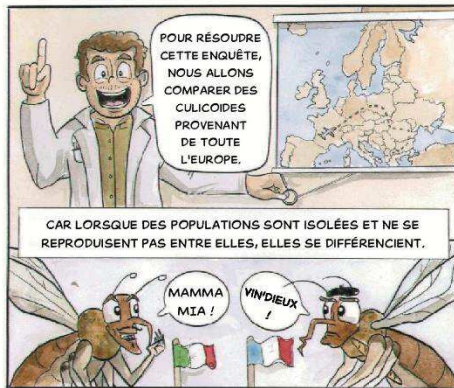
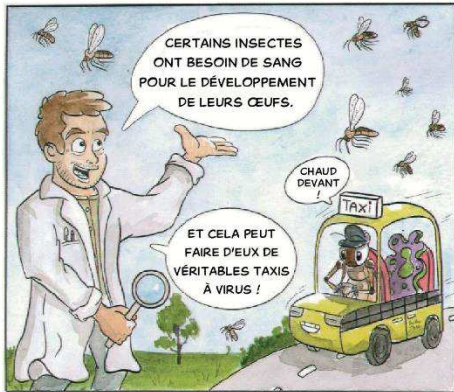
Summary

Emerging diseases and animal health: genetic diversity and dispersal of a biting midge species vector of bluetongue and Schmallenberg viruses in the Palearctic region, *Culicoides obsoletus* (Diptera: Ceratopogonidae)

Since 2006, important outbreaks of bluetongue virus (BTV) and Schmallenberg virus (SBV) have been reported in the West Palearctic region. Both viruses are transmitted by hematophagous biting midges of the genus *Culicoides* (Diptera: Ceratopogonidae). These viruses affect wild cattle, sheep, goats and wild ruminants. The dispersion of *Culicoides* corresponds to what ecologists call stratified dispersal. It results from the combination of active dispersal from near to near, over short distances, and jumps over long distances via the wind in a passive or semi-passive way, allowing dispersal jumps of several hundred kilometers. *Culicoides obsoletus* is a recognized vector species for BTV and SBV, abundant in the Palearctic region. While the dispersal of *Culicoides* by wind over water bodies is well established, it is also important to assess the dispersion over land and determine the influence of biotic and abiotic factors influencing it.

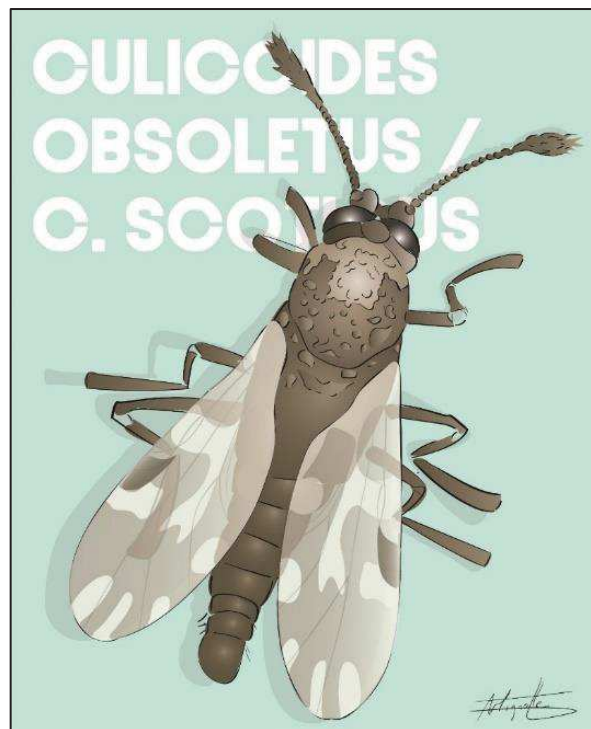
The objective of this thesis is to determine by observing the pattern of distributions of allelic frequencies how dispersion may have shaped the spatial arrangement of genetic diversity and how it is influenced by landscapes and habitat characteristics that can promote or limit the movements of *Culicoides*. Two sampling strategies were used: an Europe-wide population sampling (from Scandinavia to the Canary Islands, to Turkey), and an individual sampling at the French level. The complementary use of mitochondrial and nuclear markers at these different spatial scales has allowed us (i) to clarify the distribution and revise the taxonomic scheme of the *Obsoletus/Scoticus* Complex, (ii) to highlight important gene flows that homogenize the genetic structure of populations at the country and continent level, (iii) to quantify the impact of a wide range of environmental and meteorological factors on *Culicoides* dispersal and highlight the importance of host density, (iv) and to propose a scenario for the recent evolutionary history of European populations of *C. obsoletus*. The hypothesis is a recent post-glacial recolonization from the Balkan ice refuge and a refuge further northeast. These populations would have recolonized together western Europe, resulting in admixture of populations. Our results demonstrate a very important inland dispersal capability of *C. obsoletus*, which must be considered for future works on vector competence and epidemiological modelling of transmitted virus transmission.

Key-words : *Culicoides*, dispersal, population genetics, microsatellites, mitochondrial genes.



Remerciements

Merci à Patrick MARDULYN et Jérôme DEPAQUIT d'avoir accepté d'être rapporteurs de cette thèse. Merci également à Stéphanie MANEL, Stéphanie JACQUET et Renaud VITALIS d'être examinateurs de ces travaux. Pour leur soutien financier, je remercie PALE-Blu, VectorNet, et le CIRAD. Je tiens à remercier Erwan QUEMERE, Christine CHEVILLON, et Thierry BALDET d'avoir accepté de faire partie de mon comité de thèse. Vos conseils pendant ce rendez-vous annuel m'ont guidé et permis d'affiner la méthodologie à suivre pour atteindre nos objectifs.



Je tiens à remercier chaleureusement mes encadrantes Claire GARROS et Karine HUBER. Vous avez été essentielles à l'accomplissement de cette thèse. Je mesure la chance de vous avoir eues à mes côtés pour m'épauler. Claire, comment t'exprimer ma gratitude ? Merci d'avoir été cette oreille attentive dont j'avais besoin, et d'avoir cru en moi quand bien souvent, moi-même je n'y croyais pas. Merci pour tes encouragements, ta patience lors des corrections, mais plus encore, pour ta bonne humeur et ton sourire permanent. Ta force de caractère est inspirante pour l'éternel anxieux que je suis. Merci Karine, pour ta disponibilité, ta gentillesse et tes conseils toujours pertinents qui m'ont permis de murir scientifiquement. Merci de m'avoir guidé de façon bienveillante tout au long de ces trois années et de m'avoir recentré quand il le fallait. Tes connaissances ont sans aucun doute contribué d'une manière significative à l'amélioration de ces travaux de thèse. Merci infiniment à vous deux !

Parce que la collaboration est un point clé de la recherche, je remercie tous les membres du Service d'Epidémiologie spatiale (SPELL) pour leur accueil chaleureux à Bruxelles. Une reconnaissance particulière à Simon DELLICOUR, qui m'a aiguillé dans les analyses de génétique du paysage. Tes compétences et ta pédagogie sont une réelle source d'inspiration. Merci également à l'ensemble de l'Unité Mixte de Recherche Epidémiologie des maladies animales et zoonotiques (EPIA) de l'INRA de Clermont-Ferrand pour leur accueil et particulièrement à Maude JACQUOT, pour son attention et la réceptivité portée à toutes mes interrogations. Je remercie Bruno MATHIEU, pour son accueil à Strasbourg et son aide durant cette thèse. Je tiens à remercier Marie Pierre CHAPUIS, Karine BERTHIER, et Sylvain PIRY,

d'avoir pris le temps de me recevoir, de m'écouter et de me conseiller judicieusement sur les analyses à effectuer.

Enfin, j'exprime ma gratitude à l'ensemble du Collectif Vecteur, Laetitia, Thierry, Ignace... et plus généralement, à l'ensemble des membres de l'unité ASTRE « Animal, Santé, Territoires, Risques et Ecosystèmes », notamment Nathalie VACHIERY et Thierry LEFRANÇOIS, pour m'avoir accueilli au sein de l'unité. Merci à tous les partenaires européens avec qui j'ai eu la chance de collaborer. Merci à toutes les personnes qui ont croisé mon chemin au cours de ces trois années de recherches, que ce soit autour d'un café, d'une discussion, d'un Skype, d'un mail, d'une réunion ou encore d'une bière au soleil.

La recherche ne serait rien sans congrès, et sans les rencontres que l'on fait et que l'on entretient pendant ceux-ci. Merci à Maria, Lyson, Marion et Yannick, d'avoir fait partie de mon aventure de thèse et j'espère à bientôt !

Trois ans, ça en fait des rencontres et des souvenirs. Merci à mes colloques de canap' Clélia et Janis pour avoir supporté mes rires trop forts et mes sautes d'humeurs pendant tant de temps. J'ai été chanceux d'avoir été entouré par vos ondes positives. Et quel plaisir ne pas être le seul à porter un deuxième prénom que je qualifierais d'« original ». Merci pour tout Billy et Jeanine. Merci Cécile, d'avoir ensoleillé le bureau par ta bonne humeur, ton sourire et tes cours de ukulélé. Merci à Antoni, dont le talent en photoshop n'a d'égal que celui pour la danse. Merci à Maxime, pour avoir fait tourner l'activité familiale en Bourogne en mon absence. Merci à l'éminent numérologue Roger jr ELOIFFIN, dit le bien né, pour ses prédictions avisées. Merci à Marc HOULE pour son dynamisme. Merci à micky mike pour l'ensemble de son œuvre. Merci au déménagement de nous avoir laissé l'opportunité de laisser notre trace dans l'unité. Merci au canapé du bureau de m'avoir permis d'assoir mon statut de doctorant. En clair, merci à tous les copains de thèse, Cécile, Roger, Antoni, Marine, Maxime, Etienne, Julien, Remi, Célia¹, Célia², Rémy... vous avez rendu cette expérience absolument inoubliable.

Merci également à tous les copains Bourguignons pour leur soutien : Bubu, Mout, Bibiche, JB, Marine, Eugénie, Eliott, Léa, Emilie, Coralie... et parce que le rugby c'est un ballon avec des copains autour, quand il n'y a plus de ballon, il reste le copains, merci à vous d'être toujours là après toutes ces années : Lit's, Jules, Coco...

Merci infiniment à toute ma famille d'avoir été un soutien indéfectible, toujours à mes côtés. Un merci tout particulier à mon père, ma mère, ma Juju et mes grands-parents, merci infiniment d'être là et de m'avoir soutenu jusqu'au bout !

Enfin, merci à ma petite voix, pour son soutien, pas seulement depuis le début cette thèse mais depuis ces 10 dernières années. Merci pour ta compréhension depuis tout ce temps.

J'en oublie encore certainement, ne m'en voulez pas si votre nom n'apparaît pas, MERCI à tous !

Valorisation scientifique

Publications de rang A

Article 1

Mignotte A, Garros C, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, Tabourin L, Poujol L, Mathieu B, Ibañez-Justicia A, Deniz A, Cvetkovikj A, Purse BV, Ramilo DW, Stougiou D, Werner D, Pudar D, Petrić D, Veronesi E, Jacobs F, Kampen H, Pereira da Fonseca I, Lucientes J, Navarro J, Martinez-de la Puente J, Stefanovska J, Searle KR, Khallaayoune K, Culverwell CL, Larska M, Bourquia M, Goffredo M, Bisia M, England M, Robin M, Quaglia M, Miranda-Chueca MA, Bødker R, Estrada-Peña R, Carpenter S, Tchakarova S, Boutsini S, Sviland S, Schäfer SM, Ozoliņa Z, Segliņa Z, Vatansever Z, Huber H. (2020). **The tree that hides the forest: Cryptic diversity and phylogenetic relationships in the Palearctic species vector Complex *Culicoides obsoletus*/*C. scoticus* (Diptera: Ceratopogonidae) at the European level.** *Parasites Vectors* 13, 265, <https://doi.org/10.1186/s13071-020-04114->, (impact factor 3.43).

Article 2

Mignotte A, Garros C, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, Huber K. **Recent phylogeographic history in western Europe of one main species vector of bluetongue and Schmallenberg viruses, *Culicoides obsoletus* (Diptera: Ceratopogonidae) demonstrates important continental connectivity.** *Sera soumis dans Scientific Reports* (impact factor 5.578)

Article 3

Mignotte A, Garros C, Dellicour S, Jacquot M, Gilbert M, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, de Wavrechin M, Huber K. **Important dispersal capacity of one main species vector of bluetongue and Schmallenberg viruses *Culicoides obsoletus* (Diptera: Ceratopogonidae) revealed by landscape genetic analyses at French scale.** *Soumis en septembre 2020 dans Parasites & Vectors* (impact factor 3.43), <https://dx.doi.org/10.21203%2Frs.3.rs-75603%2Fv1>

Articles publiés hors du contexte de la thèse

Bourquia M, Garros C, Rakotoarivony I, Gardes L, Huber K, Boukhari I, Delécolle JC, Baldet T, **Mignotte A**, Lhor Y, Khallaayoune K, Balenghien T. 2019. **Update of the species checklist of *Culicoides* Latreille biting midges (Diptera: Ceratopogonidae) of Morocco. Two new records, *C. paolae* Boorman and *C. subfagineus* Delécolle and Ortega.** Parasites Vectors 12, 459, <https://doi.org/10.1186/s13071-019-3720-4>

Communications orales

Mignotte A, Reveillaud J, Huber K. **Population genetics and bacterial community structures of *Culicoides obsoletus*: a mirror of European geography ?** *Réunion annuelle de projet PALE-Blu, Rabat, Maroc, 19 septembre 2018*

Mignotte A. **The tree that hides the forest: Cryptic diversity and phylogenetic relationships in the Palearctic species vector Complex *Culicoides obsoletus*/*C. scoticus* (Diptera: Ceratopogonidae) at the European level.** *Réunion annuelle de projet PALE-Blu, Guilianova, Italie, 25 septembre 2019*

Huber K, Jacquet S, Gardès L, **Mignotte A**, Balenghien T, Bouyer J, Garros C. 2017. **Where do they come from ? where do they go ? Population genetics studies on *Culicoides* biting midges.** In : *Sove 2017 New technology conquering old vectors ? SOVE, UIB. Palma de Mallorca : SOVE, 133. ISBN 978-84-697-6086-4 International SOVE congress: New technology conquering old vectors?.* Palma de Mallorca, Espagne, 1 Octobre 2017/7 Octobre 2017.

Poster

Mignotte A, Huber K, Gardès L, Balenghien T, Duhayon M, Rakotoarivony I, Tabourin L, Martinez de la Puente J, Lucientes J, Ibañez-Justicia A, Kampen H, Werner D, England M, Carpenter S, Garros C. 2019. **Landscape genetics of *Culicoides* vector species : evaluation of cryptic diversity and gene flow in the Palearctic region.** In : *21st E-SOVE Meeting Abstract Book. Arthropod Vector Science for the benefit of society: Educate, Empathize, Engage. ESOVE. Palermo : Italie, 22 Octobre 2018/26 Octobre 2018.*

Table des matières

RESUME	5
SUMMARY	6
REMERCIEMENTS	8
VALORISATION SCIENTIFIQUE	I
TABLE DES ILLUSTRATIONS : FIGURES.....	V
TABLE DES ILLUSTRATIONS : TABLEAUX	VIII
INTRODUCTION GENERALE	1
I) L'ETUDE DE LA DISPERSION CHEZ LES ARTHROPODES VECTEURS	12
I.1) <i>Des approches directes parfois inadaptées.....</i>	12
I.2) <i>Evaluation des flux de gènes par méthodes indirectes lors de la dispersion d'insectes vecteurs</i>	13
I.3) <i>Evaluation et quantification de l'impact des facteurs environnementaux sur les flux de gènes d'insectes vecteurs.....</i>	17
I.4) <i>Des flux de gènes contemporains à la création de nouvelles espèces.....</i>	24
II) <i>CULICOIDES OBSOLETUS</i> , UN MODELE D'ETUDE DE LA DISPERSION D'ESPECES VECTRICES DE VIRUS D'INTERET VETERINAIRE	30
II.1) <i>Généralités sur la biologie et l'écologie des Culicoides</i>	30
II.2) <i>Le rôle des Culicoides dans la transmission de pathogènes en Europe du nord.....</i>	33
II.3) <i>La dispersion stratifiée chez les Culicoides.....</i>	37
II.4) <i>Culicoides obsoletus : un besoin de clarification taxonomique.....</i>	42
PROBLEMATIQUE ET OBJECTIFS DE LA THESE	46
CHAPITRE I : THE TREE THAT HIDES THE FOREST: CRYPTIC DIVERSITY AND PHYLOGENETIC RELATIONSHIPS IN THE PALAEARCTIC VECTOR OBSOLETUS/SCOTICUS COMPLEX (DIPTERA: CERATOPOGONIDAE) AT THE EUROPEAN LEVEL	49
SUPPLEMENTARY INFORMATION	64
CHAPITRE II : RECENT PHYLOGEOGRAPHIC HISTORY IN WESTERN EUROPE OF ONE MAIN SPECIES VECTOR OF BLUETONGUE AND SCHMALLEMBERG VIRUSES, <i>CULICOIDES OBSOLETUS</i> (DIPTERA: CERATOPOGONIDAE) DEMONSTRATES IMPORTANT CONTINENTAL CONNECTIVITY	73
ABSTRACT.....	76
INTRODUCTION	77
METHODS.....	80
RESULTS	84
DISCUSSION	87

CONCLUSION.....	91
FIGURES AND TABLES.....	93
SUPPLEMENTARY INFORMATION	100
REFERENCES	116
CHAPITRE III : IMPORTANT DISPERSAL CAPACITY OF ONE MAIN SPECIES VECTOR OF BLUETONGUE AND SCHMALLEMBERG VIRUSES <i>CULICOIDES OBSOLETUS</i> (DIPTERA: CERATOPOGONIDAE) REVEALED BY LANDSCAPE GENETIC ANALYSES AT FRENCH SCALE	126
ABSTRACT.....	129
BACKGROUND.....	130
METHODS.....	133
RESULTS	139
DISCUSSION	142
CONCLUSIONS.....	146
FIGURES AND TABLES	148
SUPPLEMENTARY INFORMATIONS	152
REFERENCES	164
DISCUSSION GENERALE.....	172
I) LA TAXONOMIE MOLECULAIRE DE <i>C. OBSOLETUS</i> : L'ARBRE QUI CACHE LA FORET.....	173
II) L'HISTOIRE EVOLUTIVE DE <i>C. OBSOLETUS</i> EN EUROPE : UN MORCEAU DU PUZZLE.....	178
III) DE LA DIRECTION DU VENT A LA DENSITE D'HOTES : LES FACTEURS ENVIRONNEMENTAUX QUI FAVORISENT LA DISPERSION DES <i>CULICOIDES</i>	187
IV) DES SIMULATIONS DE CIRCULATION ET UNE LUTTE ANTI VECTORIELLE ADAPTEE A DE TELS FLUX DE GENES ?	191
CONCLUSION GENERALE	195
REFERENCES.....	197
RESUME	241

Table des illustrations : figures

Introduction générale

- Figure 1.** Photos de *Culicoides*. (A et B) *Culicoides nubeculosus*, (C) Taille d'un *Culicoides* (*Culicoides impunctatus* à gauche) par rapport à un moustique (à droite) pendant le gorgement, (D) Patron alaire de *Culicoides obsoletus*, (E) Patron alaire de *Culicoides scoticus*. 30
- Figure 2.** Représentation schématique du cycle de développement et de vie des *Culicoides* (Source : Office vétérinaire fédérale 2010, Thèse Roger Venail 2014, Purse et al 2005, Purse et al 2015, Thèse Bruno Mathieu 2011, source des dessins : <http://illuminationstudios.com/>) 32
- Figure 3.** Distribution géographique des six principaux virus transmis par les *Culicoides* (Purse *et al*, 2015). AHSV ; virus de la peste équine ; AKAV, virus Akabane ; BTV, virus de la fièvre catarrhale ovine ; EHDV, virus de la maladie hémorragique ; OROV, virus de l'Oropouche ; SBV, virus de Schmallenberg. 34
- Figure 4.** Représentation schématique de la dispersion stratifiée des *Culicoides*. Le rouge implique une dispersion active et le bleu une dispersion passive. L'épaisseur des flèches rouges et bleues est représentative de l'intensité de la dispersion (Source Office vétérinaire fédérale 2010, Thèse Roger Venail 2014, Purse *et al* 2005, Purse *et al* 2015, Thèse Bruno Mathieu 2011). 41
- Figure 5.** Aire de répartition mondiale de *C. obsoletus*. Le gris indique les pays dans lesquels l'espèce est présente, le blanc indique les pays dans lesquels n'a pas été rapportée jusqu'à présent. 45

Chapitre I

- Figure S1. Maximum likelihood phylogenetic tree using 16S rDNA.** Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates). 70
- Figure S2. Maximum likelihood phylogenetic tree using *cox1*.** Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates). 71
- Figure S3 -Maximum likelihood phylogenetic tree using 28S rDNA.** Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates). 72

Chapitre II

- Figure 1. Graphical representation of the 4 scenarios simulated using an approximate Bayesian computation (ABC) approach on microsatellite data.** Details of all scenarios and parameters are shown in Supplementary Table S2. Red circles correspond to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence times. NO: population from Norway (Narvik), SP: population from Spain (Torre del campo), FR: population from France (Les Hîtres), GR: population from Greece (Velesino). ... 93
- Figure 2. Median-joining haplotype networks of *cox1* mitochondrial sequences of *C. obsoletus*.** The size of the circles is proportional to the number of individuals with that haplotype. Colors represent the country of sampled individuals. The length of the branches separating haplotypes is proportional to the number of mutational steps between them. 94

Figure 3. Mitochondrial genetic differentiation and isolation-by-distance analyses of <i>C. obsoletus</i>. Results of the smoothing of pairwise measures performed with MAPI and based on (a) $F_{ST}/(1-F_{ST})$ and (b) Isolation-by-distance analysis, Mantel test and a kernel density estimate of pairwise correlations from each sample between $(F_{ST}/(1-F_{ST}))$ and geographic distance. In panels (a), genetic dissimilarity is represented by a color scale ranging from brown (lower genetic dissimilarity) to blue (higher genetic dissimilarity). The Canary Islands populations are not considered in these analyses.	95
Figure 4. Mismatch distribution plots for 106 haplotypes of <i>cox1</i> of <i>C. obsoletus</i>. (a) The mean estimate and the 90, 95 and 99% highest posterior density (HPD) limits are indicated.	96
Figure 5. Microsatellite genetic clustering and genetic diversity of <i>C. obsoletus</i> sampled populations. Results of the genetic clustering analyses performed with STRUCTURE (a) and rarefied allelic richness of populations (b). Panels (1) corresponds to an enlargement of the geographical area (a) of the panel circled in red. One specific color has been assigned to each genetic cluster in panels (a) and (1).	97
Figure 6. Microsatellite genetic differentiation and isolation-by-distance analyses of <i>C. obsoletus</i>. Results of the smoothing of pairwise measures performed with MAPI and based on (a) $F_{ST}/(1-F_{ST})$ and (b) Isolation-by-distance analysis, Mantel test and a kernel density estimate of pairwise correlations from each sample between $(F_{ST}/(1-F_{ST}))$ and geographic distance. In panels (a), genetic dissimilarity is represented by a color scale ranging from brown (lower genetic dissimilarity) to blue (higher genetic dissimilarity). The Canary Islands population is not considered in these analyses because of its insularity.....	98
Figure S1. Identification of the optimal number of genetic clusters (K) inferred by STRUCTURE using the delta(K) and L'(K) methods. (a) and (b) panels corresponds to analyses with all 829 individuals, and (c) and (d) panels corresponds to STRUCTURE analyses on central Europe dataset.	110
Figure S2. Genetic clustering of <i>C. obsoletus</i>. (a) Results of the genetic clustering analyses performed with STRUCTURE on complete dataset for K=2 (second most probable K), and (b) Results of the genetic clustering analyses performed with STRUCTURE on dataset of central Europe for K=2 (most probable K). Panels (1) corresponds to an enlargement of the geographical area (a) of the panel circled in red. One specific color has been assigned to each genetic cluster in panels (a) and (b).	111
Figure S3. Map of genetic diversity of 2409 <i>cox1</i> sequence of <i>C. obsoletus</i>.	112
Figure S4. Neighbor-joining tree of <i>C. obsoletus</i> population based on the distance of Cavalli-Sforza & Edwards (1967) computed on microsatellite dataset. The belonging of populations to a genetic cluster defined by the STRUCTURE Analysis is represented by the colors, orange, green and blue for cluster 1, 2 and 3 respectively. Bootstraps values were calculated over 1000 replicates and are represented as percentage and only values superior to 60% are reported.....	113
Figure S5. Heatmap of $F_{ST}/(1-F_{ST})$ of all population for microsatellite dataset of <i>C. obsoletus</i> in Europe. Levels of genetic distance is represented by color gradient, red mean low level of differentiation and yellow mean high levels of differentiation. The belonging of populations to a genetic cluster defined by the STRUCTURE Analysis is represented by the colors, orange, green and blue for cluster 1, 2 and 3 respectively.....	114
Figure S6. PCA summary statistics simulated under scenario 1, the most probable scenarios. Yellow dot represents the real data set. Large dots correspond to simulated posteriors and the small circles correspond to simulated statistics from the prior statistics.....	115

Chapitre III

Figure 1. Genetic clustering and genetic differentiation of <i>C. obsoletus</i>. Results of the genetic clustering analyses performed with GENELAND (a) and STRUCTURE (b), as well as smoothing of pairwise measures performed with MAPI and based on (c) Rousset's (a_R) and (d) FCA inter-individual genetic distances (see the text for further details). One specific color has been assigned to each genetic cluster in panels (a) and (b). In panels (c) and (d), genetic dissimilarity is represented by a color scale ranging from red (lower genetic dissimilarity) to blue (higher genetic dissimilarity).....	148
Figure 2. Results of anisotropic isolation-by-distance analyses. Polar plots show the correlation between geographical projected distances by angle and inter-individual genetic distances (a_R , LKC, and FCA)...	149
Figure 3. Bearing analysis: correlation between genetic (a_R, LKC, and FCA) and geographical distances as a function of the angle between sampling sites. Circles and crosses indicate significant and non-significant values, respectively. Abbreviations "E", "W", "S" and "N" refer to "East", "West", "South" and "North", respectively.....	150
Figure S1. Environmental variables tested as potential factors that could impact inter-individual genetic differentiation of <i>C. obsoletus</i> in France (raster cell resolution: 0.04 arcmin).	158
Figure S2. Map of wind direction averaged from 2000 to 2010. Sampling sites are represented by black points. The color scale represents the wind direction from 0 to 360 degrees from North The arrowheads indicate the exact wind direction of each raster pixel.	159
Figure S3. Analytical workflow	160
Figure S4. Isolation-by-distance analyses: density plots, Mantel tests and linear regressions performed with each inter-individual genetic distance considered in this study.	161
Figure S5. Identification of the optimal number of genetic clusters (K) inferred by STRUCTURE using the delta(K) and L'(K) methods.	162
Figure S6. Population genetic structure results by clustering analyses performed STRUCTURE. One specific color has been assigned to each inferred genetic cluster.	163

Discussion générale

Figure 1. Hypothèse de scénario d'expansion mondiale de <i>C. obsoletus</i>. Les zones grises indiquent des pays dans lesquels <i>C. obsoletus</i> est présent. Les zones vertes représentent des zones d'admixture des populations. La zone bleue correspond à une zone en Europe qui n'a pas pu être incluse dans nos analyses mais qui devra être considérée dans le futur. Les cercles blancs indiquent l'emplacement de refuges glaciaires. Ne désigne une expansion démographique. Le beige désigne l'abaissement du niveau de la mer durant la glaciation.	186
---	-----

Table des illustrations : tableaux

Chapitre I

Tableau S1. Information on adult female <i>Culicoides</i> sampling sites and results of specific assignation using cox1.	65
Tableau S2. Primers used for PCRs and sequencing in this study.	67
Tableau S3. Reference sequences used for specific assignation.	68
Tableau S4. Information on all <i>Culicoides</i> sequenced and GenBank accession numbers.	69

Chapitre II

Tableau 1. Result of confidence in scenario selection based on posterior probabilities to retrace evolutionary history of <i>C. obsoletus</i> in Europe by approximate Bayesian Computation (ABC) on microsatellite data. 95% confidence intervals and type I and II errors were presented. Red circles correspond to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence between two taxa. NO: Norwegian population (Narvik), SP: Spanish population (Torre del campo), FR: French population (Les Hitres), GR: Greek population (Velesino).	99
Tableau S1. Metadata of the 2409 sequences of cox1 <i>C. obsoletus</i>	100
Tableau S2. Result of most probable scenario for all ABC simulation. 95% confidence intervals and type I and II errors were presented to. Red circles corresponds to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence between two taxa. NO: Norway population (Narvik), SP: Spain population (Torre del campo), FR: French population (Les Hitres), GR: Greece populations (Velesino).	101
Tableau S3. Demographic, historical and mutational prior parameters for ABC analyses. Effective population sizes (N) are expressed in number of diploid individuals and times of events (t) in number of generations going back to the past.	102
Tableau S4. Population genetic diversity of 2409 cox1 sequence of <i>C. obsoletus</i> in Europe. Pi: Nucleotide diversity per site, Nh: haplotype number.	103
Tableau S5. Primers of the 11 microsatellite markers used to genotype <i>C. obsoletus</i> populations. “DYE”: fluorochrome name, “Tm”: half denaturation temperature in degree, “bp”: base pairs.	105
Tableau S6. Microsatellite dataset of 829 <i>C. obsoletus</i> in Europe.	106
Tableau S7. Sampling sites and associated genetic diversity for microsatellite data set. N: sample size, Ho: observed heterozygosity, Hs: expected heterozygosity, F _{IS} : Fixation index, Ar : allelic richness, Ne : sample size. P-values obtained for the Hardy-Weinberg test were adjusted for multiple comparisons at the nominal level (5%) of 0.00017.	107
Tableau S8. Genetic diversity by locus. Ho: observed heterozygosity, Hs: expected heterozygosity, F _{IS} : Fixation index, SAD: short alleles dominance, CI: confidence indice.	108
Tableau S9. Posterior checking using all the summary statistics for ABC inferences of scenario 1, the most probable scenarios.	109

Chapitre III

Tableau 1. Results of multiple regressions on distance matrices (MRDM) and additional parameters derived from commonality analysis (CA). Pearson's correlation coefficient (r), significant regression coefficient (β), as well as unique (U) and common (C) contributions of environmental distances to the variance in the dependent variable. [C]/[R] indicate if the considered environmental raster was treated as a conductance or as a resistance factor, and k corresponds to the re-scaling parameter used to transform the initial raster file.	151
Tableau S1. Sampling sites and associated genetic diversity. N: sample size, Ho: observed heterozygosity, Hs: expected heterozygosity, Fis: Fixation index.....	152
Tableau S2. Reference sequences used for specific assignation.	153
Tableau S3. Origin and numbers of individuals used to build up the DNA library necessary for the development of the microsatellite markers.	154
Tableau S4. Primers of the 13 microsatellite markers used to genotype <i>C. obsoletus</i> populations. "DYE": fluorochrome name, "TFm": half denaturation temperature in degree, "bp": base pairs.....	155
Tableau S5. Genetic diversity by locus. Ho: observed heterozygosity, Hs: expected heterozygosity, FIS: Fixation index, SAD: short alleles dominance.....	156
Tableau S6. Results of univariate analyses: determination coefficients R²'s estimated from univariate regressions between genetic and environmental distances. [C]/[R] indicate if the considered environmental raster was respectively treated as a conductance or resistance factor for the computation of environmental distances with circuit theory. Q: Difference between environmental MRDM R ² and null raster MRDM R ² . The results for K=10 are not present here as they were non-significant.	157

Introduction générale

La dispersion est un processus universel dans le monde vivant. Que ce soit au stade œuf, juvénile ou à l'âge adulte, tout être vivant possède une phase de dispersion au cours de sa vie. La dispersion désigne le mouvement ponctuel d'individus entre le lieu de naissance et de reproduction, ou entre différents lieux de reproduction (Austerlitz *et al*, 2007; Baguette *et al*, 2013; Baguette and Van Dyck, 2007; Benton and Bowler, 2012; Cayuela *et al*, 2018; Clobert *et al*, 2009; McMichael, 2013; Semenza *et al*, 2016). Les organismes dits fixés, comme les plantes ou les coraux, peuvent disperser de manière passive via des propagules (œufs, gamètes). Les espèces mobiles, au contraire, peuvent effectuer des déplacements durant l'intégralité de leurs cycle biologique. La dispersion permet la mise en contact avec de potentiels partenaires de reproduction. Elle peut augmenter la valeur sélective de l'individu et permettre la colonisation de nouveaux territoires. Des adaptations morphologiques et comportementales favorisant la dispersion ont été sélectionnées au cours de l'évolution afin de permettre aux propagules et aux différents stades de vie de se disperser de façon active ou passive (ailes, crochet sur les akènes, expansion ailée ou hydrophobie des graines).

Dispersion, migration et conséquences génétiques

Les termes "*dispersion*" et "*migration*" sont souvent utilisés en synonyme (Broquet and Petit, 2009; Cayuela *et al*, 2018). La migration, au sens strict du terme, désigne « *le mouvement des organismes entre populations* » (Lefevre *et al*, 2016). Ces mouvements peuvent être récurrents, à double sens, aller et retour. Cependant, dans le lexique de la biologie évolutive, le terme « *migration* » est utilisé dans un sens plus précis (Lefevre *et al*, 2016), comme « *le mouvement d'individus entre des populations adjacentes et leur résidence pendant une durée suffisante pour contribuer au pool génétique des populations d'accueil permettant de réduire la différenciation génétique entre populations* » (Frankham, 2009). La migration (terme préféré à dispersion) est alors une force évolutive, car elle est capable de modifier les fréquences alléliques des populations. Un événement de dispersion générant une reproduction dans le nouvel environnement est dit efficace (Broquet and Petit, 2009; Cayuela *et al*, 2018). Les déplacements réguliers d'individus de manière quotidienne sans reproduction ne sont donc pas considérés comme de la dispersion au sens évolutif puisqu'il n'y a pas d'échange de matériel génétique (Andersson and Bodin, 2009). Lors d'un événement de dispersion active, l'échange

d'individus et de gamètes permet une contribution au pool génétique de la population accueillant les migrants et correspond à un flux de gènes entre les deux populations, tendant à homogénéiser les fréquences alléliques entre ces populations (Lowe and Allendorf, 2010; Takken *et al*, 2008). Cependant, la façon dont les individus se déplacent dans l'espace (le modèle de migration), influence fortement la distribution spatiale du polymorphisme, c'est-à-dire la structure spatiale des populations. Dans l'ensemble de ce manuscrit, nous utiliserons le mot dispersion au sens évolutif du terme, c'est-à-dire avec une contribution au pool génétique.

Le modèle le plus simple pour représenter la dispersion est conceptualisé par Wright (1931) avec le modèle en îles (Wright, 1931). Ce modèle décrit une population de taille infinie (population continent) et une infinité de sous-populations adjacentes de taille finie (population île). En raison de leur taille limitée, les sous-populations sont impactées par la dérive génétique, définie comme la modification aléatoire des fréquences des alléliques au cours des générations. Cet échantillonnage aléatoire des allèles va conduire, en l'absence de mutations, à la fixation de l'un des allèles et à une homogénéisation de la sous-population. En l'absence de migration, cela va générer de la divergence génétique entre les sous-populations. Plus la population est de petite taille, plus la dérive génétique agit rapidement. La taille d'une population détermine ainsi le rythme de la dérive génétique. Cependant, le nombre d'individus au sein de la population participant de manière active à la reproduction est rarement égal au nombre total d'individus dans la population. La taille efficace d'une population (N_e) désigne alors le nombre d'individus d'une population « idéale » de Wright-Fisher (population où la taille efficace est égale au nombre total d'individus dans la population) pour laquelle on aurait un degré de dérive génétique équivalent à celui de la population réelle (où la taille efficace est rarement égale au nombre total d'individus dans la population). Cette notion permet de quantifier le rythme auquel une population perd sa diversité génétique par dérive génétique.

En l'absence d'autres forces évolutives, la migration tend à homogénéiser les fréquences alléliques entre populations. Plus les flux de gènes seront importants entre les populations, plus la diversité et les fréquences des pools génétiques auront tendance à s'homogénéiser (Saastamoinen *et al*, 2018). Les flux de gènes entre populations limitent l'hétérogénéité génétique (Abbott, 2011; Bjornstad *et al*, 1999; Hastings *et al*, 1993; Ranta *et al*, 1997) et l'extinction de populations locales (Gotelli, 1991; Lowe and Allendorf, 2010). Il est à noter que le modèle en îles fait l'hypothèse, peu réaliste, d'un équilibre entre migration, mutation et dérive (de Meeus *et al*, 2007). La mutation correspond aux erreurs de copie lors de la duplication de

l'ADN durant la construction des gamètes. Elle va modifier les fréquences alléliques des loci qu'elle affecte. La mutation à elle seule ne représente pas une force très puissante de l'évolution des populations, car les taux de mutation sont en général assez bas (de Meeus *et al*, 2007). Cependant, c'est la seule source de nouveauté héritable et associée à la dérive et à la sélection, la mutation représente une clé de l'évolution et de la structure génétique des populations.

La notion de distance entre sous-populations a été introduite par le modèle d'isolement par la distance émis par Wright (1943). Plus réaliste que le modèle en îles, ce modèle prend en compte la dispersion des individus sur de courtes distances. La distance va ainsi limiter la probabilité que les individus s'accouplent les uns avec les autres et donc la différenciation génétique va s'accumuler avec la distance (Wright, 1943). D'autres modèles théoriques font intervenir une composante géographique où l'éloignement des sous-populations va influencer les probabilités d'échanges de gènes, comme le modèle en pas japonais "*stepping stone model*" (Kimura 1953) (où les migrants dispersent uniquement entre populations adjacentes) ou le « *neighbourhood model* » (Leblois *et al*, 2004; Rousset, 2000) (où l'apparement entre individus diminue avec la distance géographique qui les séparent, même sans barrière à la dispersion). Ainsi, le modèle de migration influence la distribution spatiale du polymorphisme. Tous ces modèles tentent de prédire l'agencement spatial de la diversité génétique (Gaggiotti and Smouse, 1996; Hanski, 1998; Hanski and Gaggiotti, 2004; Kimura, 1953; Levins, 1969). Cependant, ces modèles résumant le paysage à la distance géographique qui sépare les différentes parcelles et négligent l'hétérogénéité du paysage.

Dispersion et paysage

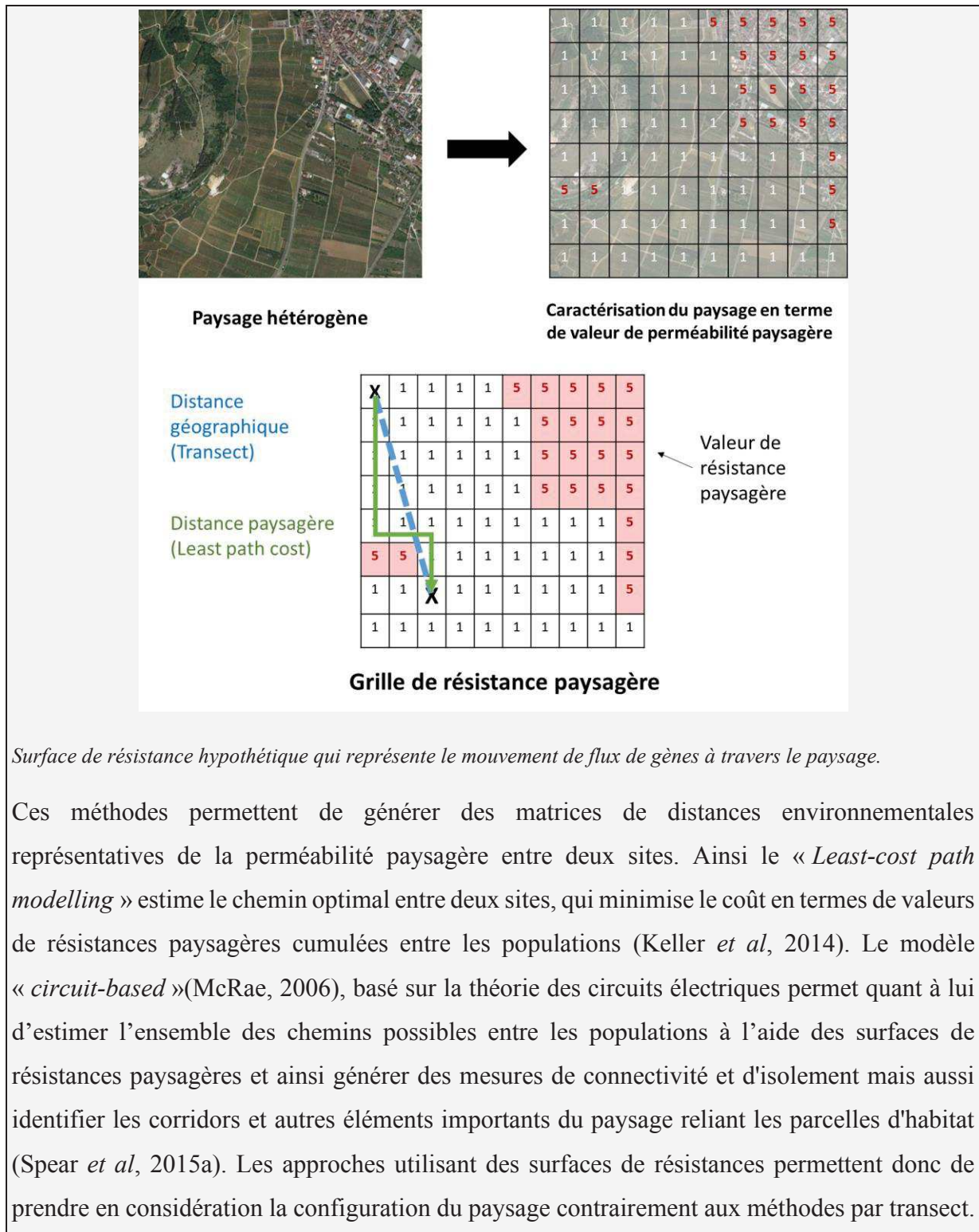
En écologie, le paysage est défini comme un ensemble d'interactions et de processus au sein d'une aire géographique spatialement hétérogène dont l'échelle et l'étendue sont pertinentes avec la résolution du phénomène (Turner, 2005). Le paysage, d'un point de vue individuel, se définit par la niche écologique propre à l'individu, c'est-à-dire un ensemble de ressources biotiques et abiotiques réparties de manière hétérogène dans l'espace. Les habitats dans lesquels les organismes vivent, influencent les processus de dynamiques démographiques, de flux génétiques et d'évolution adaptative. L'hétérogénéité du paysage est au centre de la théorie de la biogéographie insulaire de MacArthur & Wilson (1967) (MacArthur, 1967). Cette théorie formalise qu'au sein d'un paysage hétérogène, des individus vont se distribuer

spatialement en populations en fonction des parcelles d'habitats qui leurs sont favorables (les îles). Les organismes ne sont donc pas répartis de manière homogène dans le paysage (Sork, 2016; Wiens *et al*, 1993).

La théorie de la biogéographie insulaire décrit également que le nombre d'espèces dépend de la surface d'habitat disponible et de la distance géographique qui sépare l'île du continent. Ainsi, plus une île sera éloignée du continent, plus le nombre de propagules atteignant l'île sera faible. Bien qu'initialement développée pour un système insulaire, cette théorie s'applique à tous types de populations isolées au sein de zones spatiales favorables, et sera par la suite formalisée sous l'appellation de métapopulation par Levins (1969) (Levins, 1969). En écologie, la capacité de dispersion d'un organisme va influencer sa capacité de colonisation de nouveaux patches d'habitats. La biogéographie insulaire de MacArthur & Wilson et la théorie des métapopulations de Levins ont également donné naissance au modèle sources-puits de Pulliam (1988) (Pulliam, 1988). Ce modèle reprend l'idée principale de la biogéographie insulaire en y ajoutant la possibilité de sauvetage de petites populations par l'arrivée de nouveaux individus issus de plus grosses populations, permettant de renforcer le pool génétique local.

Encadré 1. Représentation et quantification de la connectivité du paysage

Les données permettant de représenter un facteur environnemental dans un paysage peuvent être de nature continue, telles que l'humidité ou la température, ou ordonnées en catégories telles que les différents types de couverture paysagère. L'utilisation de grille ou « *raster* » en anglais, permet de cartographier l'hétérogénéité paysagère. La grille représentant le paysage est alors découpée en pixels d'une taille prédéfinie appelé grain. Puis, une valeur représentative du facteur environnemental souhaité est attribuée à chaque pixel formant ainsi une grille environnementale. Différents modèles théoriques permettent de quantifier la connectivité paysagère entre populations. Les approches basées sur les transects, visent à créer un tracé en ligne droite entre deux sites puis de faire l'inventaire de chaque caractéristique paysagère pouvant limiter la dispersion (Angelone *et al*, 2011). D'autres approches utilisent des surfaces de résistances paysagères. Les surfaces de résistances sont des grilles spatiales, dans lesquelles chaque pixel ne représente pas le facteur environnemental en lui-même, mais une valeur directe du frein que représente le facteur environnemental sur le flux de gènes. Le chemin le moins coûteux calcule le chemin qui limite la résistance paysagère entre deux points (X noir) qui n'est pas toujours celui avec la distance géographique ou linéaire la plus courte.



La structure et la composition du paysage sont les moteurs de la distribution des populations et des organismes. Par conséquent, le paysage va impacter la dispersion entre les patches d'habitats favorables, c'est le concept de connectivité (Encadré 1). La connectivité paysagère structurelle représente l'agencement spatial et la quantité des parcelles favorables ainsi que les corridors

permettant de les relier (Lutz Tischendorf and Fahrig, 2000; Taylor *et al*, 1993; Uezu *et al*, 2005). La connectivité fonctionnelle désigne l'impact des éléments paysagers sur la dispersion. Selon les différents traits d'histoire de vie propres à l'organisme étudié, certains facteurs environnementaux peuvent limiter ou favoriser la dispersion (Taylor *et al*, 1993; Uezu *et al*, 2005). Lorsque le paysage facilite la dispersion d'un individu dans l'espace, on parle alors de perméabilité paysagère. Au contraire, lorsque celui-ci est ralenti par une composante environnementale, on parle de résistance paysagère. Toute modification de la structure, de la composition ou de la qualité de ce paysage peut impacter la dispersion des organismes. Par exemple, lors de la fragmentation d'un habitat, les flux de gènes sont perturbés par la scission de l'espace favorable en plusieurs espaces favorables mais de plus faible surface, séparés par des zones où l'aptitude phénotypique de l'individu est plus faible.

Les différents types de dispersion

La littérature décrit le processus de dispersion en trois étapes : l'émigration, la phase mobile entre les deux matrices paysagères et l'immigration (Baguette and Van Dyck, 2007; Clobert *et al*, 2009). L'émigration désigne la phase de départ de l'individu de sa matrice paysagère de naissance ou de reproduction. A l'inverse, l'immigration fait référence à l'installation dans la nouvelle matrice faisant office de nouveau lieu de reproduction.

La dispersion est guidée par les caractéristiques biotiques et abiotiques paysagères (Bowler and Benton, 2005; Chaine *et al*, 2013; Ronce and Clobert, 2012). Durant les phases d'émigration et d'immigration, ce sont particulièrement les facteurs biotiques comme la prédation, la compétition ou la consanguinité qui peuvent influencer le processus (Bowler and Benton, 2005; Clobert *et al*, 2004; Clobert *et al*, 2009). La dispersion est dite active quand le mouvement est intentionnel, principalement guidé par la recherche de partenaires de reproduction (Cote *et al*, 2017; Ronce, 2007). La phase mobile individuelle de transition entre deux sites est quant à elle, plutôt impactée par les caractéristiques abiotiques du paysage (Baguette *et al*, 2013; Baguette and Van Dyck, 2007; Pfluger and Balkenhol, 2014). Bien que cette phase mobile puisse être active par la marche, le vol ou la nage, des facteurs extrinsèques peuvent générer des événements de dispersion passive. L'individu dispersé n'est alors pas maître de la direction ou de l'intensité de dispersion, créant une migration à direction asymétrique (Pringle *et al*, 2011). Ces facteurs peuvent être les courants océaniques, les vents dominants ou la dispersion d'organismes vivants utilisés comme support de dispersion, permettant souvent des dispersions

sur de très grandes distances. La dispersion passive est documentée pour les pollens (Austerlitz *et al*, 2007; Dutech *et al*, 2005; Jay *et al*, 2013), les spores fongiques (Rieux *et al*, 2014), certaines espèces aquatiques (Simpson *et al*, 2014), des insectes de petite taille (Eagles *et al*, 2012) et les stades immatures de tiques par les oiseaux (Bohonak and Jenkins, 2003; Burgess *et al*, 2015; Gillespie *et al*, 2012).

La dispersion passive à longue distance chez les arthropodes

Les insectes sont particulièrement aptes à être dispersés sur de longues distances de manière passive, que ce soit par l'intervention d'un facteur environnemental, le vent ou par le mouvement d'autres espèces (incluant les activités des populations humaines). Cette dispersion passive contrebalance les capacités de dispersion active souvent limitées à quelques kilomètres (Reynolds *et al*, 2006).

La petite taille des insectes est un facteur facilitant la dispersion passive par le vent. Les vents de hautes altitudes permettent à certaines espèces de se disperser sur des distances pouvant aller jusqu'à plusieurs centaines de kilomètres. Phénomène considéré comme des événements relativement rares, certains auteurs mettent en évidence l'importance des saisons dans la fréquence des épisodes de dispersion par le vent (Chapman *et al*, 2015; Johnson *et al*, 1962; Miao *et al*, 2012; Parry *et al*, 2015; Reynolds and Reynolds, 2009; Reynolds *et al*, 2006; Sanders *et al*, 2011; Wotton *et al*, 2019). Ce type de dispersion est documenté chez de nombreuses espèces de pucerons (Hemiptera : Aphididae), guêpes (Hymenoptera : Torymidae) et *Culicoides* (Diptera : Ceratogopognidae). Elle est aussi rapportée chez des espèces plus imposantes comme les mouches noires (Diptera : Simuliidae), les criquets (Orthoptera : Acrididae) ou les syrphes (Diptera : Syrphidae) (Anderson and Sturtevant, 2011; Johnson *et al*, 1962; Kerdelhue *et al*, 2009; Lesieur *et al*, 2018; Miao *et al*, 2012; Rainey, 1973; Reynolds *et al*, 2006; Sanders *et al*, 2011; Sanogo *et al*, 2020; Sellers, 1980; Wotton *et al*, 2019).

Les caractéristiques abiotiques des courants d'air sont importantes, et toutes ne sont pas favorables à une dispersion passive. En effet, elles doivent permettre la survie des organismes durant l'épisode de dispersion passive (Reynolds *et al*, 2006). Les espèces de grande taille semblent pouvoir tolérer des vents violents (au-delà de 100 km/h) (Parry *et al*, 2015). Chez les moustiques, le vol se produit à une altitude inférieure à 10 m au-dessus du sol, bien que des individus aient pu être collectés à des altitudes supérieures (Huestis *et al*, 2019). L'altitude est un paramètre important à considérer, la vitesse du vent étant généralement plus élevée à haute

altitude (Westbrook and Isard, 1999). Les températures basses et le stress du transport à haute altitude peuvent aussi impacter la survie de l'individu pendant la phase mobile et l'émigration (Sanogo *et al*, 2020).

Le vent n'est pas le seul moyen de dispersion à longue distance chez les arthropodes. En effet, la mobilité des populations humaines aujourd'hui est grande. La distance géographique n'est plus, comme cela a pu l'être dans l'histoire évolutive, une barrière infranchissable (Gippet *et al*, 2019). Les mouvements des populations humaines et animales et l'intensification des échanges commerciaux décuplent la capacité de dispersion de certaines espèces d'arthropodes. Une des conséquences est l'introduction dans de nouveaux territoires, d'espèces exotiques vectrices ou de ravageurs qui peuvent s'y établir et proliférer (Drake and Reynolds, 2012), engendrant des conséquences économiques ou sanitaires importantes.

Un exemple d'intérêt agricole ayant eu des conséquences économiques importantes est celui de l'introduction de la chrysomèle des racines du maïs (*Diabrotica virgifera virgifera*) en Europe via les échanges commerciaux. La chrysomèle est probablement originaire d'Amérique centrale (Krysan, 1986). Son aire d'extension recouvre actuellement toute l'aire de culture du maïs en Amérique du nord où les agriculteurs américains l'ont surnommée "l'insecte à 1 milliard de dollars" (Sappington *et al*, 2006). Elle est apparue en Europe en 1992 près de l'aéroport de Belgrade (Serbie) et a été repérée en France en 2002 où elle est considérée comme une espèce invasive. Les travaux de phylogéographie ont permis de reconstruire les routes d'invasion de la chrysomèle des racines du maïs à l'aide de marqueurs génétiques. Cette étude a montré que la distribution européenne de la chrysomèle des racines du maïs résultait de plusieurs introductions à partir d'Amérique du nord vers l'Europe. Au moins trois introductions ont successivement donné naissance aux foyers détectés en Serbie en 1992, en Italie en 2000 et en France en 2002. Une autre étude a montré que les nouveaux foyers observés en Alsace en 2003 et en Île-de-France en 2005 étaient issus de deux nouvelles introductions en provenance d'Amérique du nord, portant à cinq le nombre d'introductions transatlantiques de la chrysomèle des racines du maïs. L'origine exacte des foyers européens n'a pas été identifiée, mais le nord des États-Unis apparaît comme la région d'origine la plus probable (Ciosi *et al*, 2008; Miller *et al*, 2005).

La dispersion passive des espèces vectrices et des pathogènes transmis

Parce que certains arthropodes sont hématophages et sont impliqués dans la transmission de pathogènes (Encadré 2), leur dispersion et la colonisation de nouveaux territoires permettent aussi l'extension de l'aire de distribution des pathogènes transmis (Gippet *et al*, 2019; McMichael, 2013). Les pathogènes peuvent être introduits dans les populations anciennement installées par un hôte infecté ou lors de la dispersion d'arthropodes infectés.

Le moustique-tigre (*Aedes albopictus* ou *Stegomyia albopicta* selon les auteurs), espèce de moustique de la famille des Culicidae est l'une des cent espèces les plus invasives au monde (Global Invasive Species Database), actuellement présente dans 100 pays sur les cinq continents (Lindh *et al*, 2019; Rezza *et al*, 2007; Semenza and Suk, 2018; Tatem *et al*, 2006; Venturi *et al*, 2017). Cette espèce de moustique, originaire d'Asie du sud-est, doit son expansion au flux croissant de marchandises lié au transport international. Les œufs d'*Aedes* peuvent survivre plusieurs mois en absence d'eau, ce qui facilite leur expansion par transport commercial international. Les adultes s'adaptent aussi très bien aux conditions environnementales des régions tempérées. Son installation est source de nuisance dans les métropoles. De par, les virus qu'elle transmet (au moins 22 espèces d'arbovirus), cette espèce est responsable de catastrophes sanitaires dans de nombreux territoires. Les mouvements des populations humaines ont permis l'introduction de pathogènes transmis par *Ae. albopictus* dans des zones précédemment indemnes du pathogène et du moustique. *Aedes albopictus* joue aujourd'hui un rôle majeur dans la transmission du virus de la dengue en Afrique (même si le rôle d'*Ae. aegypti* est central par ailleurs), du virus de Chikungunya, et de Zika (Gratz, 2004). Historiquement, le commerce triangulaire est reconnu pour être responsable de l'introduction d'*Aedes aegypti*, espèce de moustique afrotropicale, en Amérique du sud (Bryant *et al*, 2007), et des pathogènes transmis via les populations d'esclaves déplacées. La dispersion passive d'arthropodes infectés par les vents est documentée pour des pathogènes transmis par des moustiques ou des moucheron.

L'introduction d'animaux virémiques et la dispersion passive de moustiques infectés étaient les hypothèses classiquement avancées pour expliquer l'épidémie de fièvre de la vallée du Rift survenue en Egypte en 1977 depuis le Soudan (Sellers *et al*, 1982). A la lumière des connaissances actuelles sur l'importance des mouvements d'animaux entre l'Egypte et le Soudan, il semble plus probable de privilégier la piste du déplacement d'animaux. Sellers et ses collaborateurs ont étudié les trajectoires des vents dans l'ouest des Etats Unis et dans la province

de Manitoba au Canada pour expliquer les foyers d'encéphalomyélite équine occidentale observés chez les chevaux de cette région en 1987 (Sellers and Maarouf, 1993). Ils montrent que des femelles moustiques de l'espèce vectrice de ce virus, *Culex tarsalis*, auraient pu se disperser passivement grâce aux vents sud-nord, soufflant du Texas et de l'Oklahoma vers le nord des Etats-Unis, puis vers la province du Manitoba. Au Japon, l'introduction d'une espèce de moustique, *Culex tritaeniorhynchus*, grâce aux vents de haute altitude permet des incursions récurrentes du virus de l'encéphalite japonaise. (Duvallet *et al*, 2017; Sellers, 1980). En Australie, la dispersion passive, grâce au vent, de cette espèce peut atteindre 648 km, à une altitude allant de 50 à 310 m, permettant ainsi l'introduction du virus de l'encéphalite japonaise dans ce territoire (Duvallet *et al*, 2017).

Encadré 2 : Arboviroses, vecteur et cycle de transmission

Le terme « arboviroses », de l'anglais ARthropod-BORne VIRUSes, désigne des maladies virales infectieuses transmises par les arthropodes (Duvallet *et al*, 2017). Depuis toujours, l'humanité fait face à des épidémies virales transmises par des arthropodes, comme la fièvre jaune ou la dengue. C'est en 1877 que Manson démontre que le moustique est l'hôte obligatoire d'un ver, la filaire de Bancroft (*Wuchereria bancrofti*), celui-ci étant la cause de la filariose lymphatique. Cette découverte de la capacité de certains arthropodes à transmettre des pathogènes est aussi rapidement démontrée pour les agents responsables du paludisme (1898), de la fièvre jaune (1900) et la dengue (1903). On définit un vecteur comme un arthropode hématophage, qui assure la survie, la multiplication ou la transformation et la transmission active d'un agent pathogène infectieux ou parasitaire d'un vertébré à un autre vertébré (Duvallet *et al*, 2017). Les agents pathogènes concernés sont variés tels que les bactéries, les virus, les protozoaires, ou encore les helminthes.

La transmission biologique est qualifiée lorsque le pathogène effectue une partie ou la totalité de son cycle de réplication à l'intérieur de l'arthropode vecteur (Rodhain *et al*, 1985). La transmission mécanique se distingue de la transmission biologique par l'absence de phase de réplication ni d'évolution du pathogène dans le vecteur (Rodhain *et al*, 1985; Wagner and Fortin, 2015). Un système vectoriel comprend un agent pathogène, un ou plusieurs vecteurs, et des vertébrés faisant office d'hôtes ou de réservoirs.

C'est aussi l'hypothèse de transport passif par les vents qui est émise pour expliquer l'introduction de la fièvre catarrhale ovine dans l'ouest de la Turquie, en octobre 1977. La fièvre catarrhale ovine (FCO) est une maladie dont l'agent étiologique est transmis par des petits moucheron de la famille des Ceratopogonidae, les *Culicoides*. Bien qu'il ait été possible pour le virus d'avoir été introduit dans la région par le mouvement d'animaux infectés, des vents favorables sud-est auraient également pu transporter des moucheron infectés de Chypre, où la maladie était présente. L'étude établit précisément que dans la nuit du 14 au 15 octobre 1977, les vents sud-est auraient pu amener des moucheron infectés par le virus, après un vol de 15 h à une altitude estimée de 500 m, à une température d'environ 20 degrés. Cette hypothèse est aussi avancée pour expliquer l'introduction du virus d'Akabane, autre virus transmis par les *Culicoides*, en 1980 dans la même région (Sellers, 1980).

Les *Culicoides* sont souvent cités, à juste titre ou avec moins de prudence, comme un modèle de vecteur très facilement dispersable par les vents, expliquant les foyers de FCO (Braverman and Chechik, 1996) ou d'autres maladies transmises. Il est vrai que, contrairement à *Ae. albopictus*, toutes les études ont démontré le faible nombre d'individus capables de se disperser via les transports internationaux (Carpenter *et al*, 2013). Cependant, une étude récente de modélisation réanalyse les articles de R. F. Sellers publiés entre 1977 et 1991 et montre que la possibilité que les foyers de FCO soient initiés par la dispersion de *Culicoides* infectés à longue distance par le vent est moins plausible que Sellers ne le soutient. Sellers se base en effet uniquement sur une simple analyse de la direction du vent et ne prend pas en compte la vitesse du vent. L'approche de modélisation confirme certains scénarios de dispersion d'individus infectés, mais elle révèle aussi la nécessité d'intégrer les données sur la répartition des populations d'hôtes sensibles (et d'autres covariables) dans le processus de modélisation pour expliquer les introductions observées (Durr *et al*, 2017). Nous détaillerons dans le chapitre I et II d'autres exemples de dispersion concernant le modèle *Culicoides*.

Cette introduction a présenté les enjeux autour de la dispersion des arthropodes vecteurs de pathogènes d'intérêts pour la santé humaine ou animale. Nous allons maintenant développer les outils permettant de décrire et de caractériser la dispersion de ces organismes.

I) L'étude de la dispersion chez les arthropodes vecteurs

I.1) Des approches directes parfois inadaptées

L'approche directe la plus classique pour estimer la dispersion d'organismes est la technique dite de « capture-marquage-recapture » (CMR) décrite en 1896 par Peterson (Lebreton *et al*, 2003; O'Brien *et al*, 1985). Cette technique démographique a été développée dans le but d'estimer les tailles de populations, mais elle peut être adaptée afin d'estimer la dispersion (Muir and Kay, 1998; Slade and Blair, 2000). Le principe de cette approche, est de piéger un grand nombre d'individus puis de les marquer. Ceux-ci sont ensuite libérés en un point central et des essais de recapture sont menés à des distances connues de leur site de libération. L'espace et le temps qui séparent ces deux étapes permettent d'estimer la capacité de dispersion individuelle (Lowe and Allendorf, 2010; Waples and Gaggiotti, 2006). La taille de la population, la survie ou encore la dispersion ont ainsi pu être évaluées chez des espèces vectrices de diptères comme *Aedes aegypti* (Muir and Kay, 1998; Russell *et al*, 2005) *Aedes notoscriptus* (Watson *et al*, 2000), *Glossina palpalis gambiensis* (Vreysen *et al*, 2013), *Anopheles gambiae* (Epopa *et al*, 2017) ou encore *C. obsoletus* (Kirkeby *et al*, 2013) grâce aux méthodes de CMR. Cependant, ces techniques sont souvent difficiles à mettre en place et ne sont pas applicables de manière systématique (Landguth *et al*, 2010; Manel *et al*, 2003; Storfer *et al*, 2007). Les diptères sont petits et fragiles, et le marquage, s'il est trop intrusif, peut impacter négativement la valeur sélective de l'individu, biaisant ainsi la probabilité de recapture (Hagler and Jackson, 2001). De plus, les techniques de marquage doivent être adaptées à la biologie de chaque espèce et ne sont souvent pas transposables d'une espèce à l'autre (Hagler and Jackson, 2001).

Différents types de marquage ont été utilisés sur les diptères : peinture, poudre fluorescente, marquage par des protéines (Hagler and Jackson, 2001; Hagler *et al*, 2014; Russell *et al*, 2005). Cependant, ces marquages sont généralement chronophages, coûteux et peuvent impacter la survie des individus. Une limitation de l'approche de CMR est la taille de la population et la capacité de dispersion de l'espèce. Lorsque l'espèce étudiée se disperse de manière importante, la probabilité de recapture devient extrêmement faible (Cayuela *et al*, 2018; Lowe and Allendorf, 2010; Waples and Gaggiotti, 2006). Ainsi, lors d'une étude récente avec une approche de CMR sur des populations de *Culicoides*, le taux de recapture était très faible (61 000 individus marqués pour 12 recapturés) (Kluiters *et al*, 2015a; Kluiters *et al*, 2015b). Un

autre exemple est l'étude d'Epopa et ses collaborateurs sur des moustiques du genre *Anopheles*. La distance maximale à laquelle un individu a été recapturé est de seulement 549 mètres, alors que l'on estime leur capacité de dispersion bien supérieure (Epopa *et al*, 2017).

Plus important encore, lors de la recapture d'un individu, rien n'assure qu'il ait pu s'accoupler dans la nouvelle population. Cette recapture peut être le reflet d'un simple déplacement physique dans l'espace (de Meeus *et al*, 2007). Ainsi, la dispersion efficace estimée par une approche CMR peut être surévaluée, en intégrant des événements de dispersion ne participant pas réellement aux flux de gènes entre populations. La reproduction réussie est un critère particulièrement important dans le cas des insectes vecteurs, car la transmission de pathogènes s'effectue lors du repas de sang, qui est conditionné par la reproduction.

Au vu des limites des approches directes dans l'évaluation de dispersion sur de longues distances, une approche indirecte (Slarkin, 1985) par l'utilisation de marqueurs moléculaires polymorphes, apparaît comme une solution plus adaptée afin d'explorer la dispersion efficace d'insectes vecteurs (de Meeus *et al*, 2007; Handley *et al*, 2011).

I.2) Evaluation des flux de gènes par méthodes indirectes lors de la dispersion d'insectes vecteurs

Les approches indirectes reposent sur l'utilisation de marqueurs génétiques, c'est-à-dire une portion de l'ADN (ou un de ses produits) de l'organisme étudié. Il est important que cette portion d'ADN (ou son produit) ait la même localisation dans le génome de chaque individu (c'est-à-dire qu'elle se trouve au même endroit sur le même chromosome), d'où le terme de locus (de Meeus *et al*, 2007). L'étude de la génétique des populations des pathogènes et de leurs vecteurs est appelée épidémiologie moléculaire (Tibayrenc, 1998). La structure génétique des populations et les systèmes d'accouplement des agents pathogènes sont des phénomènes biologiques étroitement liés, qui ont des conséquences cruciales sur l'épidémiologie des maladies transmissibles (Tibayrenc and Ayala, 2002). De nombreux marqueurs moléculaires ont été utilisés pour identifier les routes d'invasion des vecteurs, les populations sources, ou mettre en évidence des facteurs environnementaux influençant les flux de gènes entre populations (Estoup and Guillemaud, 2010; Kirk *et al*, 2013; Manel *et al*, 2003). Deux catégories de marqueurs moléculaires sont couramment utilisées : les marqueurs cytoplasmiques et les marqueurs nucléaires co-dominants.

Les marqueurs cytoplasmiques sont des loci présents dans les génomes chloroplastiques (plantes) et mitochondriaux. Largement utilisé dans les études phylogéographiques d'insectes (Allio *et al*, 2017; Bellis *et al*, 2015; Jacquet *et al*, 2015), l'ADN mitochondrial (ADNmt), évolue rapidement et ne recombine pas (Arbogast, 2001; Avise *et al*, 1987). L'ADNmt possède une héritabilité uniparentale maternelle. La taille efficace de la population de ces marqueurs est difficile à appréhender car elle dépend de la dispersion spécifique au sexe, du ratio male/femelle et de la stratégie de reproduction (de Meeus *et al*, 2007; Prugnolle and de Meeus, 2002). Enfin, l'ADNmt peut parfois être soumis à sélection (Gerber *et al*, 2001). La structure génétique identifiée à partir de l'ADNmt peut alors ne pas être uniquement le reflet de processus démographiques et géographiques. Pour ces raisons, les marqueurs mitochondriaux sont de moins en moins utilisés dans les études de génétique des populations (de Meeus *et al*, 2007). Cependant, l'utilisation de marqueurs mitochondriaux a par exemple permis de détecter la structure génétique et les flux de gènes entre populations d'*Anopheles sinensis*, vecteur d'agents du paludisme (Chang *et al*, 2016) à l'aide d'approche de « *genetic landscape shape* » (Miller, 2005). Cette méthode basée sur des matrices de distance inter-individuelle génétique et géographique permet de générer une surface tridimensionnelle où les axes X et Y correspondent aux emplacements géographiques et les hauteurs de surface (axes Z) représentent les distances génétiques.

Les marqueurs nucléaires codominants permettent de distinguer les différents génotypes homozygotes et hétérozygotes (de Meeus *et al*, 2007). Malgré la démocratisation des nouvelles technologies, basées sur les techniques de séquençage plus récentes (séquençage NGS), comme le ddRADseq (Andrews *et al*, 2016), ou le séquençage de génomes entiers (Ellegren, 2014; Fuentes-Pardo and Ruzzante, 2017) permettant l'accès à de très nombreux Single Nucleotide Polymorphism (SNP), les marqueurs microsatellites restent encore très largement utilisés du fait de leur plus faible coût et de la facilité de génotypage. De plus, les SNP possèdent des taux de mutation hétérogènes qui peuvent biaiser les conclusions (Vignal *et al*, 2002). Les microsatellites correspondent à des répétitions en tandem de motifs nucléotidiques de 2 à 5 nucléotides en général. Ces marqueurs sont co-dominants, abondants dans presque tous les génomes, à hérédité biparentale et généralement neutres, c'est-à-dire non soumis à sélection (de Meeus *et al*, 2007; Holderegger *et al*, 2006; Lehmann *et al*, 1996). Très polymorphes et possédant une fréquence de mutation rapide, les marqueurs microsatellites se révèlent très informatifs, et donc utiles en épidémiologie moléculaire (de Meeus *et al*, 2007). Ces marqueurs ont été très largement utilisés pour évaluer les flux de gènes de différentes espèces d'insectes

vectrices (Bouyer *et al*, 2007; Jacquet *et al*, 2015; Onyango *et al*, 2015a; Simo *et al*, 2010; Solano *et al*, 2010).

Les marqueurs moléculaires peuvent être utilisés individuellement ou combinés à d'autres types de marqueurs (Auger-Rozenberg and Roques, 2012; Boissin *et al*, 2016; Dube *et al*, 2017; Valade *et al*, 2009). Leur utilisation lors d'approches multi-locus permet une estimation plus fine, à différents pas de temps et à plus grande échelle de la connectivité entre les populations car ils possèdent des vitesses de mutation différentes (Brito and Edwards, 2009; de Meeus *et al*, 2007; Prugnolle and de Meeus, 2002).

L'identification de la structure génétique des populations peut être effectuée par deux approches : des approches exploratoires sans *a priori*, « *Exploratory data analysis* » (EDA), et des approches basées sur des modèles de génétiques des populations, « *Model-based clustering* » (MBC) (François and Waits, 2015).

Les méthodes « *Exploratory data analysis* » sont sans *a priori* mais permettent cependant la détection de structuration génétique et de groupes d'individus génétiquement similaires (François and Waits, 2015). Parmi les méthodes EDA on peut citer les approches de type analyses en composantes principales (ACP) (Wold *et al*, 1987) comme l'analyse factorielle de correspondance (ACF) (qui s'applique à des tableaux de contingence c'est-à-dire des tableaux croisant deux variables qualitatives et qui est donc bien adaptée aux organismes diploïdes) (She *et al*, 1987) ou encore l'analyse discriminante des composantes principales (DAPC) (AFC avec une étape supplémentaire permettant d'identifier les groupes génétiques à l'aide d'analyse discriminante (AD))(Jombart *et al*, 2010). Ces méthodes permettent de positionner des groupes d'individus dans un espace multidimensionnel. Des approches spatialement explicites sont également disponibles parmi les méthodes EDA, comme l'ACP spatiale, dont les axes principaux maximisent l'autocorrélation spatiale au lieu de la corrélation (Jombart *et al*, 2008). Enfin, des approches visuelles qui permettent de cartographier la différenciation génétique, sans *a priori* de modèle, comme l'approche « *Mapping Averaged Pairwise Information* » (MAPI) (Piry *et al*, 2016) ou EEMS (Petkova *et al*, 2016). Par exemple, MAPI permet d'assigner des indices de dissimilarité moyenne sur l'ensemble des cellules de la zone d'étude et d'identifier des discontinuités significatives dans les valeurs des cellules de la grille, grâce à une procédure de randomisation non paramétrique (Piry *et al*, 2016).

Les méthodes MBC utilisent des modèles statistiques de génétique des populations pour créer des groupes panmictiques d'individus génétiquement similaires, délimitant les frontières entre populations (Corander *et al*, 2003; Manel and Holderegger, 2013; Pritchard *et al*, 2000; Storfer *et al*, 2007). Elles permettent de calculer la probabilité de chaque individu d'appartenir à une sous-population donnée (de Meeus *et al*, 2007) en mettant en relation les fréquences des allèles qui définissent la population et les fréquences trouvées chez les individus identifiés comme étant originaires de cette population. Un individu ayant une faible probabilité d'appartenir à la population dans laquelle il a été échantillonné est susceptible d'être un immigrant récent (Manel *et al*, 2005). Parmi les outils les plus utilisés, on compte les méthodes de « *clustering* » ou regroupement en français. Dans ces méthodes, l'individu est l'unité d'analyse. Le nombre de clusters génétiques distincts est déduit d'un ensemble de génotypes multilocus individuels, en estimant l'ascendance génétique à partir de populations sources non observées, tout en maximisant l'équilibre de Hardy-Weinberg (EHD) et en limitant le déséquilibre de liaisons entre loci (Manel *et al*, 2005). Les clusters déduits peuvent ensuite être utilisés pour attribuer l'appartenance, des génotypes et des individus, à un groupe. Le logiciel STRUCTURE (Falush *et al*, 2003; Pritchard *et al*, 2000), basé sur une méthode paramétrique bayésienne, est devenu l'un des plus populaires pour les études de génétique des populations et de génétique du paysage (de Meeus *et al*, 2007). D'autres approches possèdent l'avantage d'être spatialement explicites (Francois and Durand, 2010), comme GENELAND (Guillot *et al*, 2005), BAPS (Corander *et al*, 2007; Corander *et al*, 2004) ou TESS (Chen *et al*, 2007; Francois *et al*, 2006) et font l'hypothèse que les individus géographiquement proches sont plus susceptibles d'être génétiquement semblables. Les modèles intègrent alors la localisation des individus dans le modèle. La structuration spatiale de pathogènes (Carrel *et al*, 2012; Young *et al*, 2017), de parasites (Brar *et al*, 2015; Rieux *et al*, 2011) et d'hôtes (Cote *et al*, 2012; Cullingham *et al*, 2009; Frantz *et al*, 2009) ont été mise en évidence grâce à ces méthodes, en identifiant des barrières paysagères à leur dispersion. Cependant, lorsqu'il existe un isolement par la distance significatif, ces analyses de *clustering* pourraient surestimer le nombre réel de groupes génétiques (Frantz *et al*, 2009), car la plupart des méthodes de *clustering* ne prennent pas en considération l'autocorrélation spatiale (Frantz *et al*, 2009; Meirmans, 2015).

Cette section souligne l'intérêt des approches indirectes pour quantifier la diversité et la différenciation génétique intra et inter-populationnelle par le calcul de distances génétiques et l'utilisation de méthodes de regroupements d'individus ou de populations. L'identification des

facteurs environnementaux favorisant ou freinant les flux de gènes d'insectes vecteurs et donc limitant la propagation d'agents pathogènes, présente une importance capitale dans la compréhension de l'épidémiologie de maladies vectorielles (Archie *et al*, 2009; Meentemeyer *et al*, 2012).

I.3) Evaluation et quantification de l'impact des facteurs environnementaux sur les flux de gènes d'insectes vecteurs

La génétique du paysage emprunte de nombreux concepts et méthodes à d'autres disciplines (Manel *et al*, 2003; Storfer *et al*, 2007) et peut se définir comme une discipline combinant génétique des populations, écologie du paysage et techniques d'analyses spatiales. Elle permet de quantifier explicitement les effets de la composition, de la configuration et de la qualité de la matrice du paysage sur les processus microévolutifs, tels que les flux de gènes, la dérive génétique et de la sélection, en utilisant des données génétiques neutres ou adaptatives (Balkenhol and Fortin, 2015; Holderegger and Wagner, 2008). Bien que la manière la plus directe pour identifier l'impact du paysage sur la circulation de pathogènes soit de travailler sur les agents eux même (Encadré 3), la collecte en masse d'échantillons de pathogènes peut s'avérer délicate surtout lorsque leur prévalence est faible. Celle-ci nécessite souvent un travail de laboratoire très important (conservation, diagnostic, mise en culture) et de prélèvements sur le terrain sur les animaux sauvages. De plus, l'agent pathogène peut, dans certain cas, présenter une faible variation génétique. Ainsi, l'étude de la génétique des vecteurs ou des hôtes peut être utilisée pour estimer la propagation et la dynamique d'agents infectieux et se révéler très informative (Biek and Real, 2010; Kozakiewicz *et al*, 2018). Par exemple, au Kenya, des auteurs ont constaté que la différenciation génétique mitochondriale entre populations de *Aedes mcintoshi*, vecteur du virus de la fièvre de la vallée du Rift, était corrélée avec les valeurs moyennes des précipitations (Campbell and Alexander, 2017). La génétique du paysage, appliquée aux insectes vecteurs, reste encore un champ thématique peu développé, bien que certaines études commencent à émerger depuis quelques années (Bouyer *et al*, 2015; Campbell and Alexander, 2017; Chang *et al*, 2016; Hlaing *et al*, 2010).

D'un point de vue méthodologique, on peut diviser une étude de génétique du paysage en trois étapes, les deux premières pouvant être réalisées de manière simultanée. La première étape est la quantification de l'hétérogénéité du paysage afin d'appréhender de façon numérique la configuration, la composition et la qualité de la matrice pour en déduire la connectivité paysagère (Encadré 1, Encadré 3). La seconde étape est la détection spatiale de la diversité, la

structure et la différenciation génétique entre les populations par l'utilisation de marqueurs génétiques polymorphes comme nous l'avons vu précédemment. Enfin, la dernière étape consiste à trouver des relations statistiques quantifiables entre les variations génétiques et l'hétérogénéité du paysage.

Encadré 3 : La génétique du paysage appliquée aux pathogènes, aux hôtes et réservoirs

La manière la plus directe pour identifier l'impact du paysage sur la circulation de pathogènes est de travailler sur les agents eux-mêmes (Archie *et al*, 2009). L'évolution rapide des génomes viraux, combinée à la fixation rapide de ces mutations, en font de bons modèles pour étudier les processus épidémiologiques qui façonnent leur diversité. De nombreuses études se sont penchées sur l'interaction entre l'émergence de maladies infectieuses et les composantes environnementales qui peuvent y être associées (Lemey *et al*, 2010; Ostfeld *et al*, 2005). Par exemple, d'importants agents pathogènes des ongulés sauvages et domestiques ont fait l'objet d'études par l'utilisation couplée de distance phylogénétique et de résistance paysagère (Biek, 2007). Ainsi, une étude de phylogéographie du paysage par modèle linéaire généralisé a permis d'identifier les facteurs environnementaux qui ont favorisé la propagation du virus de la FCO en Europe (Jacquot *et al*, 2017). Cette étude était basée sur la diversité de 113 génomes du virus de la FCO, géoréférencés et horodatés, représentant différents sérotypes et souches du virus, provenant de pays touchés et identifiés comme source d'introduction en Europe. La densité d'hôtes bovins a été identifiée comme favorisant la propagation du virus de la FCO. Au contraire, les masses d'eau (mer, océan) et les hautes altitudes jouent le rôle de barrière dans la diffusion du virus.

Récemment, l'utilisation d'un *workflow* innovant d'analyse phylogéographique a permis de quantifier la résistance du paysage le long des voies de transmission d'un virus touchant les rats laveurs en Amérique du nord. Cette approche nécessite de reconstituer la phylogénie des lignées du virus afin d'extraire des informations spatio-temporelles, pour estimer la dispersion et définir des statistiques épidémiologiques de propagation. La résistance paysagère est ensuite quantifiée grâce au modèle classique de mouvement spatial (least-cost transect analysis).

Enfin, des analyses de corrélation par régression multiple de matrices de distance (MRDM) permettent de tester l'impact des variables environnementales sur le mode et le rythme de dispersion des agents pathogènes lors des épidémies (Dellicour *et al*, 2016). Les approches MRDM permettent de tester la corrélation entre une matrice de distance basée sur des arbres

phylogénétiques de maximum de vraisemblances et une matrice de résistance ou de conductance paysagère (Carrel *et al*, 2012; Kozakiewicz *et al*, 2018; Young *et al*, 2017).

L'épidémiologie du paysage, en s'intéressant à l'effet de l'hétérogénéité spatiale sur la circulation des pathogènes, mais aussi de leurs vecteurs ou de leurs hôtes, permet une meilleure compréhension de la transmission des maladies infectieuses (Biek and Real, 2010; Meentemeyer *et al*, 2012). L'étude de DeYoung *et al*. (2009) a, par exemple, mis en évidence chez des populations sauvages de renards gris au Texas, des flux de gènes à longue distance permettant de mieux comprendre la circulation du virus de la rage dans cette région et ainsi proposer un plan de vaccination des renards (DeYoung *et al*, 2009; Hajibabaei *et al*, 2007). Des études de génétique du paysage ont également mis en évidence que les barrières fluviales et autoroutières pouvaient parfois limiter la dispersion des femelles mouffettes rayées, porteuses du virus de la rage (Talbot *et al*, 2012). Identifier les voies de propagation de vecteurs et pathogènes, permet ainsi de reconstituer l'histoire d'épidémies récentes (Corman *et al*, 2014; Faria *et al*, 2014; Fountain-Jones *et al*, 2017; Fountain-Jones *et al*, 2018; Magee *et al*, 2015).

Deux approches sont classiquement utilisées pour quantifier les relations statistiques entre distance génétique et résistance paysagère. La première consiste à utiliser la configuration de la structure génétique obtenue, que ce soit par méthodes EDA ou MBC, (Guillot *et al*, 2005; Jombart *et al*, 2008; Pritchard *et al*, 2000) et de la comparer avec la configuration du facteur environnemental étudié (Frantz *et al*, 2012; Prunier *et al*, 2013a; Prunier *et al*, 2013b), pour tenter d'identifier de potentielles barrières ou corridors de dispersion. Cette approche a l'avantage d'être relativement simple et rapide, mais ne permet pas de quantifier les relations statistiques entre dissimilarité génétique et résistance paysagère, car elle n'utilise pas directement de données environnementales.

La seconde approche par « *gradient direct* » (Balkenhol and Fortin, 2015; Prunier *et al*, 2015) quantifie la contribution relative de l'ensemble des facteurs environnementaux testés sur la variance du jeu de données génétiques (Balkenhol *et al*, 2009; Guillot *et al*, 2009). Compte tenu de l'absence de consensus actuel sur la meilleure manière d'opérer (Zeller *et al*, 2012), de nombreuses procédures ont été développées pour quantifier la perméabilité paysagère. L'une des plus simples et plus anciennes est le test de Mantel qui teste l'absence de relation (H0) entre deux matrices de dissimilarité (ou de distance). Le test d'isolement par la distance (IBD) utilise le test de Mantel pour étudier la corrélation entre une matrice de distance géographique et une

matrice de distance génétique (Mantel, 1967; Shirk *et al*, 2010). La composante paysagère est donc résumée à la distance géographique entre les populations. Le test de Mantel souffre de nombreuses critiques (Diniz-Filho *et al*, 2013). En effet, les deux hypothèses de base de ce test, l'homoscédasticité et la linéarité, sont rarement respectées dans les analyses spatiales. De plus, les variables utilisées dans ce test (génétiques et environnementales) ne sont pas indépendantes ce qui augmente la probabilité de rejeter l'hypothèse nulle alors que cette dernière est vraie (erreur de type I) (Balkenhol *et al*, 2009; Graves *et al*, 2013; Landguth *et al*, 2010). En raison de son manque de puissance statistique, ce test est aujourd'hui moins utilisé (Manel and Holderegger, 2013).

En Asie du sud-est continentale, treize marqueurs microsatellites ont été utilisés pour identifier la structure génétique de population d'*Aedes aegypti* et tenter de déterminer les facteurs environnementaux sous-jacents (Hlaing *et al*, 2010). Ainsi, les grandes voies de transport humains ont été identifiées comme facteur permettant la similarité génétique entre sites éloignés géographiquement à l'aide de la méthode bayésienne hiérarchique implémentée dans le logiciel GESTE (GENetic STRucture inference based on genetic and Environmental data) (Foll and Gaggiotti, 2006). Cette méthode, après avoir estimée les valeurs de F_{ST} populationnelle, utilise un modèle linéaire généralisé pour mettre en relation les facteurs environnementaux et la différenciation génétique (Foll and Gaggiotti, 2006). Cependant, cette approche utilise uniquement des valeurs environnementales locales et néglige la résistance paysagère qui les sépare.

Encadré 4 : Approche populationnelle et individuelle

Le plan d'échantillonnage et le protocole d'analyse dans une étude de génétique du paysage sont guidés par la problématique à laquelle nous voulons répondre (Wagner and Fortin, 2015). L'approche classique d'échantillonnage en génétique du paysage et plus généralement en génétique des populations, est l'approche populationnelle. Roughgarden définit en 1979 une population comme « *un ensemble d'individus dont on pense qu'il fonctionne comme une unité du point de vue évolutif et écologique* ». L'approche populationnelle a donc pour objectif d'échantillonner un nombre important d'individus à un même point géographique afin d'avoir une représentation fidèle de la diversité génétique locale et donc des processus microévolutifs qui l'ont modelée. Cette approche est bien adaptée aux espèces distribuées en patch ou en métapopulations. Cependant, elle possède des limites. En effet, ce type d'approche restreint le

nombre de sites d'échantillonnage, la couverture spatiale de l'étude, et la diversité de paysage représentée, des composantes qui impactent fortement la robustesse des analyses de génétique du paysage. Le groupement d'individus au sein d'une population implique une décision a priori contestable et subjective, car il est difficile de déceler si la population étudiée peut-être ou non considérée comme une seule unité panmictique. Si cette population est subdivisée, un déficit d'hétérozygotes par rapport à la structure de Hardy-Weinberg va être observé. Cet effet est appelé effet Wahlund.

Au contraire, l'approche individuelle consiste à échantillonner un plus faible nombre d'individus dans un grand nombre de zones géographiques. Contrairement à l'approche populationnelle, la diversité génétique a plus de risques de ne pas être totalement représentative de la diversité génétique locale. En revanche, elle permet de couvrir un éventail plus grand de typologies de paysages, d'avoir une couverture spatiale plus importante, et ainsi, d'identifier de manière plus précise géographiquement, des variations de structure ou de flux génétiques (Luximon *et al*, 2014; Prunier *et al*, 2013a). Cette approche est adaptée aux organismes vivants ne possédant pas un mode de vie grégaire, mais étant plutôt répartis de manière uniforme dans le paysage (Richardson *et al*, 2016).

Des méthodes par gradient direct ont été développées plus récemment, comme les régressions multiples sur matrices de distances (MRDM) qui permettent de prendre en compte simultanément plusieurs vecteurs de distance explicatifs (Smouse *et al*, 1986) et des modèles de maximum de vraisemblance, des effets de population (MLPE) (Clarke *et al*, 2002; Van Strien *et al*, 2012) permettant de modéliser explicitement la dépendance entre des observations par paires. Dans ce type d'analyses, les variables testées sont introduites sous forme de surface de résistance paysagère ou de matrices des distances basées sur les modèles « *least-cost path* » (Adriaensen *et al*, 2003), « *circuit-based* » (McRae and Shah, 2008; McRae and Beier, 2007) et « *transect-based approaches* » (Angelone *et al*, 2011; Emaresi *et al*, 2009)(Encadré 1). Ces surfaces indiquent comment chaque cellule de la grille paysagère limite ou favorise le flux de gènes (McRae, 2006; Spear *et al*, 2010; Spear *et al*, 2015b; Storfer *et al*, 2010). Par exemple, Bouyer et ses collaborateurs, utilisent des cartes de friction environnementales pour étudier la dispersion de *Glossina palpalis gambiensis*, vecteur de trypanosomoses en Afrique de l'ouest. Une carte de friction est une représentation graphique qui cartographie la résistance paysagère et qui permet d'en déduire le chemin optimal minimisant la résistance paysagère entre deux zones géographiques. Cette approche a permis d'identifier les barrières environnementales à la

dispersion de ce vecteur d'intérêt (Bouyer *et al*, 2015). Les cartes de friction restent cependant à ce jour peu appliquées aux espèces d'insectes (Zeller *et al*, 2012).

Les tests basés sur les MRDM font l'objet de critiques similaires à celles faites au test de Mantel, notamment à cause de la multicolinéarité entre les prédicteurs (Manel and Holderegger, 2013; Prunier *et al*, 2015; Wagner and Fortin, 2015). Dans le cas des MRDM (Balkenhol *et al*, 2014; Lichstein, 2006; Wang, 2013), les différentes matrices de distances environnementales de résistance ou perméabilité paysagère sont, en général, toutes dépendantes de la distance géographique ; les analyses multiples deviennent ainsi redondantes et non indépendantes (Cayuela *et al*, 2018). L'analyse de commonalité permet l'identification de la part commune et unique des différents facteurs environnementaux. La part unique (U) quantifie la proportion de la variance, expliquée de manière unique par une variable explicative. Si cette valeur est négligeable, l'ajout du prédicteur n'améliore pas ou peu le modèle. Au contraire, la part commune (C) représente la proportion de la variance, expliquée conjointement par plusieurs variables explicatives. L'analyse de commonalité corrige alors la multicollinéarité inhérente aux approches corrélatives (Prunier *et al*, 2017; Prunier *et al*, 2015). Une approche complémentaire récente permet également, pour se défaire de l'omniprésence de la composante liée à la distance géographique, d'utiliser d'un raster nul, qui n'est autre qu'une grille où l'ensemble des pixels possèdent la même valeur de résistance. Seule alors compte la distance géographique dans le calcul du chemin le plus court entre deux populations. Ainsi, il est possible de détecter les facteurs environnementaux dont la résistance ou la perméabilité paysagère explique mieux la dissimilarité génétique que la distance géographique seule (Dellicour *et al*, 2019).

Cette section souligne l'importance d'une connaissance fine de l'influence des facteurs paysagers sur les flux de gènes des agents pathogènes, de leurs hôtes et de leurs vecteurs. Les outils de génétique des populations et du paysage permettent ainsi de mieux appréhender la manière dont les maladies se transmettent et d'élaborer des stratégies de lutte ou de vaccination plus adaptées (Biek and Real, 2010; Manel and Holderegger, 2013; Segelbacher *et al*, 2010). Cependant, si la structure génétique observée résulte de processus microévolutifs récents, elle est également le témoin de phénomènes évolutifs plus anciens.

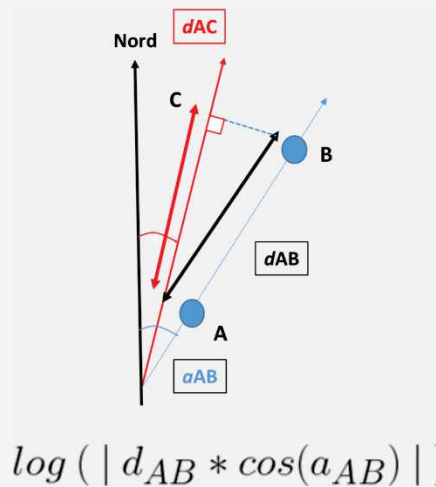
Encadré 5 : L'isolement par la distance anisotropique

La théorie de l'isolement par distance (IBD), introduite par Wright (1943), décrit l'accumulation de la variation génétique locale sous l'hypothèse d'une dispersion spatiale limitée (Slatkin, 1993). Selon cette théorie, les mesures de différenciation génétique par paires devraient augmenter avec la distance géographique. Cependant, l'isolement par la distance classique n'intègre pas de composante directionnelle. En effet, le modèle d'isolement par la distance suppose une homogénéité des variables paysagères ou environnementales, ce qui n'est pas le cas pour la plupart des systèmes naturels (Duvallat *et al*, 2017; Wagner and Fortin, 2005). Or, de nombreux phénomènes anisotropes, c'est-à-dire dont les propriétés varient selon la direction considérée, permettent la dispersion passive d'individus, comme par exemple le vent ou les courants (Bertin *et al*, 2015). Ainsi, en intégrant une dimension spatiale à l'analyse des distances génétiques il est possible d'identifier l'orientation la plus propice au flux de gènes, c'est-à-dire celle pour laquelle l'accumulation de la différenciation génétique est la plus importante (Rosenberg, 2010). Développées à l'origine dans le cadre d'études de génétique des populations humaines (Amit Roy, 2014; Harrup *et al*, 2016; Sokal *et al*, 1993), ces techniques ont été principalement utilisées pour d'étudier la dispersion de pollens (Dutech *et al*, 2005) ou de spores fongiques (Rieux *et al*, 2014). L'utilisation de ces approches sur des espèces d'insectes de petite taille pouvant être dispersées passivement par le vent, pouvant permettre d'identifier les grandes trajectoires de dispersion.

Ainsi, des méthodes permettant d'identifier des flux de gènes directionnels ont émergé, tels que l'isolement par la distance anisotropique (Jay *et al*, 2013), les « *bearing correlogram* » (Rosenberg, 2010), les semi-variogrammes directionnels (Wagner and Fortin, 2005) ou les « *bearing analysis* », disponibles dans le logiciel Pattern Analysis, Spatial Statistics and Geographic Exegesis (PASSaGE). Ces approches permettent d'identifier l'angle entre populations qui maximise et minimise la corrélation entre matrices de distances génétiques et géographiques (Jay *et al*, 2013). Ce type d'analyse nécessite de calculer la distance projetée entre deux sites d'échantillonnage.

Dans le calcul de la distance géographique projetée en fonction de l'angle entre les populations, d_{AB} correspond à la distance géographique entre les populations A et B, et α_{AB} correspond à l'angle entre les populations A et B. La matrice de distance projetée est par la suite calculée pour chaque angle entre tous les sites d'échantillonnage, en utilisant la formule ci-dessus. Puis, une régression linéaire entre les distances génétiques et les distances géographiques projetées

obtenues pour chaque angle est effectuée. L'angle qui maximisera le R^2 de cette régression avec un coefficient de régression positif est considéré comme l'angle maximisant le signal d'isolement par distance.



Représentation schématique du calcul de la distance géographique projetée en fonction d'un angle

I.4) Des flux de gènes contemporains à la création de nouvelles espèces

La fragmentation de populations en petits isolats entraîne une diminution de la taille de la population et à une dérive génétique plus rapide avec pour conséquence, une différenciation génétique pouvant aboutir à la création de lignées génétiques différentes. Si l'isolement géographique et l'absence de flux de gènes sont prolongés, la sélection et la dérive peuvent être à l'origine d'événements de spéciation. La sélection naturelle augmente la fréquence de variantes génétiques avantageuses dans une population, donnant naissance à des organismes avec une plus forte valeur sélective, ce qui peut conduire à de la spéciation. La spéciation résulte d'un processus évolutif aboutissant à l'accumulation de mutations dans une population sur une longue période de temps, menant possiblement à une nouvelle espèce. La création d'entités génétiques distinctes est conditionnée par l'apparition d'une incompatibilité génétique, comportementale ou écologique constituant un isolement reproducteur. Ainsi, même en cas de contact physique entre les lignées, les flux de gènes deviennent impossibles. Une espèce est un groupe d'organismes qui peuvent se reproduire les uns avec les autres pour produire une descendance fertile et qui sont isolés des autres organismes sur le plan de la reproduction. Des processus de spéciation en sympatrie, définis comme des événements de spéciation sans l'intervention de barrières aux flux de gènes, interviennent chez certaines espèces d'insectes.

Cependant, le mode de spéciation le plus répandu chez les espèces d'insectes généralistes intervient lorsque des populations biologiques de la même espèce sont isolées les unes des autres par une barrière géographique, au point que cela empêche ou interfère avec le flux de gènes (Bush and Butlin, 2004). Il s'agit là de spéciation allopatrique, dont le plus illustre exemple est la spéciation des pinsons de Darwin sur les îles des Galápagos.

L'espèce est une unité fondamentale de la biologie, permettant d'organiser, classifier et grouper la diversité du vivant (De Queiroz, 2007; Mayr, 1982). Cependant, le concept d'espèce ne répond pas toujours à une définition unifiée et partagée (De Queiroz, 2007; Mayden, 1999). De l'inflation taxonomique se produit lorsque les chercheurs élèvent les sous-espèces ou les variantes géographiques au niveau des espèces en appliquant des concepts d'espèces alternatives. Cette subjectivité nuit à la stabilité taxonomique. De plus, les espèces ne sont pas des entités évolutives figées et permanentes. Au contraire, ce sont des groupes de populations en constante évolution. Pour éviter ces écueils, des méthodes de délimitation et de description des espèces sont nécessaires. Un large éventail de nouvelles méthodes de délimitation des espèces a été récemment développé, et ces dernières sont de plus en plus utilisées en taxonomie (Flot, 2015). Les méthodes de délimitation d'espèces servent à identifier les entités appartenant à la même espèce. Ces méthodes permettent d'avoir l'assurance que tous les individus utilisés dans une étude appartiennent à la même unité taxonomique, évitant de biaiser les résultats (Bortolus, 2008).

De nombreuses méthodes de délimitation d'espèces ont été mises au point, mais les plus utilisées sont basées sur l'utilisation de dendrogrammes (Fontaneto *et al*, 2015). Les méthodes basées sur le modèle généralisé de coalescence mixte de Yule (GMYC) (Pons *et al*, 2006) et sur le « *Poisson tree processes* » (PTP) (Zhang *et al*, 2013) identifient les points de transition entre diversité interspécifique et intraspécifique à l'aide d'arbres phylogénétiques (Fujisawa and Barraclough, 2013). Les deux méthodes se distinguent par l'estimation du taux de spéciation (Dellicour and Flot, 2018). La méthode GMYC utilise un arbre phylogénétique ultra-métrique et considère la longueur des branches de l'arbre comme proportionnelle au temps de divergence. Ainsi, cette approche vise à rechercher dans cet arbre ultra-métrique, le seuil auquel les divergences entre taxons représentent des événements de spéciation ou de coalescence (Pons *et al*, 2006). En revanche, la méthode PTP considère la longueur des branches proportionnelle au nombre de substitutions (Zhang *et al*, 2013), sans nécessité d'arbres ultra-métriques, rendant la méthode plus rapide et tout aussi efficace selon des études de modélisation (Dellicour and Flot, 2018). De plus, la méthode GMYC peut parfois surestimer le nombre d'espèces suite à un

biais de certains paramètres de l'analyse, comme les différences dans la taille de la population ou les taux de spéciation (Dellicour and Flot, 2015; Esselstyn *et al*, 2012; Fujisawa and Barraclough, 2013; Talavera *et al*, 2013). Des implémentations bayésiennes de la méthode GMYC, (bGMYC) (Reid and Carstens, 2012) et de la méthode PTP (bPTP) ont également été développées.

Depuis de nombreuses années, l'ADNmt est utilisé comme marqueur moléculaire de choix pour étudier les unités évolutives correspondant à l'espèce, en particulier chez les espèces difficilement différenciables morphologiquement. L'ADNmt est un outil de taxonomie moléculaire précieux grâce à l'approche de « *barcoding* » moléculaire, qui permet d'identifier des unités taxonomiques opérationnelles moléculaires (MOTU). Le pourcentage de divergence nucléotidique mitochondriale permet d'identifier la frontière entre la divergence intra-spécifique et inter-spécifique. Une approche de délimitation d'espèces est basée sur ce principe de *barcoding* moléculaire de l'ADNmt, nommé l'ABGD (*automatic barcode gap discovery*) (Puillandre *et al*, 2012). L'approche ABGD (Puillandre *et al*, 2012) repose sur le « *barcoding gap* » c'est-à-dire l'écart qui sépare les distances génétiques inter- et intra-spécifiques déduites de l'ADNmt. Cette délimitation est particulièrement adaptée pour des jeux de données possédant une richesse spécifique importante (Puillandre *et al*, 2012).

La plupart de ces approches sont unilocus, principalement réalisées avec le marqueur moléculaire *cox1*. Cependant, nous avons vu précédemment que l'ADNmt possède une héritabilité uniparentale maternelle. La taille efficace de population des marqueurs mitochondriaux est donc inférieure à celle des loci du génome nucléaire. Par conséquent, la perte de diversité génétique, via la dérive génétique, aura tendance à être plus rapide pour les loci mitochondriaux, créant ainsi des haplotypes mitochondriaux différents (de Meeus *et al*, 2007; Prugnolle and de Meeus, 2002). La diversité génétique nucléaire aura normalement tendance à être mieux conservée. Ainsi, si une forte différenciation mitochondriale est détectée entre des lignées, il est important sans preuve morphologique ou écologique de différenciation ou d'isolement reproducteur, d'intégrer des marqueurs nucléaires à l'analyse. Il est donc conseillé d'utiliser pour la délimitation d'espèces, une approche multi-locus comme le Bayesian Poisson Tree Processes (bPTP), qui permet d'améliorer la résolution de l'analyse (Bakhoun *et al*, 2018; Fujita *et al*, 2012; Leache and Fujita, 2010; Pentinsaari *et al*, 2017). De plus, la diversité des algorithmes, des paramètres d'analyses et des implémentations disponibles pour effectuer des délimitations d'espèces, peuvent être source de résultats différents (Dellicour and Flot, 2015). Il n'existe donc pas de solution optimale, mais l'utilisation combinée de plusieurs

méthodes différentes, appliquées sur le même modèle biologique, est conseillée. Si un patron commun de délimitation émerge, l'hypothèse de délimitation devrait être correcte (Dellicour and Flot, 2018).

L'ADN mitochondrial et nucléaire peuvent également être utilisés pour détecter des expansions démographiques, des divergences entre lignées et des phénomènes d'admixture entre populations. La phylogéographie désigne justement l'étude des principes et processus qui gouvernent la distribution des lignées au sein des espèces. Elle étudie les phénomènes génétiques et démographiques, en particulier les phénomènes de spéciation, ayant conduit à la distribution et à la structuration géographique actuelle des populations (Arbogast, 2001). L'utilisation combinée de ces deux types de loci qui, de par leur héritabilité, réagissent différemment aux fluctuations démographiques, permet d'identifier des scénarii évolutifs de fragmentation des populations parfois complexes. La fragmentation du paysage, causée par les épisodes glaciaires par exemple, a été de nombreuses fois mise en évidence comme la cause, chez de nombreuses espèces européennes, de divergences entre lignées génétiques actuelles (Hewitt, 1999; Hewitt, 2004; Provan and Bennett, 2008). Durant ces épisodes glaciaires, l'ADN nucléaire aura tendance à moins diverger que l'ADNmt. L'étude d'espèces phylogénétiquement proches ayant récemment divergé peut permettre de mieux appréhender la différenciation génétique contemporaine. C'est le cas pour certaines espèces de termites dont l'étude phylogéographique révèle de multiples événements de colonisation post-glaciaires en Europe (Lefebvre *et al*, 2016). Lorsque des espèces ayant divergé récemment en allopatrie, présentent un isolement reproductif incomplet, l'apparition d'hybrides est encore possible lors du contact secondaire entre les deux espèces. Lors des recolonisations post-glaciaires en Europe, depuis les principaux refuges glaciaires que sont les péninsules ibérique, italienne et balkanique, des lignées divergentes sont entrées de nouveau en contact dans les Alpes et les Pyrénées, créant ainsi des zones d'hybridation (Hewitt, 2001). Lors de recolonisation spatiale, la taille de la population de front étant réduite, la diversité est sous échantillonnée par rapport à la population originale, c'est le goulot d'étranglement génétique (Nei *et al*, 1975). Cette perte de diversité génétique est accentuée par la dérive génétique générée par l'effet fondateur (Excoffier *et al*, 2009; Sexton *et al*, 2009). La différence de diversité génétique peut permettre ainsi d'identifier des routes d'expansion géographique en identifiant le front de propagation et la population source.

Lors de phases d'expansion géographique, comme c'est le cas lors de recolonisation post-glaciaire, la dispersion joue un rôle capital (Richardson *et al*, 2000). Cependant, si les individus

parviennent à franchir la barrière géographique, qui correspond à la phase de transport, il reste toujours la barrière reproductive et de survie, correspondant à l'établissement des populations dans le nouvel environnement (Garcia-Diaz, 2016; Richardson *et al*, 2000). La plasticité phénotypique (Richards *et al*, 2006) et la capacité d'adaptation de l'espèce aux nouvelles conditions biotiques et abiotiques conditionnent le succès de cette étape. Enfin, si toutes ces barrières sont franchies avec succès et que la population nouvellement établie est en expansion, la population peut devenir envahissante. La phylogéographie a permis d'identifier des routes d'invasions d'insectes vecteurs responsables d'importants dommages économiques et de santé publique (Aukema *et al*, 2011; Juliano and Lounibos, 2005; Kenis and Branco, 2010; Kenis *et al*, 2017). Retracer les routes d'invasions et déterminer l'origine des populations est capital pour prévenir de nouvelles introductions et établir une gestion appropriée (Estoup and Guillemaud, 2010). Par exemple, une étude de phylogéographie utilisant, entre autres, des approches d'Approximate Bayesian computation (ABC) a permis d'invalider l'hypothèse classiquement admise que *Culicoides imicola*, principal vecteur de la FCO en zone afrotropicale, avait étendu récemment son aire de répartition au bassin méditerranéen, en lien avec les changements climatiques actuels. La nouvelle hypothèse proposée soutient une présence ancienne de *C. imicola* dans le bassin méditerranéen sur sa côte africaine et européenne. Cette extension de son aire de répartition serait survenue suite aux changements climatiques survenus au cours du Pléistocène et/ou de l'Holocène précoce (Jacquet *et al*, 2015).

L'*Approximate Bayesian computation* (ABC) est une approche de statistiques bayésiennes qui peut être appliquée à l'analyse de données génétiques et qui permet d'estimer la vraisemblance de modèles retraçant l'histoire évolutive et démographique de populations (Beaumont, 2010; Beaumont *et al*, 2002; Bertorelle *et al*, 2010; Csillery *et al*, 2010; Lopes and Beaumont, 2010). Les statistiques bayésiennes permettent, en effet, de déterminer la distribution *a posteriori* d'un modèle ou du paramètre θ en connaissant les données observées D en suivant l'équation suivante : $P(\theta|D) \propto P(D|\theta)P(\theta)$ où $P(\theta)$ est la distribution *a priori* du modèle ou du paramètre. La probabilité $P(D|\theta)$ d'observer les données D étant donné le paramètre (ou le modèle) θ correspond à la vraisemblance de θ . Cette vraisemblance étant très difficile à estimer pour des modèles complexes, il est possible de calculer une distribution *a posteriori* en remplaçant la vraisemblance par une approximation en utilisant les statistiques, résumées sur des jeux de données simulés selon différents modèles.

Tout d'abord, cette approche nécessite de créer des scénarios évolutifs cohérents et d'estimer les valeurs et les distributions *a priori* (priors) des paramètres historiques, démographiques et

génétiques du modèle biologique et des marqueurs moléculaires associés comme les taux de mutations des marqueurs, les tailles efficaces des populations et les dates de divergences. Des jeux de données génétiques, correspondant à chacun des scénarii décrits précédemment, sont alors simulés en très grands nombres à l'aide de modèles stochastiques. Les paramètres servant à alimenter ces simulations sont tirés aléatoirement dans la distribution des *priors*. Ces jeux de données sont ensuite résumés à une série de statistiques intra ou inter-populationnelles préalablement choisies, comme l'hétérozygotie observée et attendue ou les F_{ST} entre populations. Ces statistiques, calculées sur les jeux de données simulés, sont ensuite comparées à celles calculées sur le jeu de données observé. Les simulations les plus éloignées des observations sont écartées (Beaumont *et al*, 2002). Une régression logistique sur les jeux de données simulés conservés, permet de calculer la probabilité, *a posteriori* de chaque scénario. Enfin, le calcul des erreurs de type I (faux positif) et les erreurs de type II (faux négatif) concernant le choix du scénario optimal est réalisé pour vérifier la puissance de l'analyse effectuée. Enfin, il convient de déterminer si le scénario sélectionné et les distributions *a posteriori* des paramètres inférés reproduisent convenablement les données observées. Ceci se fait en contrôlant la concordance entre le scénario sélectionné et les données observées, en simulant des données à partir de valeurs de paramètres tirées dans les distributions *a posteriori* (« *model checking* »). Les méthodes ABC sont donc des outils précieux en biologie évolutive qui permettent d'explorer les histoires démographiques, parfois complexes, des populations naturelles et qui permettent d'estimer des paramètres démographiques sous un modèle d'évolution donné (Alvarado-Serrano *et al*, 2015; Beaumont, 2010; Rougemont and Bernatchez, 2018).

L'utilisation couplée de méthodes de phylogénie, de délimitation d'espèces et de phylogéographie par des approches multilocus apparaissent ainsi comme des outils pertinents pour établir l'histoire évolutive des arthropodes vecteurs, parfois difficilement accessibles de par leurs capacités de dispersion passive importantes.

II) *Culicoides obsoletus*, un modèle d'étude de la dispersion d'espèces vectrices de virus d'intérêt vétérinaire

II.1) Généralités sur la biologie et l'écologie des *Culicoides*

Les *Culicoides* (Diptera : Ceratopogonidae) sont de petits moucheron nématocères de 1 à 3 mm de long, associés à une grande diversité d'hôtes vertébrés dont les ruminants sauvages (Lorca-Oro *et al*, 2014) et domestiques (Ayllon *et al*, 2014). Le régime alimentaire des mâles et des femelles est floricole (Goetghebuer, 1952; Zimmer *et al*, 2014a). Seules les femelles sont hémaphages, les repas de sang permettant d'assurer la maturation des œufs (Barber *et al*, 2018).

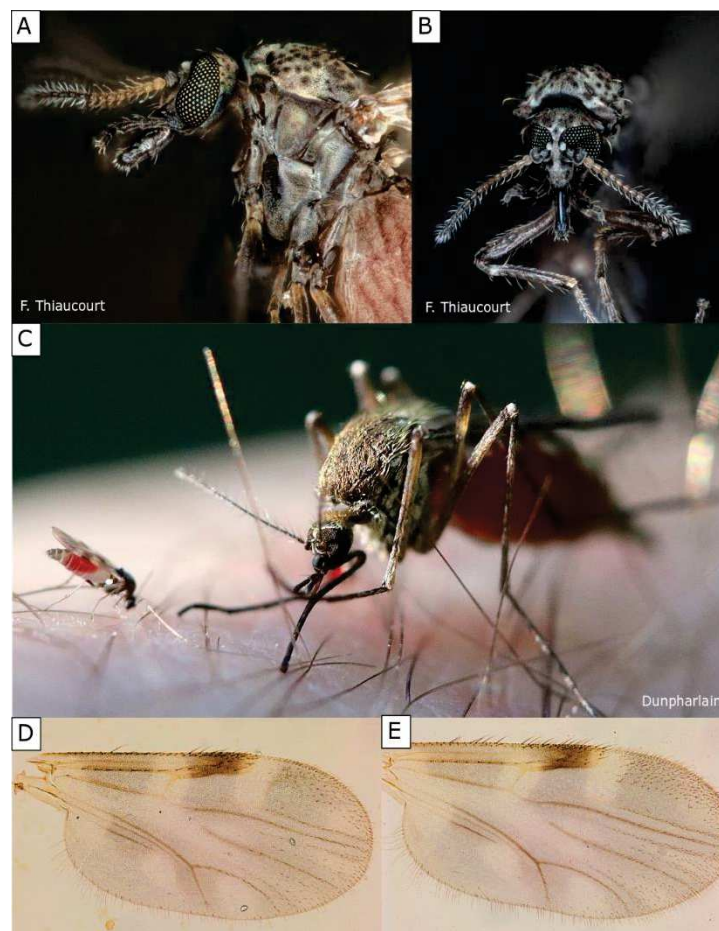


Figure 1. Photos de *Culicoides*. (A et B) *Culicoides nubeculosus*, (C) Taille d'un *Culicoides* (*Culicoides impunctatus* à gauche) par rapport à un moustique (à droite) pendant le gorgement, (D) Patron alaire de *Culicoides obsoletus*, (E) Patron alaire de *Culicoides scoticus*.

Les *Culicoides* sont présents dans quasiment toutes les régions du monde à l'exception des régions polaires et d'îles isolées des masses continentales (Borkent, 2016). Les *Culicoides* sont présents sur un large gradient d'altitude, et certains spécimens ont été capturés jusqu'à 4 200 m au Tibet et dans les Andes (Garros and Balenghien, 2017). Cette large distribution géographique et altitudinale montre la plasticité biologique et écologique de ce genre (Werner *et al*, 2020).

Le cycle de vie comprend sept stades, les œufs, 4 stades larvaires, le stade nymphal et le stade adulte (Figure 2). La durée de chaque stade et des phases inter-stades est dépendante des conditions environnementales. Dans des conditions optimales, notamment de température, le cycle de vie entier peut être réalisé en moins d'un mois, mais il peut aussi atteindre plusieurs mois pour certaines espèces tempérées exposées à des basses températures (Purse *et al*, 2015). D'aspects fusiformes, les œufs des *Culicoides* sont pondus au sol. Les habitats larvaires sont généralement décrits comme des zones humides voire semi-aquatiques, riches en matière organique d'origine animale ou végétale : fumier, litières de forêt, fèces, boue, abords de point d'eau, mares, étangs, rivières, plages, eaux stagnantes, flaques, abords d'abreuvoirs, fruits en décomposition (Bakhoun *et al*, 2016; Carpenter *et al*, 2009; Garros and Balenghien, 2017; Gonzalez *et al*, 2013; Jenkins and Young, 2010). Les quatre stades larvaires de *Culicoides* sont vermiformes, apneustiques, mobiles (Purse *et al*, 2015) et se nourrissent de nématodes, bactéries et de matière organique (Borkent, 2014). Le stade nymphal est très peu mobile et reste à la surface de l'habitat larvaire. Les stades immatures sont très sensibles à la dessiccation et des conditions défavorables ont des conséquences fortes pour la survie des populations. Après l'émergence, pour la plupart des espèces, l'accouplement s'effectue après un vol nuptial, sur un substrat ou en essaim (Duvallat *et al*, 2017; Viennet *et al*, 2013). Les femelles stockent les spermatophores dans leurs spermathèques (nombre variant entre 1 à 3), leur permettant d'engendrer une descendance tout au long de leur vie sans avoir à effectuer de nouveau un accouplement. Après l'accouplement et avant chaque ponte, la femelle se met en quête d'un hôte vertébré pour effectuer un repas de sang. Ce repas de sang permettra d'apporter l'azote et les protéines nécessaires au développement et à la maturation des œufs (entre 10 et 100 par ponte). La durée du cycle gonotrophique, représentant l'intervalle de temps entre deux repas de sang, est dépendante des conditions environnementales. Pour la plupart des espèces, les femelles sont crépusculaires avec un pic d'activité au crépuscule ou à l'aube (Mellor *et al*, 2000; Purse *et al*, 2015). Les espèces paléarctiques utilisent principalement la végétation comme lieu de repos durant la journée (Zimmer *et al*, 2014b). Avec deux générations par an, la plupart des espèces distribuées dans l'ouest de la région paléarctique sont bivoltines (pic de population au

printemps et en été) (Blackwell *et al*, 1994; Garros and Balenghien, 2017; Meiswinkel *et al*, 2014). Lorsque les conditions environnementales ne sont pas favorables à l'émergence, certaines espèces de *Culicoides*, peuvent effectuer une diapause estivo-hivernale, permettant aux larves d'émerger à la saison suivante (Garros and Balenghien, 2017).

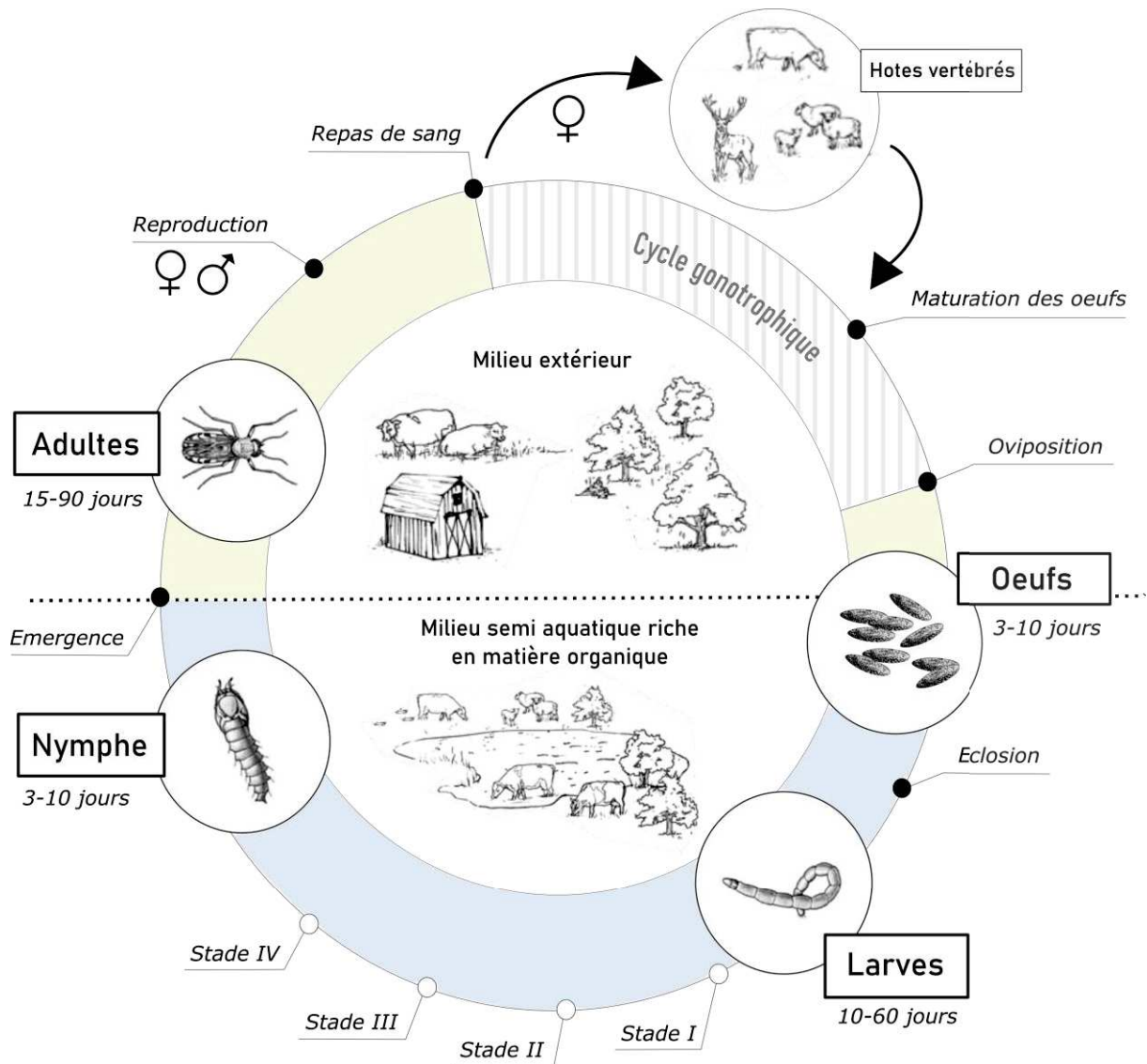


Figure 2. Représentation schématique du cycle de développement et de vie des *Culicoides* (Source : Office vétérinaire fédérale 2010, Thèse Roger Venail 2014, Purse et al 2005, Purse et al 2015, Thèse Bruno Mathieu 2011, source des dessins : <http://illuminationstudios.com/>)

Principalement mammophiles (hommes, équins, ovins, bovins, caprins, cervidés), les *Culicoides* se gorgent aussi sur d'autres types de vertébrés comme des oiseaux, des reptiles ou des amphibiens (Carpenter *et al*, 2008c; Gerry *et al*, 2009; Linden *et al*, 2010; Mullens *et al*, 2010; Viennet *et al*, 2013; Votypka *et al*, 2009). De façon plus anecdotique, des spécimens gorgés on a retrouvés sur des diptères eux-mêmes gorgés (Ma *et al*, 2013). La caractérisation

des préférences trophiques des espèces de *Culicoides* est compliquée par la difficulté de collecte de femelles gorgées sur les hôtes et le manque de connaissances sur les lieux de repos (Viennet *et al*, 2013). *Culicoides obsoletus*, *C. turanicus* et *C. impunctatus* effectuent des repas sur une large gamme d'hôtes, présentant un comportement opportuniste et généraliste. Au contraire, certaines espèces sont spécialistes, telles que *Culicoides maritimus*, *C. pictipennis* et *C. circumscriptus*, connues pour se gorger préférentiellement sur des oiseaux domestiques et sauvages (Blackwell *et al*, 1994). *Culicoides bolitinos*, *C. brevitarsis* et *C. chiopterus* sont décrits comme fortement inféodés aux bovins, que ce soit pour le repas de sang ou le site d'oviposition.

Cette section souligne qu'il reste encore beaucoup d'inconnues concernant la bio-écologie des *Culicoides*. Cela s'explique par la difficulté de mettre en élevage les espèces d'intérêt et de travailler *in naturae* sur les populations sauvages.

II.2) Le rôle des *Culicoides* dans la transmission de pathogènes en Europe du nord

Le rôle nuisible des *Culicoides* est lié à l'hématophagie des femelles. La piqûre des *Culicoides* est douloureuse et la salive peut causer des prurits et des allergies, dont la dermite estivale récidivante chez les chevaux (Anderson *et al*, 1988; Kleider and Lees, 1984) ou les moutons (Yeruham *et al*, 2004). Cette nuisance affecte aussi les activités humaines (foresterie, tourisme) et peut être si intense qu'elle a conduit à nommer certaines espèces en lien avec cette nuisance, comme *C. irritans* ou *C. diabolicus* (Auriault, 1979; Garros and Balenghien, 2017; Itoua *et al*, 1987). En s'attaquant aux populations humaines, *C. furens* a longtemps nuit aux activités touristiques sur les plages des Caraïbes et de la Floride (Linley and Davies, 1971). En Ecosse, le harcèlement incessant de *C. impunctatus*, limite les activités extérieures telles que la sylviculture, l'agriculture et le tourisme durant les pics d'abondance (Blackwell, 2001; Hendry and Godwin, 1988).

Compte tenu de leur hématophagie, les femelles de *Culicoides* sont responsables de la transmission de près de 15 espèces de protozoaires et de 25 espèces de filaires et 65 espèces de virus (Carpenter *et al*, 2013; Garros and Balenghien, 2017; Mellor *et al*, 2000; Mullens *et al*, 2010). Cependant, c'est bien leur rôle dans la transmission de virus d'importance vétérinaire et

médicale qui font des *Culicoides* des vecteurs d'intérêt pour la communauté scientifique (Figure 3).

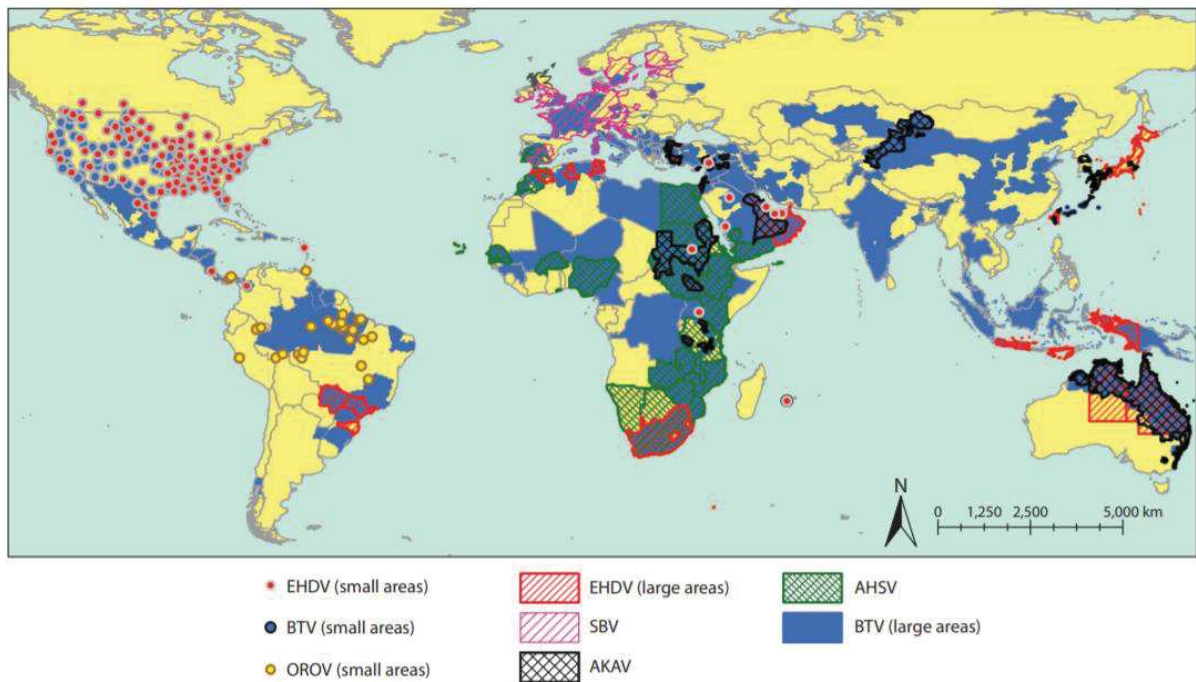


Figure 3. Distribution géographique des six principaux virus transmis par les *Culicoides* (Purse *et al*, 2015). AHSV ; virus de la peste équine ; AKAV, virus Akabane ; BTV, virus de la fièvre catarrhale ovine ; EHDV, virus de la maladie hémorragique ; OROV, virus de l'Oropouche ; SBV, virus de Schmallenberg.

Dans la région ouest paléarctique, deux arboviroses transmises par les *Culicoides* sont d'intérêts sanitaire et économique : la fièvre catarrhale ovine (FCO) et la maladie de Schmallenberg (SBV) (Beer *et al*, 2013; Carpenter *et al*, 2013; Maclachlan, 2011; Maclachlan and Guthrie, 2010). Quelques épizooties sporadiques de peste équine (PE), impliquant une transmission par *C. imicola*, ont été observées en Espagne en 1966 et 1987, mais cette maladie reste enzootique au continent africain depuis les années 1990 (Wilson and Mellor, 2009) (au moment de la rédaction de ce manuscrit, deux émergences en Asie étaient notifiées à l'OIE).

La FCO est une arbovirose non contagieuse, transmise exclusivement par les *Culicoides* (Mellor *et al*, 2000). Largement répandue dans le monde, la FCO affecte les ruminants sauvages et domestiques (Mellor *et al*, 2000; Ruiz-Fons *et al*, 2014; Talavera *et al*, 2015). Appartenant au genre *Orbivirus* (*Reoviridae*), le virus de la FCO est un virus à ARN double brin. Actuellement, 29 sérotypes de FCO sont recensés au niveau mondial (Bumbarov *et al*, 2020), les derniers ayant été respectivement identifiés chez des chèvres en Corse en 2014 (BTV-27)

(Schulz *et al*, 2016; Zientara *et al*, 2014), isolés à partir d'un vaccin au Moyen-Orient (BTV-28) (Bumbarov *et al*, 2020; Sun *et al*, 2016), et détectés chez un alpage en Afrique du sud (BTV-29) (Wright, 2014). La transmission s'effectue entre hôtes par piqûres de *Culicoides* cependant, une transmission verticale (Chauhan *et al*, 2014; Dubovi *et al*, 2013; Santman-Berends *et al*, 2010) ou horizontale par ingestion de tissus infectés est aussi rapportée (Backx *et al*, 2009; Darpel *et al*, 2009; Lopez-Olvera *et al*, 2010; Menzies *et al*, 2008)

La fièvre catarrhale ovine est aussi appelée maladie de la langue bleue en raison des hémorragies et des œdèmes faciaux qu'elle provoque. Les autres symptômes sont une forte hyperthermie, une fonte musculaire, des avortements, des écoulements nasaux et une forte salivation (Backx *et al*, 2007; Sperlova and Zendulkova, 2011). L'issue de la maladie est variable selon l'espèce touchée et le sérotype concerné, allant de l'absence de signes cliniques au décès. Le taux de mortalité oscille de 5 à 10%, pouvant s'élever à 40% dans de mauvaises conditions sanitaires. Après l'apparition des premiers symptômes, la mort peut survenir dans les dix jours qui suivent (Sperlova and Zendulkova, 2011). La fièvre catarrhale ovine est considérée comme enzootique dans les régions tropicales et tempérées situées entre les latitudes 40°S et 35°N (Walton, 2004). Cependant, des épizooties périodiques ont eu lieu principalement dans les régions adjacentes jusqu'à 58°N (Barcelo and Miranda, 2018; Carpenter *et al*, 2009; Walton, 2004; Wilson and Mellor, 2009).

Avant 1998, seules quelques incursions épizootiques avaient été constatées dans le bassin méditerranéen (Espagne, Portugal, Grèce et Chypre) (Carpenter *et al*, 2013). De 1998 à 2005, six sérotypes sont entrés en Europe du sud via l'Afrique du nord à l'ouest et via la Turquie et Chypre à l'est et se sont étendus à 12 pays (Purse *et al*, 2005). A l'issue de cette période, la limite septentrionale de la FCO dans la région s'était déplacée de 800 km. Parallèlement, *C. imicola*, le principal vecteur en Afrique et dans le bassin méditerranéen, a été collecté dans des régions au-delà de sa limite nord précédemment connue atteignant à son maximum septentrional le nord de l'Espagne, le sud de la France et le nord de l'Italie (Purse *et al*, 2005). En 2006, le sérotype 8 est entré en Europe du nord (Belgique, Allemagne et Pays-Bas), dans une zone hors de l'aire de répartition de *C. imicola*. Le virus a survécu à l'hiver et s'est étendu en 2007 à 14 pays, de l'Espagne à la Norvège (Carpenter *et al*, 2009; Garros and Balenghien, 2017; Schwartz-Cornil *et al*, 2008). Parallèlement, le sérotype 1 a été introduit dans le sud de la France via la péninsule Ibérique avec une incursion en Bretagne via l'importation d'ovins infectés (Carpenter *et al*, 2013; Sperlova and Zendulkova, 2011). Les incursions de ces différents sérotypes de FCO en Europe ont été catastrophiques sur le plan sanitaire et

économique (Ganter, 2014; Gethmann *et al*, 2020; Mellor *et al*, 2000; Wilson and Mellor, 2009; Zientara and Sánchez-Vizcaíno, 2013).

En 2011, suit l'émergence du virus de Schmallenberg (SBV), un nouveau virus pour la science transmis par les *Culicoides* (Barber *et al*, 2018). La maladie de Schmallenberg affecte comme la FCO, les ruminants domestiques et sauvages et possède des symptômes, une épidémiologie et une distribution comparable à celle de la FCO (Afonso *et al*, 2014; Garros and Balenghien, 2017; Maclachlan *et al*, 2019; Sick *et al*, 2019; Temizel *et al*, 2009). Cette arbovirose a également causé d'importantes pertes économiques pour les éleveurs de ruminants domestiques notamment par les mortalités périnatales chez les bovins et les ovins (Maclachlan *et al*, 2019).

L'émergence et la propagation massive de ces deux maladies dans la région ouest paléarctique, hors de la distribution de *C. imicola*, ont soulevé des questions sur la compétence vectorielle des *Culicoides* autochtones de la région ouest paléarctique (Barber *et al*, 2018). Des études d'infection expérimentale ou de recherche de virus chez les espèces abondantes et distribuées sur la région paléarctique ont confirmé qu'elles étaient compétentes pour la transmission des deux virus (Barber *et al*, 2018; Sick *et al*, 2019; Veronesi *et al*, 2013). En particulier, les espèces du Groupe Obsoletus, *C. obsoletus*, *C. scoticus*, *C. dewulfi* et *C. chiopterus*, sont identifiées comme des espèces vectrices majeures pour les virus de la FCO et de Schmallenberg en région paléarctique, suite à la détection ou à l'isolement du virus dans des populations sauvages (Barber *et al*, 2018; De Regge *et al*, 2012; Sick *et al*, 2019; Veronesi *et al*, 2013). Des infections expérimentales en laboratoire ont également confirmé la compétence vectorielle de *C. obsoletus* et *C. scoticus* (Carpenter *et al*, 2008a).

Ces événements ont soulevé de nombreuses questions sur les facteurs ayant permis l'expansion de la FCO. Il a été constaté que la propagation du virus de la FCO est facilitée dans des conditions favorables à l'activité des *Culicoides* et à la réplication virale. Par exemple, une température ambiante élevée tend à réduire la période d'incubation du virus (Wittmann *et al*, 2002). Les arguments avancés sont fortement liés aux changements climatiques globaux (Baylis, 2017; Purse *et al*, 2005), le changement du paysage agricole (Purse *et al*, 2005), la densité des hôtes (Jacquot *et al*, 2017), et la compétence vectorielle des espèces de *Culicoides* autochtones (notamment les espèces des Groupes Obsoletus et Pulicaris) (Goffredo *et al*, 2016; Torina *et al*, 2004).

Un facteur clé de la compréhension de l'émergence et de la large circulation des deux virus est la caractérisation de la dispersion des *Culicoides* à différentes échelles spatiales et au-dessus de différents types de milieux (continents et masses d'eaux).

II.3) La dispersion stratifiée chez les *Culicoides*

La littérature a longtemps considéré les *Culicoides* comme des diptères n'ayant pas d'importantes capacités de dispersion. Cette hypothèse se fonde sur leur petite taille ne leur permettant pas d'effectuer de vol actif de plus de quelques centaines de mètres (Murray, 1987). La dispersion active des femelles *Culicoides* est principalement motivée par la reproduction : la recherche de partenaires pour la reproduction, d'hôtes pour le repas de sang ou de sites de ponte pour l'oviposition (Kluiters *et al*, 2015b; Mellor *et al*, 2000). Compte tenu de leur taille et des limitations techniques qui en découlent, peu d'études de capture-marquage-recapture ont été menées sur les *Culicoides*.

Une première étude est réalisée en 1981 par Lillie *et al*. (Lillie *et al*, 1981) avec un marquage à la poudre fluorescente. Elle montre que des femelles de *C. variipennis* peuvent se disperser jusqu'à 4 kilomètres du point de lâcher d'origine, 36 heures après le lâcher. Une autre étude met en évidence une distance de recapture record chez *C. mohave*. Une femelle a été recapturée à 6 km du point de lâcher. Cependant, l'environnement désertique de la Californie du sud où est menée cette étude, les conditions environnementales et la très faible densité d'hôtes rendent peu généralisable les estimations de cette étude (Brenner *et al*, 1984). La dispersion des *Culicoides* est dépendante de facteurs environnementaux comme la vitesse et la direction du vent qui peuvent favoriser ou limiter la dispersion chez les *Culicoides* (Braverman and Chechik, 1996; Mellor *et al*, 1983). Un vent, au-delà de 3 mètres/seconde, pourrait limiter la dispersion de *C. schultzei* et de *C. imicola* (Walker, 1977). D'autres facteurs comme la température (si inférieure à 13°C, ou supérieure à 35°C), les précipitations (Braverman and Chechik, 1996) et des vitesses de vent élevées peuvent réduire l'activité de vol des *Culicoides* (Baylis *et al*, 2010; Carpenter *et al*, 2008b; Gloster *et al*, 2008).

En région paléarctique, une étude récente s'est intéressée à l'effet de certaines variables environnementales et climatiques sur l'abondance des *Culicoides* (Kirkeby *et al*, 2013), en ciblant deux groupes d'espèces (les Groupes *Obsoletus* et *Pulicaris*). Les auteurs utilisent un marquage avec un colorant d'isothiocyanate de fluorescéine. La distance de recapture maximale

observée est de 1750 m du site de lâcher. Cette étude met en évidence un effet positif et significatif de la vitesse du vent entre 1 et 5 m/s sur l'abondance des *Culicoides*. Les auteurs émettent l'hypothèse que, lors d'un vent de plus forte intensité, la dispersion active est limitée par la difficulté à localiser les hôtes par leur odeur. Les précipitations et les turbulences, estimées par les changements de direction du vent d'au moins 22,5 degrés ont, au contraire, des effets négatifs et significatifs sur l'abondances des *Culicoides* (Kirkeby *et al*, 2013). Ces conclusions sont en opposition avec celles d'une étude précédente, montrant une diminution de l'abondance de *Culicoides* lorsque la vitesse du vent est supérieure à 3 m/s (Carpenter *et al*, 2008c). Cependant, il est important de noter une différence importante entre les deux études. Le niveau d'identification taxonomique est au groupe pour l'étude de Kirkerby et al (2013), et non pas à l'espèce, ce qui limite la portée des conclusions. En effet, il est connu que la saisonnalité ou l'utilisation de l'habitat peuvent différer au niveau de l'espèce, même au sein du même groupe d'espèces (Searle *et al*, 2014). Une étude pharaonique est par la suite publiée par Kluiters et ses collaborateurs (Kluiters *et al*, 2015a; Kluiters *et al*, 2015b). Ils utilisent de la poussière fluorescente pour marquer plus de 61 000 *Culicoides* afin d'étudier la dispersion des membres du Groupe Obsoletus au Pays de Galles. Malgré le très grand nombre d'individus relâchés, le taux de recapture est de 0.02%. La distance parcourue varie entre 1 et 2,5 km, avec une dispersion plus importante chez les femelles (2,21 km en moyenne contre seulement 1 km pour les mâles). Le point important à noter est que des captures ont été effectuées à la fois dans le sens du vent mais aussi contre la direction du vent. Cette étude conclue à une possible dispersion passive et active des *Culicoides* de proche en proche entre fermes (Kluiters *et al*, 2015a; Kluiters *et al*, 2015b). Sander et ses collaborateurs (Sanders *et al*, 2017) utilisent une autre approche de marquage pour quantifier la dispersion de *Culicoides* en Angleterre, avec une technique d'immunomarquage à l'ovalbumine. Le marquage est effectué dans cette étude directement par exposition de contact et par pulvérisation sur les sites de repos et d'oviposition supposés, ce qui peut engendrer un nombre important de contamination. Les auteurs ont une analyse plus poussée des données météorologiques leur permettant d'aller plus loin sur l'analyse du rôle du vent. Ils mettent en évidence que 84,4 % des vols de plus d'un kilomètre ont eu lieu dans le sens du vent alors que seulement 15,6 % des vols ont été effectués face au vent (Sanders and Carpenter, 2014; Sanders *et al*, 2017).

Les études sur les *Culicoides* appliquant une méthodologie de capture-marquage-recapture restent délicates compte tenu de la taille de l'insecte, du taux de mortalité des *Culicoides*

capturés et du faible taux de recapture (Kluiters *et al*, 2015b). Les trois exemples développés ci-dessus nous permettent de conclure que les capacités de dispersion des *Culicoides* seraient comprises entre 1 et 2,5 km (Kirkeby *et al*, 2013; Kluiters *et al*, 2015b; Sanders and Carpenter, 2014), avec une variation liée au sexe, que le vol se ferait de manière omnidirectionnelle, avec des trajectoires dans le sens du vent, mais également contre le vent (Brenner *et al*, 1984; Kirkeby *et al*, 2013). Si les capacités de dispersion horizontales actives restent faibles, des études de piégeages ont montré la présence des *Culicoides* dans des colonnes d'air à plusieurs centaines de mètres d'altitude (Bishop *et al*, 2000; Hardy and Cheng, 1986; Murray, 1987; Sanders *et al*, 2011), indiquant un potentiel dispersion verticale passive.

D'autres arguments existent qui soutiennent une importante dispersion passive sur de longues distances au-dessus des masses d'eau, grâce aux vents. Ces événements de dispersion par le vent de *Culicoides*, ont permis la colonisation et l'installation de nouvelles populations dans des zones indemnes (Johansen *et al*, 2003) comme en France depuis la Corse (Jacquet *et al*, 2016a). Le transport de *Culicoides* infectés aurait permis l'introduction de sérotypes du virus de la FCO dans des territoires insulaires ou entre régions géographiques éloignées (Aguilar-Vega *et al*, 2019; Alba *et al*, 2004; Hendrickx *et al*, 2008; Jacquet *et al*, 2016a; Sedda *et al*, 2015). De telles introductions ont ainsi été démontrées de la Sardaigne aux Baléares (Alba *et al*, 2004), mais également en Irlande (McGrath *et al*, 2018) et au Royaume-Uni (Burgin *et al*, 2017).

Ces études s'appuient sur des approches de modélisation où les trajectoires de dispersion sont estimées par des modèles de dispersion atmosphérique de particules fines comme HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) (McGrath *et al*, 2018), NAME (Numerical Atmospheric-dispersion Modelling Environment) (Andrew Jones, 2007; Eagles *et al*, 2014), ou MATCH (Multiscale Atmospheric Transport and Chemistry) (Agren *et al*, 2010). Initialement, ces modèles sont conçus pour simuler le transport de gaz ou des particules fines en suspension dans l'air. Jacquet et ses collaborateurs (Jacquet *et al*, 2016b) mettent en évidence, grâce à l'utilisation combinée de paramètres de génétique des populations classique, de méthodes bayésiennes ABC et d'un modèle de dispersion par le vent (NAME), le rôle des populations de *C. imicola* de Corse comme source d'introduction la plus probable de *Culicoides* dans les Pyrénées-Orientales via des événements de dispersion à longue distance par le vent de plus de 500 km (Jacquet *et al*, 2016b).

Les études de modélisation de la dispersion appliquées à des territoires terrestres sont plus rares. En 2015, Sedda et ses collaborateurs utilisent les données de contamination des fermes par le

virus de la FCO collectées au cours de l'épizootie de 2006 en Europe pour les comparer à des modèles mécanistes et stochastiques de dispersion de *Culicoides* par le vent (Sedda *et al*, 2012; Sedda *et al*, 2015). Ce modèle permet d'expliquer 94 % des données d'infections entre fermes en incluant des vols dans le sens du vent mais également à contre-courant, le modèle prenant en compte uniquement le vol dans le sens du vent n'étant pas assez informatif (Sedda and Rogers, 2013). Les auteurs montrent que près de 54 % des foyers observés auraient été causés par la dispersion du vecteur sur moins de 5 km, et 8 % des infections causées par un déplacement du vecteur sur plus de 31 km. Ainsi, ils concluent que les infections à longue distance causées par les *Culicoides* sont rares. De ce fait, la dynamique d'infection serait principalement effectuée par de la dispersion de proche en proche illustrée avec et contre le vent. Enfin, une étude de modélisation a permis de quantifier le rôle de la dispersion des vecteurs et des déplacements d'animaux dans la transmission entre fermes du virus de Schmallenberg (SBV) et de la FCO en Grande-Bretagne (Sumner *et al*, 2017). Les résultats de l'étude indiquent que respectivement 90% et 98% des événements de transmission entre exploitations de FCO et de SBV seraient dus à la dispersion des *Culicoides*. Sur la base de leurs résultats, les auteurs concluent que l'impact des mesures de restrictions de mouvement des animaux sur la circulation du virus dépend des hypothèses posées sur la capacité de dispersion des vecteurs. Si la dispersion se produit principalement à l'échelle locale (99 % de la transmission dans une zone <25 km), les restrictions de mouvement seront efficaces pour réduire la circulation du virus, mais si la dispersion se produit fréquemment sur de plus longues distances (99 % de la transmission dans une zone au-delà de 50 km), alors elles ne le seront pas.

Il a également été décrit que les *Culicoides* pouvaient s'élever activement dans les vents en altitude puis être poussées par les courants de vents de manière passive (Drake and Reynolds, 2012; Huestis *et al*, 2019). La dispersion par le vent est alors qualifiée de semi-passive.

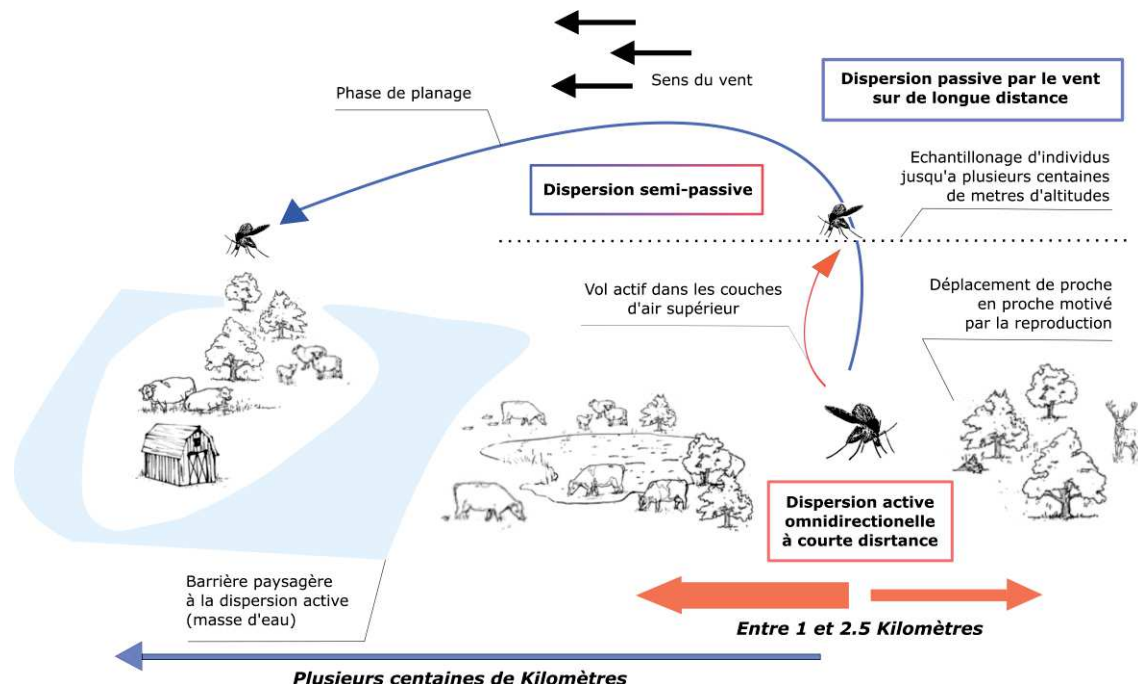


Figure 4. Représentation schématique de la dispersion stratifiée des *Culicoides*. Le rouge implique une dispersion active et le bleu une dispersion passive. L'épaisseur des flèches rouges et bleues est représentative de l'intensité de la dispersion (Source Office vétérinaire fédérale 2010, Thèse Roger Venail 2014, Purse *et al* 2005, Purse *et al* 2015, Thèse Bruno Mathieu 2011).

En résumé, la dispersion des *Culicoides* est qualifiée de « stratifiée » (Figure 4), car elle associe de la dispersion active à faible distance de proche en proche et de la dispersion passive par le vent sur de longues distances, plus ou moins rare selon les masses traversées. Cette section souligne (i) la forte divergence d'opinions sur l'importance de la dispersion par le vent (ii) le fait que les capacités de dispersion à longue distance au-dessus des masses continentales ont été peu évaluées que ce soit par approche directe (CMR) ou indirecte (génétique, modélisation) (Burgin *et al*, 2013) (iii) et l'absence d'études moléculaires visant à étudier la dispersion de *C. obsoletus* à l'aide d'outils de génétique des populations et du paysage.

II.4) *Culicoides obsoletus* : un besoin de clarification taxonomique

L'application des méthodes de la taxonomie moléculaire au genre *Culicoides* est relativement récente, en lien avec les émergences et l'évolution des situations sanitaires dans les régions touchées par les maladies transmises. Compte tenu de cet intérêt récent en Europe, l'étude de la taxonomie et de la phylogénie des *Culicoides* paléarctique est encore naissante. Il reste donc de nombreux questionnements sur le statut spécifique de certaines entités et les relations phylogénétiques entre les différentes espèces de *Culicoides*.

Le genre *Culicoides* (Diptera: Ceratopogonidae), défini par Latreille en 1809, regroupe actuellement environ 1 400 espèces décrites, groupées en 31 sous-genres (Borkent, 2016). Plus de 10% des espèces ne sont pas affiliées à un groupe et 38 complexes d'espèces ou groupes ne sont pas affiliés à un sous-genre (Borkent, 2016). Les groupes d'espèces sont pour la plupart définis sur la base de similarités morphologiques sans réelle signification phylogénétique. La configuration des patrons alaires, des palpes maxillaires, des pièces génitales ou des antennes sont des critères diagnostiques pour l'identification à l'espèce (Augot *et al*, 2010; Delecolle, 1985; Hajd Henni *et al*, 2014; Talavera *et al*, 2017). La notion de groupe ou de complexe d'espèces n'a pas réellement été définie dans le genre *Culicoides* (Borkent, 2016; Harrup *et al*, 2015). Enfin, pour compléter l'imbroglio taxonomique, de nombreuses espèces sont dites cryptiques ou jumelles, c'est-à-dire morphologiquement identiques mais moléculairement différentes. Ainsi, la taxonomie du genre *Culicoides* est problématique, et nécessite une révision systématique (Harrup *et al*, 2015)

Le Groupe *Obsoletus* est défini dans la littérature comme un groupe d'espèces avec une morphologie marquée par des ailes tachetées mal définies et une deuxième cellule radiale avec une tâche lumineuse. Neuf espèces valides composent le Groupe *Obsoletus* tel que décrit par Mathieu en 2011 : *C. obsoletus* (Meigen), 1818 ; *C. sinanoensis* Tokunaga, 1937 ; *C. scoticus* Downes & Kettle, 1952 ; *C. montanus* Shakirzjanova, 1962 ; *C. gornostaevae* Mirzaeva, 1984 ; *C. abchazicus* Dzhafarov, 1964 ; *C. filicinus* Gornostaeva & Gachegova, 1972 ; *C.alachua* Jamnback & Wirth, 1963 ; et *C. sanguisuga* (Coquillet, 1901). Parmi ces espèces, *C. abchazicus*, *C. montanus*, *C. obsoletus* et *C. scoticus* sont sympatriques en région ouest paléarctique. *Culicoides obsoletus* est considéré comme une espèce holarctique car présente à la fois dans les régions néarctique et paléarctique (Mathieu, 2011). Abondantes dans les élevages de ruminants domestiques (Baldet *et al*, 2008; Meiswinkel *et al*, 2014), *C. scoticus* et *C. obsoletus* sont des espèces dites « jumelles » car les femelles sont morphologiquement très

proches avec un chevauchement de valeurs au niveau de la taille des spermathèques et des taches alaires très similaires (Figure 1). *Culicoidesalachua* et *C. sanguisuga* sont exclusivement présentes dans la région néarctique (Mathieu, 2011). Les cycles de vie et les comportements trophiques de la plupart des espèces du Groupe Obsoletus ne sont pas bien décrits. Historiquement associés au Groupe Obsoletus sur la base de critères morphologiques et de leur sympatrie, *C. chiopterus* et *C. dewulfi* ont été écartées du Groupe Obsoletus récemment suite à des études moléculaires et de morphométrie géométrique des ailes (Ander *et al*, 2013; Gomulski *et al*, 2005; Hajd Henni *et al*, 2014; Kluiters *et al*, 2016; Meiswinkel *et al*, 2004). Au sein du Groupe Obsoletus, un complexe d'espèces, nommé lui aussi Obsoletus, a été défini et inclut *C. obsoletus*, *C. montanus*, *C. abchazicus* et *C. scoticus* (Mathieu, 2011). Récemment, des études moléculaires au sein du Groupe Obsoletus basées sur le polymorphisme de marqueurs mitochondriaux, ont montré l'existence d'espèces cryptiques : *C. obsoletus* clade 'O2', *C. obsoletus* clade 'O3' respectivement en Suède et en Suisse (Ander *et al*, 2013), et *C. obsoletus* clade « dark » aux Pays-Bas (Elbers and Meiswinkel, 2014; Meiswinkel *et al*, 2015).

Afin de respecter les règles posées par Harbach pour le genre *Anopheles* pour les catégories infrasubgénériques (Harbach, 1994), nous utiliserons pour ces travaux de thèse le terme "Complexe Obsoletus/Scoticus" pour désigner l'ensemble d'espèces regroupant *C. obsoletus*, *C. abchazicus*, *C. scoticus* et *C. montanus* ainsi que toutes les unités taxonomiques opérationnelles décrites récemment dans la littérature (*C. obsoletus* clade O2, *C. obsoletus* clade O3, *C. obsoletus* clade « dark »).

Encadré 6 : Les différents marqueurs moléculaires utilisés pour estimer les relations phylogénétiques des *Culicoides*

Étant donné la difficulté d'identification morphologique et l'imbricatio taxonomique actuel, des techniques d'amplification en chaîne par polymérase (PCR) avec amorces spécifiques, de spectrométrie de masse (Kaufmann *et al*, 2012) de PCR en temps réel (Monaco *et al*, 2010; Talavera *et al*, 2015; Wenk *et al*, 2012), ou de puces à ADN (Deblauwe *et al*, 2012; Duvallat *et al*, 2017), ont été mises en place. Divers marqueurs moléculaires ont été utilisés pour l'identification moléculaire des *Culicoides*, notamment les gènes du cytochrome c oxydase mitochondrial sous-unité I (cox1, région du code à barres), sous-unité II (COII), l'ADN ribosomique 16S (Debeila, 2010; Emaresi *et al*, 2009; Suk and Semenza, 2011; Tabachnick, 2004), l'ADN ribosomique 28S (Bakhoun *et al*, 2018), la cytochrome oxydase B (Cytb), la

région d'espacement interne transcrite 1 (ITS1) (Cetre-Sossah *et al*, 2004; Duvallet *et al*, 2017; Jacquot *et al*, 2017; Mathieu, 2011; Mathieu *et al*, 2007) et l'ADNr ITS2 (Gomulski *et al*, 2005; Gomulski *et al*, 2006). Parmi ce panel de marqueurs moléculaires, le plus utilisé est la sous-unité 1 de la cytochrome c oxydase mitochondriale (cox1) (Nolan *et al*, 2007). Parmi les trois sous-unités de cytochrome oxydase dans l'ADN mitochondrial, le gène cox1 est le plus conservé (Zhang and Hewitt, 1997), et est couramment utilisé dans les études de phylogénie chez les arthropodes (Ander *et al*, 2013; Hebert *et al*, 2003). Ce fragment du gène mitochondrial cox1 a été utilisé pour développer des tests PCR avec amorces spécifiques (Cetre-Sossah *et al*, 2004; Garros and Balenghien, 2017; Mathieu, 2011; Nolan *et al*, 2007; Schwenkenbecher *et al*, 2009), ou pour identifier des espèces morphologiquement indifférenciables (Pages *et al*, 2009; Pages and Sarto, 2005; Schwenkenbecher *et al*, 2009). La région cox1 a été utilisée pour étudier les relations phylogénétiques des *Culicoides* afrotropicaux, comme *C. imicola* (Dallas *et al*, 2003; Linton *et al*, 2002; Suk and Semenza, 2011).

Culicoides obsoletus est présent sur les trois continents de l'hémisphère nord, l'Amérique, l'Europe et l'Asie (Borkent, 2016), avec un centre de distribution qui semble être en Europe. Des notifications de l'espèce existent pour une zone allant de l'ouest du continent américain, la Colombie britannique, Canada (Anderson *et al*, 1991), la Russie (Sprygin *et al*, 2014), la Chine (Gao *et al*, 2020) et la Corée du sud (Cho and Chong, 1974) à l'est, l'Afrique du nord au sud, (Maroc (Bourquia *et al*, 2019), Tunisie (Sghaier *et al*, 2017), Algérie (Hakima *et al*, 2020)), jusqu'au Portugal (Capela *et al*, 2003) et à l'Espagne (Miranda *et al*, 2003) pour l'ouest de l'Europe, et la Finlande et la Suède pour le nord de l'Europe (Mignotte *et al*, 2020). Cette espèce est également présente dans les territoires insulaires européens, comme Les Baléares (Miranda *et al*, 2003), les Canaries (Martinez-de la Puente *et al*, 2012), et la Corse (Venail *et al*, 2012). Cette espèce est considérée comme paléarctique et néarctique (Borkent, 2016; Mathieu, 2011). De nombreuses espèces synonymes ont été décrites comme *C. varius* ((Winnertz), 1852) en Allemagne, *C. yezoensis* ((Matsumura), 1911) en Russie, *C. lacteinervis* (Kieffer, 1919) en Ukraine, *C. kabyliensis* (Kieffer, 1922) en Algérie ou *C. intermedius* (Okada, 1941) au Japon.



Figure 5. Aire de répartition mondiale de *C. obsoleteus*. Le gris indique les pays dans lesquels l'espèce est présente, le blanc indique les pays dans lesquels n'a pas été rapportée jusqu'à présent.

Son comportement trophique est décrit comme généraliste et ses préférences écologiques des stades immatures sont peu exigeantes, lui permettant d'être présent dans de nombreux habitats. *Culicoides obsoleteus* est abondant dans toute l'Europe centrale et septentrionale (France (Cuellar *et al*, 2018), Roumanie (Hristescu *et al*, 2020), Slovaquie (Sarvasova *et al*, 2014), et péninsule des Balkans (Goffredo *et al*, 2004; Mignotte *et al*, 2020; Pudar *et al*, 2018). Il est cependant moins abondant sur ses fronts septentrionaux et méridionaux en Europe de l'ouest, Norvège et Finlande (Mignotte *et al*, 2020) et Maroc (Bourquia *et al*, 2019), Tunisie (Sghaier *et al*, 2017), ou Algérie (Hakima *et al*, 2020).

Dans ce contexte de flou taxonomique, et sachant sa vaste répartition géographique, il apparaît primordial de clarifier la situation phylogénétique du Groupe Obsoleteus.

Problématique et objectifs de la thèse

L'Europe connaît depuis quinze ans des épizooties récurrentes de virus d'intérêt vétérinaire, le virus de Schmallenberg (SBV) et le virus de la Fièvre Catarrhale Ovine (FCO). Ces virus ont été responsables de crises sanitaires sans précédent en Europe. Des espèces de *Culicoides* paléarctiques autochtones et notamment celles appartenant au Groupe Obsoletus sont connues comme étant les vecteurs principaux de ces virus dans la région ouest paléarctique. Les analyses épidémiologiques des épizooties de FCO ou de SBV, couplées à des approches de modélisation soulignent le rôle : (i) du vent dans la dispersion à longue et à courte distance des *Culicoides* et donc des virus qu'ils transmettent (Aguilar-Vega *et al*, 2019; Burgin *et al*, 2017; Burgin *et al*, 2013; Ducheyne *et al*, 2007; Jacquet *et al*, 2016a; McGrath *et al*, 2018), (ii) que la circulation du virus de la FCO s'effectue majoritairement de proche en proche (Sedda *et al*, 2012; Sedda *et al*, 2015), (iii) et qu'un arrêt des mouvements animaux comme mesure de contrôle n'aurait pas affecté la transmission du virus de Schmallenberg (Gubbins *et al*, 2014).

L'utilisation de la méthode CMR a permis de mettre en évidence le caractère omnidirectionnel de la dispersion des *Culicoides*, à la fois dans le sens du vent mais aussi contre la direction du vent à courte distance. De plus, la dispersion des *Culicoides* sur de très longues distances (plusieurs centaines de km) au-dessus des mers en utilisant les courants d'airs chauds et de faible vitesse est bien documentée. Celle-ci est facilitée par le fait que les masses d'eau présentent peu de résistance aux flux d'air horizontaux. Au-dessus des terres, du fait des phénomènes de friction liés à la rugosité du terrain ou à la présence de barrières physiques (relief), les phénomènes de dispersion se font à de plus courtes distances (de quelques km à plusieurs dizaines de km) et sont composés à la fois de mouvements actifs et de déplacements passifs des *Culicoides*. L'ensemble de ces éléments permet de faire l'hypothèse que la dispersion des *Culicoides* correspond à ce que les écologues appellent la dispersion stratifiée.

Si le rôle des vents et des mouvements d'air locaux a un rôle indéniable dans la dispersion des *Culicoides*, à l'échelle locale, d'autres facteurs jouent également un rôle prépondérant. La distribution des hôtes dans le paysage impacte la dispersion active des *Culicoides* sur de petites distances. La présence d'hôtes conditionne l'accès au repas de sang indispensable pour le développement des œufs et constitue des milieux extrêmement favorables aux *Culicoides* en offrant des gîtes larvaires nécessaires au développement des immatures. Ainsi, la configuration et la composition du paysage (couvert végétal, densité d'hôtes) influencent la dispersion des

Culicoides à faible échelle. Les conditions climatiques, comme la température, la vitesse du vent, l'humidité, ou les précipitations, impactent la survie des *Culicoides* et donc par conséquent leurs capacité de dispersion.

Dans un contexte actuel et mondial favorable à la propagation de maladies vectorielles, il est important de comprendre l'impact de la dispersion stratifiée des *Culicoides* sur la structure des populations, ainsi que l'influence des facteurs biotiques et abiotiques.

L'objectif de ce travail de thèse est de déterminer par l'observation de patrons de distributions des fréquences alléliques, comment la dispersion, mais également des événements évolutifs plus anciens, ont pu façonner l'arrangement spatial de la diversité génétique et notamment comment elle est influencée par les paysages et par les caractéristiques des habitats qui peuvent selon le cas favoriser ou limiter les mouvements des *Culicoides*. Le Groupe *Obsoletus* est un groupe d'espèces difficilement différenciables morphologiquement, notamment pour les femelles. Au sein de ce groupe, le Complexe *Obsoletus/Scoticus*, tel qu'il se présente actuellement dans la littérature, est confus et nécessite une révision. De plus, les espèces cryptiques sympatriques peuvent présenter des compétences vectorielles différentes et peuvent ainsi biaiser les enquêtes épidémiologiques mais également nos conclusions sur l'étude de la dispersion des *Culicoides*. Il nous donc est apparu nécessaire d'évaluer la diversité intra- et interspécifique au sein de ce complexe

Le **premier chapitre** de cette thèse est consacré à l'identification et à la description de la diversité cryptique du Complexe *Obsoletus/Scoticus* à l'échelle européenne. Des analyses de phylogénie moléculaire et de délimitation d'espèces avec une approche multi-marqueurs ont permis d'explorer la diversité génétique au sein du Complexe *Obsoletus/Scoticus* dans la région ouest paléarctique sur un transect allant de la Scandinavie aux îles Canaries (du nord au sud) et des îles Canaries à la Turquie (d'ouest en est). Cette étude a permis de cartographier la diversité d'espèces valides ou cryptiques du complexe, de clarifier les relations phylogénétiques au sein du Complexe *Obsoletus/Scoticus* et de remettre en question le statut taxonomique de certains clades nouvellement décrits. Nous proposons ici une redéfinition du complexe d'espèce *Obsoletus/Scoticus*.

La suite de ce travail s'est concentrée exclusivement sur l'espèce *C. obsoletus*.

Le **second chapitre** est dédié à l'identification de flux de gènes entre populations de l'espèce *C. obsoletus* à l'échelle continentale. Il s'intéresse à décrire l'histoire évolutive et à

émettre des hypothèses sur les événements historiques ayant façonné le schéma actuel de structuration génétique ainsi qu'à caractériser l'influence des phénomènes de dispersion à longue distance sur la structure génétique des populations. Pour cela, 11 marqueurs microsatellites nouvellement développés et le gène mitochondrial *cox1*, ont été utilisés pour identifier la structure génétique et phylogéographique de *C. obsoletus* à l'échelle européenne.

La dispersion des *Culicoides*, est influencée par la composition et la configuration du paysage qui va selon ses caractéristiques, favoriser (conductance) ou freiner (résistance) la dispersion de ces diptères. Le **troisième chapitre** vise à identifier les facteurs environnementaux qui favorisent ou limitent les flux génétiques de *C. obsoletus* à l'échelle française en utilisant un cadre méthodologique innovant intégrant des approches spatiales, statistiques et de génétique des populations. Pour cela, une approche individuelle a été développée afin de détecter des schémas spatiaux de différenciation génétique et les facteurs environnementaux qui en sont la cause.

Enfin, l'intégralité des résultats obtenus durant ces travaux de thèse sont synthétisés puis discutés, en relation avec des questions évolutives et écologiques. Un regard critique sera apporté sur les résultats et des axes de développements méthodologiques seront proposés. Les nouveaux éléments de compréhension sur la structuration et les capacités de dispersion des *Culicoides* paléarctiques seront analysés en relation avec l'évaluation du risque épidémiologique et les stratégies de lutte. Nos résultats fournissent des clés pour le développement de modèles prédictifs de dispersion des *Culicoides* pouvant faciliter la surveillance et les mesures de contrôle

Chapitre I : The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palaearctic vector *Obsoletus/Scoticus* Complex (Diptera: Ceratopogonidae) at the European level

Une bonne connaissance de la diversité des espèces vectrices est capitale pour la mise en place de méthodes de lutte anti-vectorielle adaptées. Encore faut-il être sûr que nous avons affaire à la bonne espèce et qu'il s'agit bien d'une seule espèce. C'est dans cette optique que prend place ce premier chapitre.

A la suite des épizooties massives du virus de la FCO dans le nord de l'Europe en 2006, puis de la maladie de Schmallenberg en 2011, en l'absence de *C. imicola*, *C. obsoletus* a très rapidement été identifiée comme l'une des espèces vectrices majeures en Europe du nord en raison de son abondance et de sa compétence vectorielle confirmée.

Mais, derrière *C. obsoletus*, se cachent de nombreuses espèces, valides ou nouvellement et sommairement décrites. D'ailleurs, cette profusion décourage de nombreux auteurs comme en témoignent la plupart des études qui se limitent à l'identification des individus au niveau du Groupe *Obsoletus* et non à l'espèce. Neuf espèces valides composent le Groupe *Obsoletus* tel que décrit par Mathieu en 2011, et seulement six sont sympatriques en région ouest paléarctique : *C. obsoletus* (Meigen), 1818 ; *C. scoticus* Downes & Kettle, 1952 ; *C. montanus* Shakirzjanova, 1962 ; *C. gornostaevae* Mirzaeva, 1984 ; *C. filicinus* Gornostaeva & Gachegova, 1972 ; et *C. abchazicus* Dzhafarov, 1964. Parmi ce groupe, *C. scoticus* et *C. obsoletus* sont des espèces dites « jumelles », car les femelles sont quasi indifférentiables morphologiquement, hormis pour certains experts du genre. Un complexe d'espèces, nommé *Obsoletus*, a été défini et inclut *C. obsoletus*, *C. montanus*, *C. abchazicus* et *C. scoticus*. De plus, des études de taxonomie moléculaire ont montré récemment l'existence d'entités cryptiques : *C. obsoletus* clade 'O2', *C. obsoletus* clade 'O3' respectivement en Suède et en Suisse et *C. obsoletus* clade « dark » aux Pays-Bas. Il en résulte, un ensemble d'espèces cryptiques ou jumelles que nous appellerons "Complexe *Obsoletus/Scoticus*" pour désigner un ensemble regroupant *C. obsoletus*, *C. abchazicus*, *C. scoticus* et *C. montanus* ainsi que toutes les unités taxonomiques opérationnelles décrites récemment dans la littérature (*C. obsoletus* clade O2, *C. obsoletus* clade O3, *C. obsoletus* clade « dark »).

Ce premier chapitre présente une étude de taxonomie moléculaire multi-marqueurs du Complexe *Obsoletus/Scoticus* qui permet de clarifier l'imbricatio taxonomique actuel. Pour la première fois dans l'étude de ce complexe, le jeu de données inclut des individus issus de 20 pays de la région ouest paléarctique, de la Scandinavie aux îles Canaries (du nord au sud) et des îles Canaries à la Turquie (d'ouest en est). Nous avons utilisé conjointement deux marqueurs mitochondriaux (Cox1 et rDNA 16S) et un marqueur nucléaire (rDNA 28S). Cette approche multi-marqueurs a permis (i) d'identifier le statut spécifique des spécimens collectés et de

sélectionner les individus *C. obsoletus* pour la suite de nos travaux, (ii) de cartographier la diversité du Complexe Obsoletus/Scoticus à l'échelle ouest paléarctique, et (iii) de réviser le schéma taxonomique du Complexe Obsoletus/Scoticus. Nous confirmons le niveau important de diversité cryptique trouvée au sein du Complexe Obsoletus/Scoticus. Les méthodes de délimitation d'espèces nous ont permis de délimiter cinq MOTUs et (i) d'apporter la preuve de la validité taxonomique de *C. obsoletus* clade O2 et *C. obsoletus* clade dark (ii) d'identifier des individus appartenant à une espèce non encore décrite ou non présente dans les bases de données et (iii) de remettre en question le statut taxonomique de *C. montanus*.

Nos travaux apportent pour la première fois une clarification taxonomique à l'échelle européenne permettant une redéfinition du complexe d'espèces Obsoletus/Scoticus. Cette étude a permis également de cartographier la diversité des espèces valides ou cryptiques du complexe et de remettre en question le statut taxonomique de certains clades nouvellement décrits.

Ce chapitre a été accepté et publié dans le journal *Parasites and Vectors* :

DOI: <https://doi.org/10.1186/s13071-020-04114-1>

RESEARCH

Open Access



The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palaearctic vector *Obsoletus/Scoticus* Complex (Diptera: Ceratopogonidae) at the European level

Antoine Mignotte^{1,2*}, Claire Garros^{1,2*}, Laetitia Gardès^{1,3}, Thomas Balenghien^{1,2,4}, Maxime Duhayon^{1,2}, Ignace Rakotoarivony^{1,2}, Laura Tabourin^{1,2}, Léa Poujol^{1,2}, Bruno Mathieu⁵, Adolfo Ibañez-Justicia^{6†}, Ahmet Deniz^{7†}, Aleksandar Cvetkovikj^{8†}, Bethan V. Purse^{9†}, David W. Ramilo^{10†}, Despoina Stougiou^{11†}, Doreen Werner^{12†}, Dubravka Pudar^{13†}, Dušan Petrić^{13†}, Eva Veronesi^{14†}, Frans Jacobs^{6†}, Helge Kampen^{15†}, Isabel Pereira da Fonseca^{10†}, Javier Lucientes^{16†}, Javier Navarro^{17†}, Josue Martinez de la Puente^{18,19†}, Jovana Stefanovska^{8†}, Kate R. Searle^{20†}, Khalid Khallaayoune^{4†}, C. Lorna Culverwell^{21†}, Magdalena Larska^{22†}, Maria Bourquia^{2,4†}, Maria Goffredo^{23†}, Marina Bisia^{11†}, Marion England^{24†}, Matthew Robin^{25†}, Michela Quaglia^{23†}, Miguel Ángel Miranda-Chueca^{26†}, René Bødker^{27†}, Rosa Estrada-Peña^{16†}, Simon Carpenter^{24†}, Simona Tchakarova^{28†}, Sofia Boutsini^{11†}, Ståle Sviland^{29†}, Stefanie M. Schäfer^{9†}, Zanda Ozoliņa^{30†}, Zanda Segliņa^{30†}, Zati Vatansever^{7†} and Karine Huber¹

Abstract

Background: *Culicoides obsoletus* is an abundant and widely distributed Holarctic biting midge species, involved in the transmission of bluetongue virus (BTV) and Schmallenberg virus (SBV) to wild and domestic ruminants. Females of this vector species are often reported jointly with two morphologically very close species, *C. scoticus* and *C. montanus*, forming the *Obsoletus/Scoticus* Complex. Recently, cryptic diversity within *C. obsoletus* was reported in geographically distant sites. Clear delineation of species and characterization of genetic variability is mandatory to revise their taxonomic status and assess the vector role of each taxonomic entity. Our objectives were to characterize and map the cryptic diversity within the *Obsoletus/Scoticus* Complex.

Methods: Portion of the *cox1* mitochondrial gene of 3763 individuals belonging to the *Obsoletus/Scoticus* Complex was sequenced. Populations from 20 countries along a Palaearctic Mediterranean transect covering Scandinavia to

*Correspondence: antoine.mignotte@cirad.fr; claire.garros@cirad.fr

†All of the authors participated equally by providing samples

¹ ASTRE, Univ Montpellier, Cirad, INRAE, Montpellier, France

² Cirad, UMR ASTRE, 34398 Montpellier, France

Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Canary islands (North to South) and Canary islands to Turkey (West to East) were included. Genetic diversity based on *cox1* barcoding was supported by 16S rDNA mitochondrial gene sequences and a gene coding for ribosomal 28S rDNA. Species delimitation using a multi-marker methodology was used to revise the current taxonomic scheme of the *Obsoletus/Scoticus* Complex.

Results: Our analysis showed the existence of three phylogenetic clades (*C. obsoletus* clade O2, *C. obsoletus* clade dark and one not yet named and identified) within *C. obsoletus*. These analyses also revealed two intra-specific clades within *C. scoticus* and raised questions about the taxonomic status of *C. montanus*.

Conclusions: To our knowledge, our study provides the first genetic characterization of the *Obsoletus/Scoticus* Complex on a large geographical scale and allows a revision of the current taxonomic classification for an important group of vector species of livestock viruses in the Palaearctic region.

Keywords: *Culicoides* spp., Cryptic species, Phylogeny, Taxonomy, Species delimitation, Palaearctic Region, Biting midge

Background

In 2006, northern Europe faced massive outbreaks of bluetongue disease (BTV), a *Culicoides*-borne viral infection which affects wild and domestic ruminants. This was followed by the emergence of Schmallenberg virus (SBV) in 2011, another *Culicoides*-borne virus, which also caused important economic losses for farmers of domestic ruminants [1]. Both, the emergence and massive spread of these diseases over the Palaearctic region raised questions about the vector competence of native Palaearctic biting midges members of the genus *Culicoides* [2, 3]. Quickly following these epizootics, studies confirmed *Culicoides* species of the subgenus *Avaritia* as the most likely vector species and particularly those of the *Obsoletus* group (see below) [3–5]. *Culicoides obsoletus*, *C. scoticus*, *C. dewulfi* and *C. chiopterus* are reported to be implicated in BTV and SBV transmission in Europe based on virus detection or isolation in field-collected populations [6]. Laboratory experimental infections have also confirmed the vector competence of *C. obsoletus* and *C. scoticus*, highlighting species variations in their competence level [7].

The literature defines the *Obsoletus* Group as a group of species with a similar morphology, especially for the characters commonly used for the identification of these insect vectors, namely wing spot pattern (poorly defined spotted wings and a second radial cell with a light spot) and distinctive male genitalia [8–10]. Adults of species in this Group are known to be abundant, widespread across central and northern Europe, and are characterized by long seasonal occurrence [11, 12]. At present, the group is an artificial taxonomic entity still poorly defined, with no real consensus on the included species, with variable internal groupings and naming. Indeed, the world catalogue of *Culicoides* does not account for levels below subgenus and does not identify species complexes as groups [13]. At present, the *Obsoletus* Group is composed of nine valid species: *C. obsoletus* (Meigen), 1818; *C. sinanoensis*

Tokunaga, 1937; *C. scoticus* Downes & Kettle, 1952; *C. montanus* Shakirzjanova, 1962; *C. gornostaevae* Mirzaeva, 1984; *C. abchazicus* Dzhafarov, 1964; *C. filicinus* Gornostaeva & Gachegova, 1972; *C.alachua* Jamnback & Wirth, 1963; and *C. sanguisuga* (Coquillett, 1901). The latter two are the only species belonging to this group exclusively present in the Nearctic region, while the others, are sympatric in the Palaearctic region [8]. *Culicoides obsoletus* is considered Holarctic because it is present in both Nearctic and Palaearctic regions [8]. Combined studies of the geometric morphometry of wings, coupled with molecular analysis have excluded *C. chiopterus* and *C. dewulfi* from the *Obsoletus* Group, despite previously being considered as part of the Group, based only on morphological features [9, 10, 14–16]. Life-cycles and trophic behaviors for most of the species in the *Obsoletus* Group are not well described or vary greatly. For instance, ecological niches of *C. chiopterus* and *C. obsoletus* are suspected to be different, although these two species are phylogenetically very close within the subgenus *Avaritia*. *Culicoides obsoletus* is a widespread generalist species and occupies a wide range of larval habitats such as forest litter, silage residue, tree holes or manure [17]. *Culicoides obsoletus* shows opportunistic trophic preferences and is able to take blood meals on various hosts species (man, sheep, goat, cow, horse or rodent) and, occasionally, on birds [18]. *Culicoides chiopterus* is a more specialist species, found engorged almost exclusively on cattle blood with larvae associated with cattle dung [19, 20].

In addition to the group, Meiswinkel et al. [21], defined the *Obsoletus* Complex to close taxa with very similar female adult morphology, including *C. obsoletus*, *C. montanus* and *C. scoticus*. Several authors have recently reported the existence of cryptic diversity within *C. obsoletus*, namely the *C. obsoletus* clade 'O2' and *C. obsoletus* clade 'O3' in Sweden and Switzerland [15], and the *C. obsoletus* clade 'dark' in the Netherlands [21, 22]. We will use the term "Obsoletus/Scoticus Complex" here to refer

to the cryptic species *C. obsoletus*, *C. scoticus* and *C. montanus* as well as all previously described operational taxonomic units in the literature (see above). The terminology “Obsoletus/Scoticus Complex” is written according to the rules defined by Harbach [23] for infrasubgeneric categories within the genus *Anopheles*. Considering that sympatric cryptic species may exhibit different vector competence and may confound epidemiological investigations, it is mandatory to assess the intra- and interspecific diversity within the Obsoletus/Scoticus Complex.

Given the difficulty of separating these species based on morphological identification, various molecular markers have been used to overcome specific identification problems, including 16S ribosomal DNA [24, 25], 28S ribosomal DNA [26], cytochrome oxidase *b* (*cytb*), the internal transcribed spacer region 1 (ITS1) [27] and ITS2 rDNA [28]. However, the DNA region primarily used to infer phylogenetic relationships in this complex has been the mitochondrial cytochrome *c* oxidase subunit 1 (*cox1*) [29]. Most of the diversity within the Obsoletus/Scoticus Complex has been identified using *cox1* [29–33]. Despite the diversity of markers used to characterize the phylogeny of the complex, few studies have used a multi-marker approach [34]. This study, therefore, will integrate a multi-marker approach in order to strengthen genetic reconstruction of the Obsoletus/Scoticus Complex.

The Obsoletus/Scoticus Complex, as currently presented in the literature, is confused and needs taxonomic revision. We characterized and mapped the genetic diversity of the Obsoletus/Scoticus Complex along a Palearctic-Mediterranean transect covering Scandinavia to Canary islands (North to South) and Canary islands to Turkey (West to East). Our main objectives were to identify and describe the cryptic diversity observed within the Obsoletus/Scoticus Complex over a wide geographical area using molecular analyses and to question the taxonomic status of some newly described clades. In order to achieve these objectives, we conducted a molecular analysis which combined multi-marker sequencing, phylogenetic analyses and species delimitation to explore the genetic diversity of the Obsoletus/Scoticus Complex in the western European portion of the Palearctic region.

Methods

Culicoides capture and morphological identification

Biting midges were collected at 68 sites located in 20 countries in the western European portion of the Palearctic region, between 2009 and 2017, using national surveillance networks for *Culicoides* populations or local collections (Additional file 1: Table S1). Collections were made overnight with Onderstepoort Veterinary Institute (OVI) light traps set at farms near horses, cattle or sheep and all insects were stored in 70% ethanol. Morphological

identification to the species level of adult *Culicoides* spp. was performed under a binocular microscope using the available identification keys [35, 36].

DNA extraction, amplification and sequencing

DNA was extracted from a total of 3883 adult females belonging to the Obsoletus/Scoticus Complex using the NucleoSpin® DNA kit RapidLyse (Macherey-Nagel, Duren, Germany), following the manufacturer's instructions. An additional step was added, before extraction, for all individuals (specimens were ground in 50 µl of 1× PBS buffer). DNA samples are available upon request. Fragments of *cox1* were amplified for the 3883 individuals. After sequence cleaning, 3763 sequences of *cox1* were obtained [dataset *cox1* (1)]. Fragments of 16S and 28S rDNA were amplified on individuals chosen to be representative of the entire species diversity resulting from *cox1* [dataset *cox1* (2)] to reinforce mitochondrial gene sequences. All primer sequences as well as the information relating to them are present in Additional file 2: Table S2. PCR's were performed in a 25 µl reaction volume. The PCR mix contained 1× Qiagen buffer, 1 mM MgCl₂, 0.25 mM of each dNTP, 0.2 µM of each primer, 1.25 U Qiagen Taq Polymerase and 0.7 ng/µl genomic DNA for all genes. PCR programs included one-step of 5 cycles before a second step with 35 cycles for 16S rDNA and 28S rDNA. PCR amplification conditions were: an initial denaturation step at 94 °C for 5 min followed by 5 cycles of 94 °C for 30 s; 45 °C for *cox1*, 42 °C for 16S rDNA or 55 °C for 28S rDNA for 40 s; 72 °C for 1 min; 35 cycles of 94 °C for 30 s; 51 °C for *cox1*, 55 °C for 16S rDNA or 50 °C for 28S rDNA for 30 s; 72 °C for 1 min; and a final extension step at 72 °C for 10 min. For each amplification reaction, negative controls were carried out. The PCR products were visualized on 1.5% agarose gels with a GelRed® Nucleic Acid Gel Stain, staining after migration of 90 min at 130 V by electrophoresis for quality control. After purifications, carried out by the sequencing service provider, the remaining 20 µl were sequenced with the same forward primers used for PCR (<https://www.genewiz.com>).

Sequence analyses

The reference sequences of *cox1* used to identify individuals to species are available in Additional file 3: Table S3. A total of 3763 *cox1* sequences (Additional file 4: Table S4) from female adults morphologically identified as belonging to the Obsoletus/Scoticus Complex were obtained after deletion of short and poor-quality sequences. The *cox1* alignment was used to identify all *Culicoides* to species- or clade-level

within the complex using the reference sequences. Among the reference sequences used to specifically assign our *Culicoides* samples, some sequences previously identified as *C. obsoletus* O1 and O3 [15] were included in our analysis in order to cover the diversity of the clades described in the literature within the Obsoletus/Scoticus Complex. After comparison with other sequences, it appears that the sequences named *C. obsoletus* O1 were actually *C. obsoletus* and that *C. obsoletus* O3 belongs to *C. obsoletus* clade dark. For this purpose, *cox1* sequences were aligned with reference sequences [dataset *cox1* (1)]. A phylogenetic tree based on maximum likelihood method allowed designation of a species name to each sequence if the latter belonged to a monophyletic clade with strong support (bootstrap < 900) that included a reference sequence. Thus, *Culicoides* were sequenced for markers 16S rDNA and 28S rDNA to support the phylogenetic reconstruction of the complex. All *cox1* [dataset *cox1* (2)], 16S rDNA and 28S rDNA sequences were independently aligned with the MUSCLE [37] algorithm available in the software GENEIOUS v.6.0.5 (Biomatters, <http://www.geneious.com>). Genetic diversity indices, haplotype and nucleotide diversity were evaluated using DNASP v.5.10 [38]. Alignments with gaps were cleaned using the software GBLOCKS 0.91b [39]. To assess genetic distance between clades and species within, barcoding gap bar chart using R software was performed with *ggplot2* [40] and *ggthemes* packages. Intra- and interspecific genetic differences based on the Kimura 2-Parameter (K2P) distance model [41, 42] were calculated with MEGAX [43]. In order to map the specific diversity of the Obsoletus Group, the R software version 3.6.0 was used with the *Leaflet* version 2.0.2 and *shiny* packages version 1.4.0.

Phylogenetic inferences

Phylogenetic trees were constructed for the three markers using maximum-likelihood (ML) and Bayesian inference (BI). Bayesian inference analyses were conducted on MRBAYES version 3.2.6, with tree sampling every 1000 generations in order to calculate posterior probabilities (PP) and 10 million generations. Optimal sequence evolutionary models for each analysis were obtained with Bayesian information criterion (BIC) using jMODEL-TEST. Maximum-likelihood analyses were conducted on PHYML 3.0. The ML analyses were conducted with the best model selected using 1000 bootstrap replicates for each dataset to investigate the level of support at each node, with starting tree determined by BioNJ analysis. After independent analysis of each gene, alignment of *cox1* and 16S rDNA were concatenated, and analyzed following ML and BI methods.

Species delimitation methods

Two species delimitation methods were applied. The first method was a Bayesian implementation of classical GMYC method, Bayesian General Mixed Yule Coalescent (bGMYC) [44]. The single-locus ultrametric gene trees used for bGMYC methods were created with BEAST 1.8.0 [45] under a strict clock model, a Yule Process Tree Model of speciation, and a random starting tree. This analysis was carried out with default prior distribution, without outgroups and with 10 million generations sampled every 1000 cycles with HKY + G substitution model for 16S rDNA, and with HKY + I for 28S rDNA [46]. The software TREEANNOTATOR v1.8.2 was used to find Ultrametric maximum clade credibility (MCC). Single-threshold GMYC analyses were conducted with *splits* package in R.

The second species delimitation method used was the Bayesian Poisson Tree Processes (bPTP) through the web server PTP (<http://species.h-its.org/ptp/>) [47] with 100,000 MCMC generations and a thinning parameter of 100 on a maximum likelihood phylogenetic tree constructed with *cox1* and 16S rDNA genes concatenated with *C. dewulfi* as outgroup.

Results

Molecular analysis

In total, 3763 sequences were obtained for *cox1*, 95 for 16S rDNA and 95 for 28S rDNA (Table 1). All sequences were deposited in GenBank (Additional file 4: Table S4). No stop codons, insertions or deletions were found in any of the *cox1* sequences, indicating functional mitochondrial products.

Within the selected samples present in our data set, the most abundant species was *C. obsoletus*, with 2416 individuals sampled, representing 68% of all *Culicoides* caught (Fig. 1). Species diversity within the complex varied according to the latitudes of the sampling sites (Fig. 1). The most sampled species in northern Europe (Norway and Finland) was the *C. obsoletus* clade O2 with 62% (162 individuals) and 58% (147 individuals) of this clade, respectively, in each country. For eastern Europe (Latvia, Poland, Serbia, Bulgaria, Macedonia, Greece and Turkey) the most sampled species was *C. obsoletus*, representing, for example, 99% (162 individuals) of all *Culicoides* sampled in Greece. However, a population in Poland (Wronka) appears to be an exception with 100% (35 individuals) of *C. obsoletus* clade O2. Three individuals, belonging to a phylogenetic clade unidentifiable by our reference sequences close to *C. obsoletus* dark, were reported from Latvia. The latter sequences are identical to sequences present in the BOLD database (accession numbers: GMGRC1056-13, GMGRC1000-13, GMGRD2587-13) of *Culicoides* collected in Bavaria,

Table 1 Sequence statistics for four gene fragments used to reconstruct the phylogeny of the Obsoletus/Scoticus complex

Dataset	n	Length (bp)	S	C+G (%)	h	Hd (SD)	π (SD)	Nucleotide model (under BIC)	Implemented model (BI)
cox1 (1)	3763	512–627	141	34.5	228	0.890 (0.003)	0.06299 (0.00473)		
cox1 (2)	95	528–623	146	33.7	77	0.994 (0.003)	0.0921 (0.00489)	TPM2uf+I+G	nst = 6; rates = invgamma
16S rDNA	95	263	51	15.8	13	0.851 (0.02)	0.04791 (0.00446)	HKY+G	nst = 2; rates = gamma
28S rDNA	95	576	25	39.8	24	0.911 (0.015)	0.00753 (0.00099)	HKY+I	nst = 2; rates = propinv
Concatened genes	95	731	197	32.8	90	0.998 (0.002)	0.04594 (0.00297)	TPM2uf+I+G	nst = 6; rates = invgamma

Abbreviations: n, number of individuals; h, number of haplotypes; Hd, haplotype (gene) diversity; π, nucleotide diversity; S, number of polymorphic sites; SD, standard deviation; bp, base pairs; BIC, Bayesian information criterion; BI, Bayesian inference

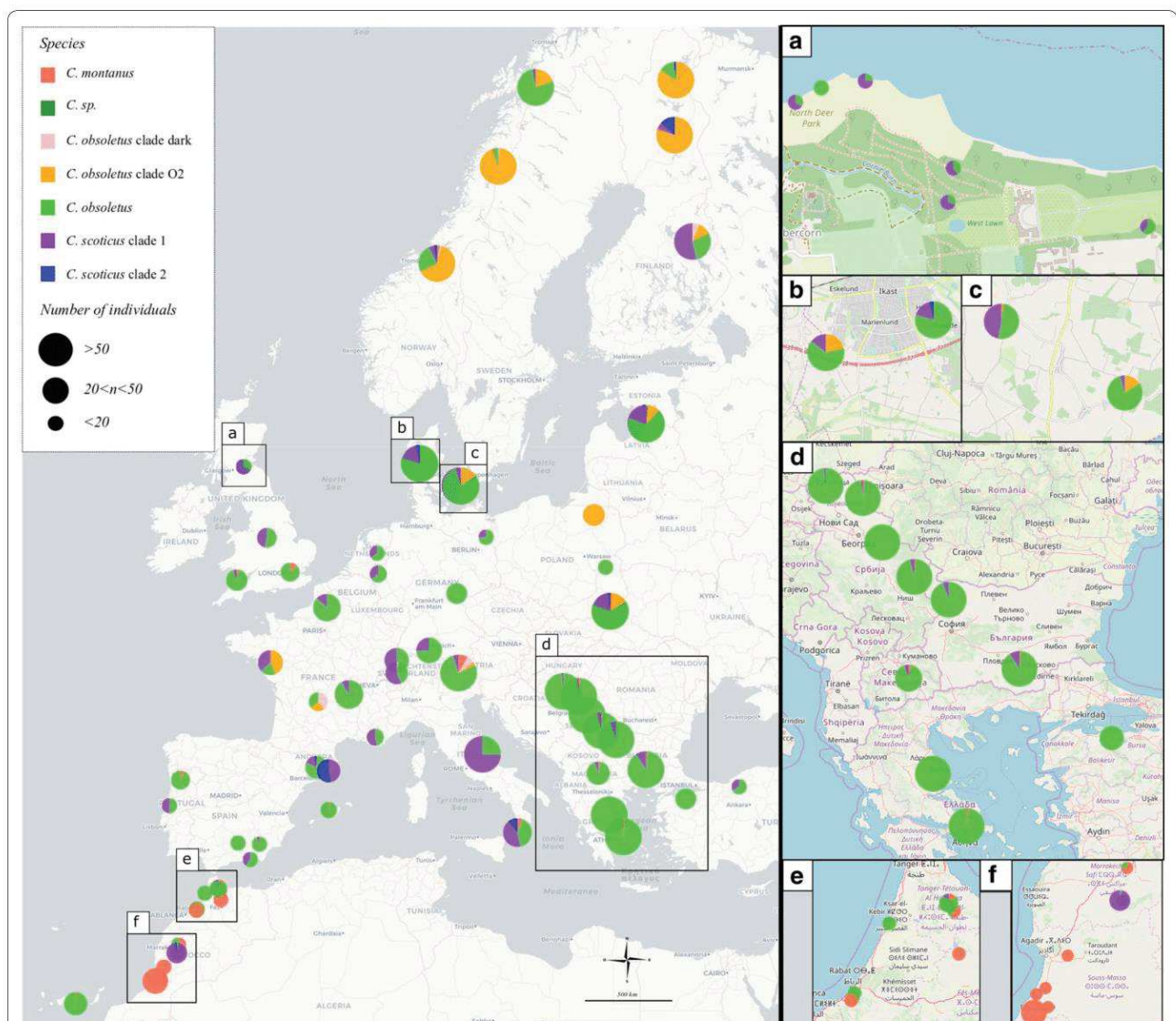


Fig. 1 Map of biting midges sampling sites represented by the number of individuals per clade within the Obsoletus/Scoticus Complex. The different clades identified are shown in different colours. The size of circles on Europe map correspond to the number of individuals per clade within the cryptic species. The numbered maps on the right correspond to magnifications of some study areas. The sites of sampling too close and thus the pie charts this superimposed are to be taken into account in the numbered magnifications and not on the main map. Magnifications: **a** Scotland sample sites; **b, c** Denmark sample sites; **d** Balkans sample sites; **e, f** Morocco sample sites

Germany. Finally, western and central Europe (Portugal, Spain, UK, France, Italy, Netherlands, Germany, Switzerland and Denmark) had the higher species diversity of all the species found in the Palaearctic transect. The most sampled species within this area was *C. obsoletus* with 913 individuals, representing 64% of the *Culicoides* sampled. However, unlike in eastern Europe, *C. scoticus* clade 1 was also found in significant numbers with 386 individuals, or 27% of the samples. *Culicoides obsoletus* clade dark was rarely reported in Europe with only 26 individuals found in France, Denmark, Finland, Italy, Latvia, Norway and Switzerland. *Culicoides montanus* was found in relatively high proportion in Morocco with 54% (80 individuals) of samples, whereas it was much rarer and more sporadic in European countries.

A total of 228 different *cox1* haplotypes were identified. Haplotype diversity varied from 0.829 for *C. obsoletus* clade O2 to 0.517 for *C. scoticus* clade 1 and *C. obsoletus* clade dark (Table 2). In accordance with mitochondrial genes in insects [48], the GC composition ranged from 16% for *16S* rDNA alignment to 39.8% for *28S* rDNA, with a strong AT bias (Table 2). *Culicoides obsoletus* showed very high intraspecific diversity with 94 different haplotypes, representing nearly 43% of the total haplotypic diversity encountered in this study. The maximum interspecific genetic distance (Fig. 2) was reached between *C. dewulfi* and other taxonomic units, with a minimum of 17% of genetic distance between this outgroup and all other members inside the complex. *Culicoides dewulfi* is used here as an outgroup. Genetic distances of the same level as the other intraspecific distances were observed between *C. scoticus* clade 1 and *C. scoticus* clade 2, and between *C. obsoletus* and *C. montanus*, with a maximum of 2% and 4%, respectively. Similar interspecific genetic distances, were observed between all the other clades within the complex, with a minimum of 8% distance. All species had mean intraspecific distances of less than 1%, other than *C. scoticus* clade 2.

Among the 228 unique *cox1* haplotypes, 95 were selected to represent the specific diversity of the complex. Using the same individuals, two alignments of 95 sequences of *16S* rDNA (Additional file 5: Figure S1.) and *28S* rDNA were constructed. In order to compare tree topology and to concatenate markers, a second *cox1* alignment of the 95 sequences was performed (Additional file 6: Figure S2.). The *cox1* dataset had a greater haplotype diversity than *16S* rDNA and *28S* rDNA datasets, with 0.994 vs 0.851 and 0.911 respectively. *28S* rDNA was more monomorphic than *cox1* and *16S* rDNA, with 25, 146 and 51 polymorphic sites, respectively.

Phylogenetic analysis

Information on the alignments used for the construction of phylogenetic trees is provided in Table 1. Species with confirmed taxonomic validity (*C. montanus*, *C. obsoletus* and *C. scoticus*) and cryptic taxa (*C. obsoletus* clade O2, *C. obsoletus* clade dark) highlighted by previous studies were strongly supported (bootstrap > 90%) (Fig. 3). *Culicoides obsoletus* clade O2 constituted a monophyletic clade with strong support (bootstrap > 90%). *Culicoides obsoletus* and *C. montanus* formed a monophyletic clade. *Culicoides scoticus*, another species considered valid, showed two phylogenetic clades, i.e. *C. scoticus* clade 1 and *C. scoticus* clade 2. A monophyletic clade close to *C. obsoletus* clade dark, was strongly supported by bootstrap values. Topologies of phylogenetic trees constructed via maximum likelihood (Fig. 3a) and Bayesian inference (BI) analyses (Fig. 3b) were congruent for alignment of full haplotype diversity.

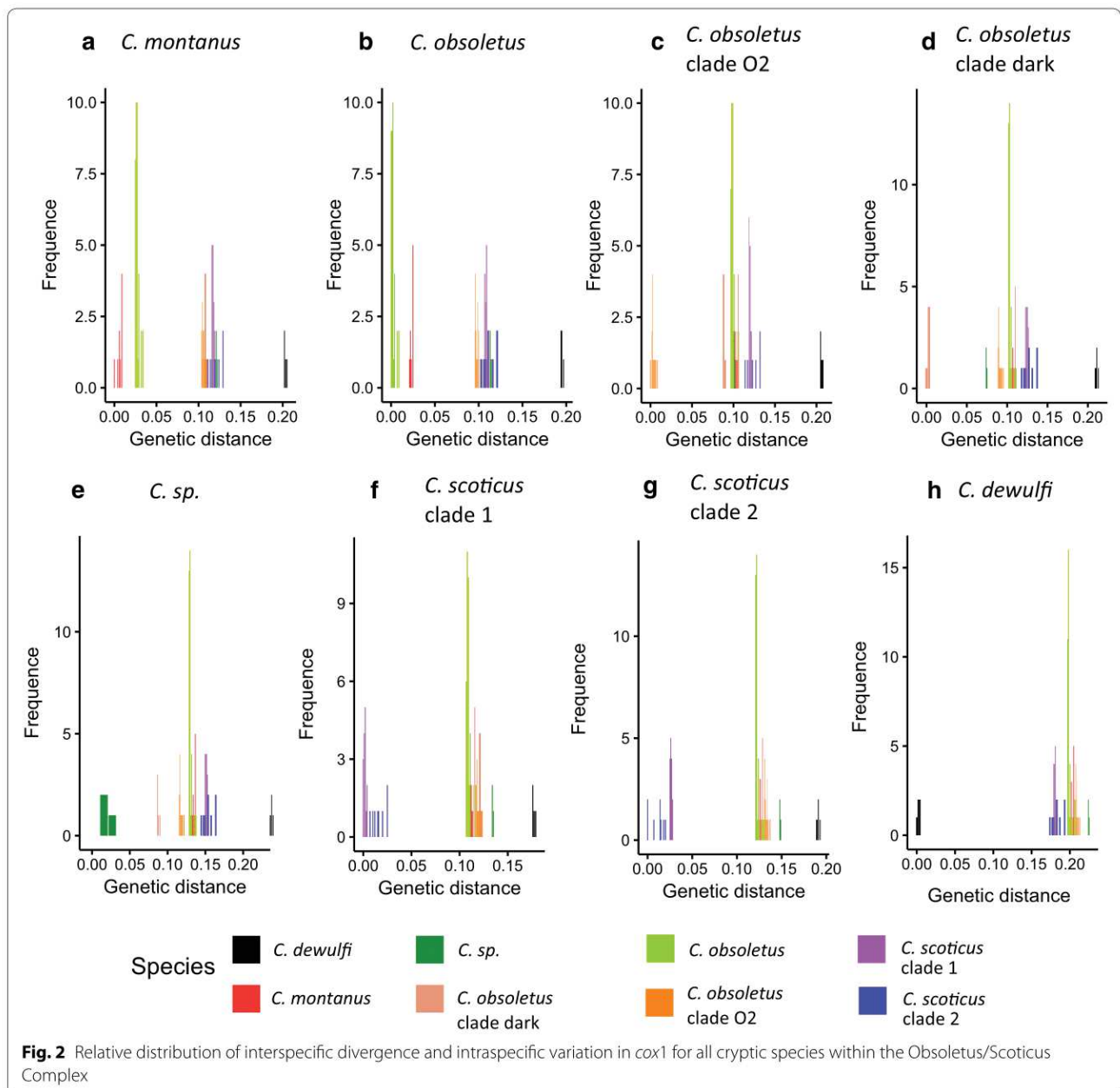
Species delimitation

Using the *cox1* dataset and bGMYC method for species delimitation (Fig. 4), 6 molecular operational taxonomic units (MOTUs) were observed: *C. montanus*, *C. obsoletus*, *C. scoticus*, *C. obsoletus* clade O2, *C. obsoletus* clade dark and *Culicoides* sp. within the Obsoletus/Scoticus Complex. Based on the *16S* rDNA

Table 2 Genetic diversity indices for mitochondrial *cox1* gene segment of *Culicoides* spp. in the Obsoletus/Scoticus Complex

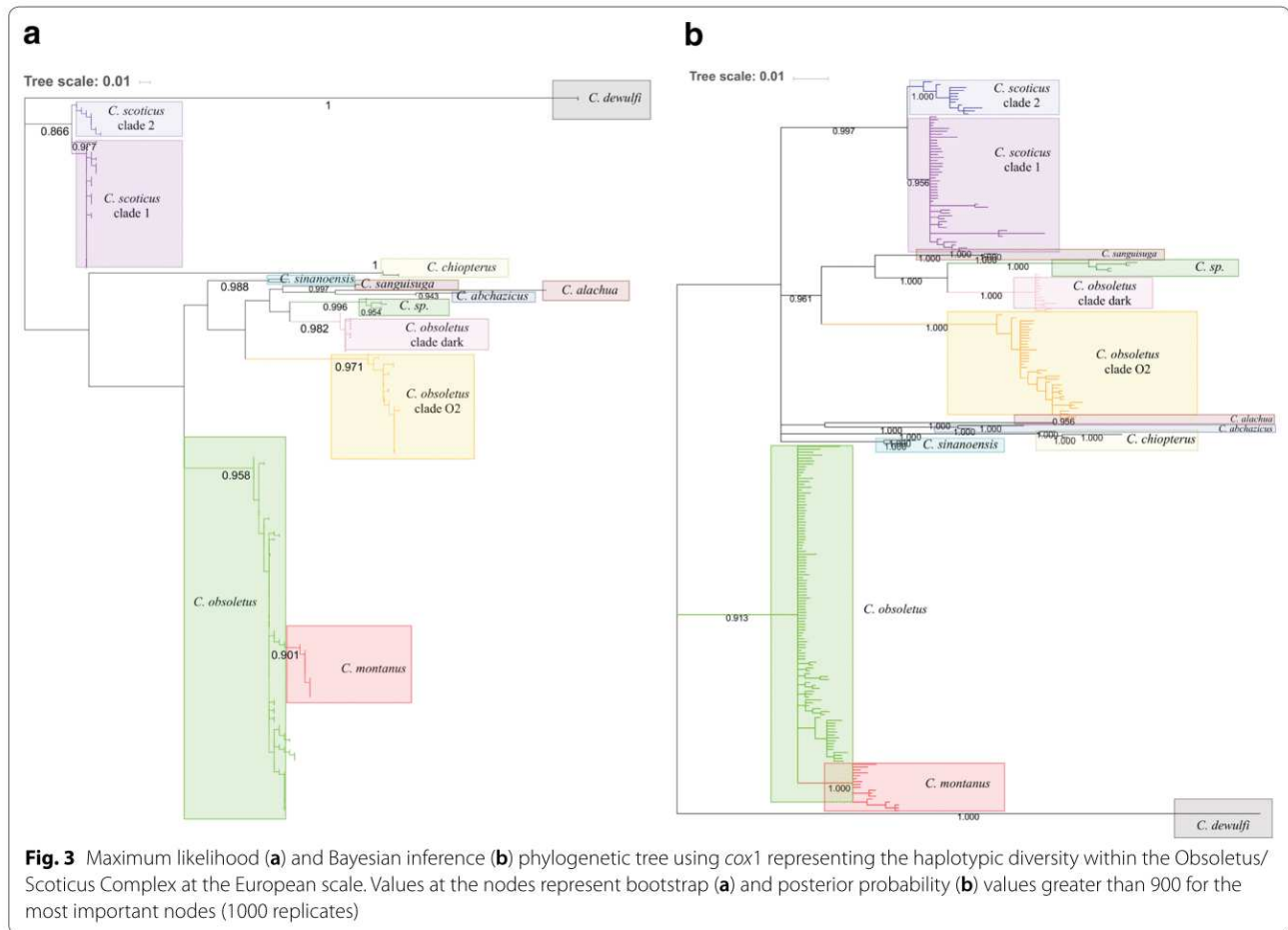
Species	n	h	Hd (SD)	π (SD)	S	C+G (%)
<i>C. montanus</i>	106	13	0.642 (0.039)	0.00726 (0.00317)	66	0.334
<i>C. obsoletus</i>	2416	106	0.773 (0.005)	0.00450 (nd)	72	0.324
<i>C. obsoletus</i> clade O2	512	38	0.829 (0.012)	0.00447 (nd)	53	0.310
<i>C. obsoletus</i> clade dark	26	6	0.517 (0.113)	0.00191 (0.00058)	7	0.315
<i>Culicoides</i> sp.	3	2	0.667 (0.314)	0.00749 (0.00353)	6	0.331
<i>C. scoticus</i> clade 1	645	61	0.562 (0.023)	0.00463 (nd)	93	0.347
<i>C. scoticus</i> clade 2	55	12	0.753 (0.046)	0.00499 (0.00093)	16	0.328

Abbreviations: n, number of individuals; h, number of haplotypes; Hd, haplotype (gene) diversity; π , nucleotide diversity; nd, not determined; S, number of polymorphic sites; SD, standard deviation



dataset analysed with the same delimitation method, three MOTUs were characterised including (i) two clades within *C. scoticus*; (ii) *Culicoides* sp., *C. obsoletus* clade dark and *C. obsoletus* clade O2; and (iii) *C. obsoletus* and *C. montanus*. Species delimitation with the 28S rDNA and 16S rDNA dataset had a lower resolution compared to *cox1*. The 28S rDNA and 16S rDNA datasets were invaluable in identifying cryptic diversity within the Obsolete/Scoticus Complex. Indeed, the tree generated with the 28S rDNA dataset showed very low polymorphisms and resolution signal (Additional

file 7: Figure S3.). The bPTP method was conducted on concatenated dataset of *cox1* and 16S rDNA genes. The best statistical support for species delimitation was for concatenated dataset with bPTP. MOTUs found with the concatenated alignment and *cox1* were the same, except for the delimitation of *C. montanus* and *C. obsoletus*. Indeed, using molecular delineation based on bPTP analysis, MOTUs were observed for *C. scoticus*, *C. obsoletus* clade O2, *C. obsoletus* clade dark, and *Culicoides* sp. Using the bPTP method we were not able to distinguish *C. obsoletus* and *C. montanus* as well as



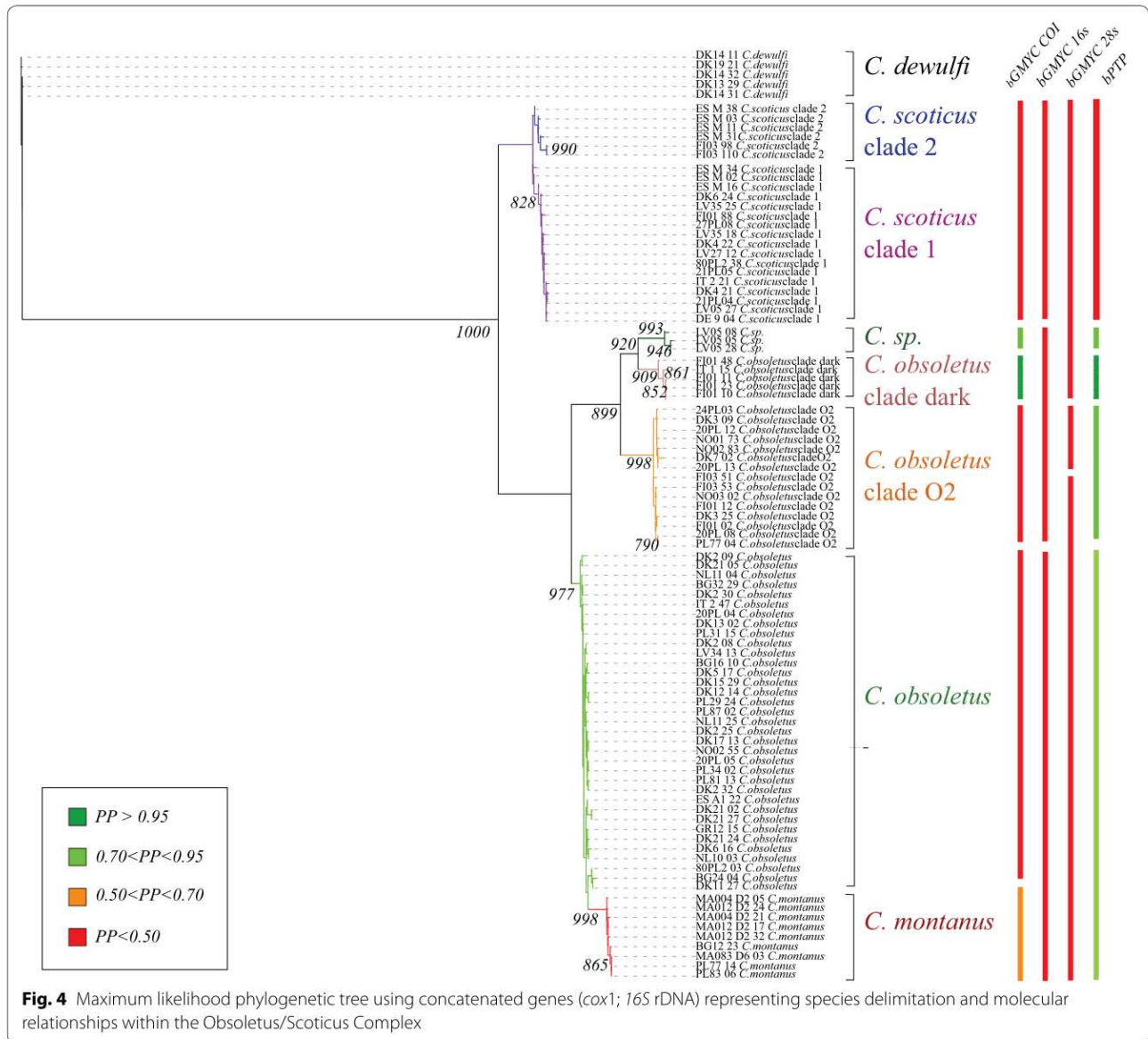
C. scoticus clades 1 and 2, which were not classified as separate species by either method.

Discussion

Our study investigated the genetic diversity of the Obsoletus/Scoticus Complex at the European level (68 sampling sites across 20 Palaearctic countries with 3763 *cox1* sequences). For the first time, the sample selection covers the whole known western Palaearctic distribution area of these species. The complementary use of the mitochondrial 16S rDNA and nuclear 28S rDNA genes confirms the important level of cryptic diversity found within the Obsoletus/Scoticus Complex. Indeed species delimitation methods allowed us to delineate five MOTUs and: (i) to provide evidence of the taxonomic validity of *C. obsoletus* clade O2 and *C. obsoletus* clade dark (ii) to identify individuals belonging to a species not yet described or not present into the databases; and (iii) to question the taxonomic status of *C. montanus*.

Species assignment at the European scale showed variation in the distribution of the cryptic diversity of

Obsoletus/Scoticus Complex. This result confirmed a previous study by Möhlmann et al. [49], who found a strong latitudinal effect on the relative abundance of species of the Obsoletus/Scoticus Complex. However, the previous study was carried out with few individuals from a relatively small number of countries and sampling sites [49]. This contrasts with our study, the first to be conducted at a European scale with a large data set sufficient to provide a more precise idea of the cryptic diversity within the Obsoletus/Scoticus Complex. The latitudinal variation in the relative abundance of the different cryptic species in the Obsoletus/Scoticus Complex, could be due to a wide range of factors like different ecological niches, or differences in the availability of hosts and breeding sites [50]. For example, a study conducted in Italy showed that *C. scoticus* collection sites were dominated by areas of natural vegetation or forest, at medium altitudes, preferably in wilder and more pristine environments [51]. However, the heterogeneity of *Culicoides* collection dates may also explain these variations in specific diversity within the complex. For instance, a study conducted in



Sweden, found a seasonal variation in *Culicoides* community structure [52]. The omnipresence of *C. obsoletus* makes it the dominant species in Europe, confirming its status as a generalist species, which tolerates a wide range of eco-climatic conditions. The dominant species also varies according to geographical location. For example, *C. obsoletus* clade O2 is the most sampled species in Nordic countries while *C. montanus* prevails in Morocco. France, Italy and Spain appear to have the highest specific diversity of *Culicoides* belonging to the Obsoletus/Scoticus Complex. These three countries bring together all the cryptic diversity known so far, except the new, unidentified, *Culicoides* taxon that has only been found in Latvia regarding our dataset. This could be due to the significant

diversity of ecological niches as well as the high density of hosts in these countries and this species diversity variation is in line with general patterns of latitudinal increase in species richness [53].

Our phylogenetic analysis allowed us to define seven well supported phylogenetic clades. Some of them correspond to species with taxonomic validity (*C. montanus*, *C. obsoletus* and *C. scoticus*), some more recently described phylogenetic clades (*C. obsoletus* clade dark and *C. obsoletus* clade O2) [52, 54] and some clades never described before (*Culicoides* sp. and *C. scoticus* clade 2).

Without taking into account *C. scoticus*, the phylogenetic reconstruction produced herein, confirmed the presence of two divergent groups; one consisting of

C. obsoletus and *C. montanus* and the other of *C. obsoletus* clade dark, *C. obsoletus* clade O2 and a clade not yet described in the literature. We were unable to identify these sequences due to the absence of reference sequences identified at the species level in the sequence databases. However, three sequences, from specimens collected in Bavaria in Germany, present in the BOLD database are identical. Further sampling and sequencing of *Culicoides* from eastern Europe are necessary in order to associate morphological features with this new cryptic species. The large number of *Culicoides* processed during this study made it difficult to use non-destructive DNA extraction techniques. Non-destructive techniques would be necessary to couple morphological criteria with genetic analysis, in order to identify this species.

We also described a second clade phylogenetically very close to *C. scoticus*, *C. scoticus* clade 2 [55]. However, given the small genetic distance observed between these two clades, they can be considered as intraspecific variation within *C. scoticus*.

According to this study, *C. obsoletus* clade dark appears as a true cryptic species, with high phylogenetic support. Meiswinkel et al. [21] hypothesized that *C. obsoletus* clade dark could be *C. gornostaevae* Mirzaeva, 1984, but *C. gornostaevae* was significantly larger and had a distribution restricted to the boreal zone of Siberia. However, this species has been recently reported from Norway, Poland and Sweden [56] but the lack of *C. gornostaevae* reference sequences in publicly available databases prevented comparisons with *C. obsoletus* clade dark.

Culicoides obsoletus clade O2 was also strongly supported by phylogenetic and species delimitation analysis, and thus could be considered a cryptic species within the Obsoletus/Scoticus Complex. However, our sample showed a high abundance of *C. obsoletus* clade O2 mainly at high latitudes, whereas it had initially been identified for the first time further south in the Swiss Alps [54, 56, 57] and France.

Little is known about the ecology of *C. montanus*. We found this species over a large geographical area, from Morocco to a few individuals in Norway. In spite of this, we have identified only a very small number of *C. montanus* except in Morocco where this species constitutes the majority. This latter result is in accordance with the fact that sites in Italy, where *C. montanus* is predominant are characterised by a high land surface temperature, higher than the values registered in the *C. obsoletus* and *C. scoticus* sites [51]. *Culicoides obsoletus* and *C. scoticus* are considered to be sibling species because of diagnostic female morphological characters, which are difficult to observe or overlap [16], as opposed to the morphological identification of males which is simpler [58]. If the genetic analyses

by Pages & Sarto [58] confirm that both *C. obsoletus* and *C. scoticus* are distinct species, the question remains open for *C. montanus*. Previous phylogenetic studies based on *cox1* indeed showed limited genetic distance between *C. obsoletus* and *C. montanus* and in studies based on ITS sequences *C. montanus* always appears in one of the subclades of *C. obsoletus* [59]. In our multi-marker phylogenetic tree, *C. montanus* and *C. obsoletus* formed a monophyletic clade. In addition, the genetic distance between *C. obsoletus* and *C. montanus* was of the same order of magnitude as some intraspecific distances.

The number of putative species defined within the Obsoletus/Scoticus Complex varied depending on the molecular markers and species delimitation methods used (Fig. 4), particularly pertaining to the status of *C. obsoletus* and *C. montanus*. Unlike the other methods, the bGMYC method, based on *cox1*, distinguished these two species. This can be explained by the fact that some parameters of the analysis (i.e. priors), like differences in population size or speciation rates, can bias the GMYC method by overestimating the number of species [46, 60–62]. Insufficient sampling, high gene flow or a recent speciation event are also likely explanations for the differences in results between phylogenetic trees and species delimitation results [63, 64]. Although subject to the same constraints, it has been shown that the bPTP method is significantly more robust [65]. Moreover, methods based on multiple loci improve discovery, resolution and stability of species delimitation [66, 67]. Furthermore, studies have shown that species delineation in insects is more appropriate with multilocus species delimitation methods [68–71]. These arguments allowed us to validate the species delimitation scheme produced with the bPTP method. This analysis coupled with the low level of genetic distances observed between *C. montanus* and *C. obsoletus* led us to question the taxonomic status of *C. montanus*. This could be the beginning of a speciation process of *C. montanus* within the clade of *C. obsoletus*. However, *C. montanus* was originally described from Kazakhstan and present in central Asia [57]. Therefore, examination of individuals sampled in this geographical area could strengthen our conclusions.

In the light of our conclusions, in-depth morphological analyses with the deposition of reference individuals will have to be carried out in order to decide whether or not to definitively rule out the taxonomic status of the cryptic species making up this complex. Indeed, although adult females are not morphologically distinguishable on a routine basis, males are easier to identify using their genitalia and pupal differences and can provide evidence of morphological differences between the species.

Conclusions

This study provides clarification of the distribution pattern of species belonging to the *Obsoletus/Scoticus* Complex, using a dataset based on samples from western Palaearctic and Mediterranean transect. Strong variations in latitudinal cryptic species diversity was observed. This study clarifies the phylogenetic relationships between species belonging to the *Obsoletus/Scoticus* Complex. We identified and validated five MOTUs, *C. obsoletus*, *C. scoticus*, *C. obsoletus* clade O2, *C. obsoletus* clade dark and a MOTUs corresponding to an unidentified species. The latter three species have not been formally described but our results confirm that they should be considered as species in their own right. More detailed studies of their morphology and ecology are needed to provide more detailed descriptions of these species. Furthermore, our results raise questions concerning the taxonomic status of *C. montanus*, which was previously considered as a taxonomically valid species.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13071-020-04114-1>.

Additional file 1: Table S1. Information on adult female *Culicoides* sampling sites and results of specific assignment using *cox1*.

Additional file 2: Table S2 Primers used for PCRs and sequencing in this study.

Additional file 3: Table S3. Reference sequences used for specific assignment.

Additional file 4: Table S4. Information on all *Culicoides* sequenced and GenBank accession numbers.

Additional file 5: Figure S1. Maximum likelihood phylogenetic tree using 16S rDNA. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

Additional file 6: Figure S2. Maximum likelihood phylogenetic tree using *cox1*. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

Additional file 7: Figure S3. Maximum likelihood phylogenetic tree using 28S rDNA. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

Abbreviations

BTV: bluetongue virus; SBV: Schmallenberg virus; *cytb*: cytochrome oxidase b; ITS1: internal transcribed spacer region 1; ITS2: internal transcribed spacer region 2; *cox1*: cytochrome c oxidase subunit 1; N: number of individuals; h: number of haplotypes; Hd: haplotype (gene) diversity; π : nucleotide diversity; nd: not determined; S: number of polymorphic sites; bGMYC: Bayesian General Mixed Yule Coalescent; MOTUs: molecular operational taxonomic units; bPTP: Bayesian Poisson Tree Processes.

Acknowledgements

The authors are grateful to all European partners who assisted in *Culicoides* sampling and shipping. We acknowledge Bruno Mathieu for his friendly cooperation and for allowing us access to use some of his sequences as references for our study.

Authors' contributions

AM, KH and CG designed the study. AIJ, AD, AC, BVP, DWR, DS, DW, DPu, DPe, EV, FJ, HK, IPF, JL, JN, JMP, JS, KRS, KK, CLC, ML, MBo, MG, MBi, ME, MR, MO,

MAMC, RB, REP, SC, ST, SB, SS, SMS, ZO, ZS and ZV provided samples of *Culicoides* collected in their respective countries. IR, MD contributed the identification of *Culicoides*. AM, LG, LP and LT performed DNA extraction, amplification, quality control and sequencing. BM provided some sequences as references for our study. AM analysed the data. KH and CG contributed to the manuscript written first by AM. All authors read and approved on the final manuscript.

Funding

This study was partially funded by the EU grant H2020-727393 PALE-Blu and by the VectorNet project (OC/EFSA/AHAW/2013/02-FWC1) funded by the European Centre for Disease Prevention and Control (ECDC) and the European Food Safety Authority (EFSA). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission, of the ECDC or of the EFSA.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its additional files. The newly generated sequences were submitted in the GenBank database under the accession numbers MT170026-MT173788.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ ASTRE, Univ Montpellier, Cirad, INRAE, Montpellier, France. ² Cirad, UMR ASTRE, 34398 Montpellier, France. ³ Cirad, UMR ASTRE, 97170 Petit-Bourg, Guadeloupe, France. ⁴ Institut Agronomique et Vétérinaire Hassan II, Unité Parasitologie et Maladies Parasitaires, 10100 Rabat, Morocco. ⁵ Institute of Parasitology and Tropical Pathology of Strasbourg, Université de Strasbourg, DIHP UR 7292, 67000 Strasbourg, France. ⁶ Centre for Monitoring of Vectors, National Reference Centre, Netherlands Food and Consumer Product Safety Authority, Wageningen, The Netherlands. ⁷ Veterinary Control Central Research Institute, Ankara, Turkey. ⁸ Department of Parasitology and Parasitic Diseases, Faculty of Veterinary Medicine, Ss. Cyril and Methodius University in Skopje, Skopje, Republic of North Macedonia. ⁹ Centre for Ecology, Centre for Ecology & Hydrology, Wallingford OX10 8BB, UK. ¹⁰ CIISA-Centro de Investigação Interdisciplinar em Sanidade Animal, Faculdade de Medicina Veterinária, Universidade de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal. ¹¹ Department of Parasitology-Parasitic Diseases, Entomology & Bee Health, Veterinary Centre of Athens, Athens, Greece. ¹² Leibniz-Centre for Agricultural Landscape Research, Müncheberg, Germany. ¹³ Faculty of Agriculture, University of Novi Sad, Novi Sad, Serbia. ¹⁴ National Centre for Vector Entomology, Institute of Parasitology, University of Zürich, Zürich, Switzerland. ¹⁵ Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Greifswald, Germany. ¹⁶ Department of Animal Pathology, AgriFood Institute of Aragón (IA2) Veterinary Faculty, 50013 Zaragoza, Spain. ¹⁷ Departamento de Microbiología, Laboratorio de Producción y Sanidad Animal de Granada, Junta de Andalucía, Granada, Spain. ¹⁸ Doñana Biological Station, CSIC, Sevilla, Spain. ¹⁹ Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. ²⁰ Centre for Ecology & Hydrology, Edinburgh OX10 8BB, UK. ²¹ Department of Virology, University of Helsinki, Medicum, Haartmaninkatu 3, Helsinki 00014, Finland. ²² National Veterinary Research Institute, Pulawy, Poland. ²³ Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise 'G. Caporale', Campo Boario, 64100 Teramo, Italy. ²⁴ The Pirbright Institute, Pirbright, UK. ²⁵ Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, Leahurst, Chester High Road, Neston, Cheshire CH64 7TE, UK. ²⁶ Applied Zoology and Animal Conservation Research Group, University of the Balearic Islands UIB, Palma, Spain. ²⁷ University of Copenhagen, Copenhagen, Denmark. ²⁸ National Diagnostic and Research Veterinary Medical Institute, Sofia, Bulgaria. ²⁹ Norwegian Veterinary Institute, Oslo, Norway. ³⁰ Institute of Food safety, Animal Health and Environment 'BIOR', Riga, Latvia.

Received: 13 January 2020 Accepted: 29 April 2020
Published online: 20 May 2020

References

- Carpenter S, Groschup MH, Garros C, Felipe-Bauer ML, Purse BV. *Culicoides* biting midges, arboviruses and public health in Europe. *Antiviral Res.* 2013;100:102–13.
- Wilson AJ, Mellor PS. Bluetongue in Europe: past, present and future. *Phil Trans R Soc Lond B.* 2009;364:2669–81.
- Carpenter S, Wilson A, Mellor PS. *Culicoides* and the emergence of bluetongue virus in northern Europe. *Trends Microbiol.* 2009;17:172–8.
- Hoffmann B, Eschbaumer M, Beer M. Real-time quantitative reverse transcription-PCR assays specifically detecting bluetongue virus serotypes 1, 6, and 8. *J Clin Microbiol.* 2009;47:2992–4.
- Savini G, MacLachlan NJ, Sanchez-Vizcaino JM, Zientara S. Vaccines against bluetongue in Europe. *Comp Immunol Microbiol Infect Dis.* 2008;31:101–20.
- Carpenter S, Lunt HL, Arav D, Venter GJ, Mellor PS. Oral susceptibility to bluetongue virus of *Culicoides* (Diptera: Ceratopogonidae) from the United Kingdom. *J Med Entomol.* 2006;43:73–8.
- Carpenter S, McArthur C, Selby R, Ward R, Nolan DV, Luntz AJ, et al. Experimental infection studies of UK *Culicoides* species midges with bluetongue virus serotypes 8 and 9. *Vet Rec.* 2008;163:589–92.
- Mathieu B. Les espèces de *Culicoides* du sous-genre *Avaritia* (Diptera: Ceratopogonidae) dans le monde: révision systématique et taxonomique des espèces d'intérêt dans la transmission d'Orbivirus. *École doctorale Sciences de la vie et de la santé Université de Strasbourg*; 2011.
- Meiswinkel R, Gomulski LM, Delecolle JC, Goffredo M, Gasperi G. The taxonomy of *Culicoides* vector complexes—unfinished business. *Vet Ital.* 2004;40:151–9.
- Hajd Henni L, Sauvage F, Ninio C, Depaquit J, Augot D. Wing geometry as a tool for discrimination of *Obsoletus* group (Diptera: Ceratopogonidae: *Culicoides*) in France. *Infect Genet Evol.* 2014;21:110–7.
- Mehlhorn H, Walldorf V, Klimpel S, Jahn B, Jaeger F, Eschweiler J, et al. First occurrence of *Culicoides obsoletus*-transmitted bluetongue virus epidemic in central Europe. *Parasitol Res.* 2007;101:219–28.
- Savini G, Goffredo M, Monaco F, Di Gennaro A, Cafiero MA, Baldi L, et al. Bluetongue virus isolations from midges belonging to the *Obsoletus* complex (*Culicoides*, Diptera: Ceratopogonidae) in Italy. *Vet Rec.* 2005;157:133–9.
- Borkent A. World species of biting midges (Diptera: Ceratopogonidae). *Salmon Arm: American Museum of Natural History, and Instituto Nacional de Biodiversidad*; 2016.
- Gomulski LM, Meiswinkel R, Delecolle JC, Goffredo M, Gasperi G. Phylogenetic relationships of the subgenus *Avaritia* Fox, 1955 including *Culicoides obsoletus* (Diptera, Ceratopogonidae) in Italy based on internal transcribed spacer 2 ribosomal DNA sequences. *Syst Entomol.* 2005;30:619–31.
- Ander M, Troell K, Chirico J. Barcoding of biting midges in the genus *Culicoides*: a tool for species determination. *Med Vet Entomol.* 2013;27:323–31.
- Kluiters G, Pages N, Carpenter S, Gardes L, Guis H, Baylis M, et al. Morphometric discrimination of two sympatric sibling species in the Palaearctic region, *Culicoides obsoletus* Meigen and *C. scoticus* Downes & Kettle (Diptera: Ceratopogonidae), vectors of bluetongue and Schmallenberg viruses. *Parasit Vectors.* 2016;9:262.
- Ninio C, Augot D, Delecolle JC, Dufour B, Depaquit J. Contribution to the knowledge of *Culicoides* (Diptera: Ceratopogonidae) host preferences in France. *Parasitol Res.* 2011;108:657–63.
- Martinez-de la Puente J, Figuerola J, Soriguer R. Fur or feather? Feeding preferences of species of *Culicoides* biting midges in Europe. *Trends Parasitol.* 2015;31:16–22.
- Garros C, Gardes L, Allene X, Rakotoarivony I, Viennet E, Rossi S, et al. Adaptation of a species-specific multiplex PCR assay for the identification of blood meal source in *Culicoides* (Ceratopogonidae: Diptera): applications on Palaearctic biting midge species, vectors of Orbiviruses. *Infect Genet Evol.* 2011;11:1103–10.
- Kettle DS, Lawson JWH. The early stages of British biting midges *Culicoides* Latreille (Diptera: Ceratopogonidae) and allied Genera. *Bull Entomol Res.* 2009;43:421–67.
- Meiswinkel R, De Bree F, Bossers-De Vries R, Elbers AR. An unrecognized species of the *Culicoides obsoletus* complex feeding on livestock in The Netherlands. *Vet Parasitol.* 2015;207:324–8.
- Elbers AR, Meiswinkel R. *Culicoides* (Diptera: Ceratopogonidae) host preferences and biting rates in the Netherlands: comparing cattle, sheep and the black-light suction trap. *Vet Parasitol.* 2014;205:330–7.
- Harbach RE. Review of the internal classification of the genus *Anopheles* (Diptera, Culicidae) - the foundation for comparative systematics and phylogenetic research. *Bull Entomol Res.* 1994;84:331–42.
- Debeila TJ. Characterisation of selected *Culicoides* population in South Africa using genetic markers. PhD Thesis, Faculty of Veterinary Science, University of Pretoria; 2010.
- Tabachnick W. *Culicoides* and the global epidemiology of bluetongue virus infection. *Vet Ital.* 2004;40:144–50.
- Bakhomou MT, Labuschagne K, Huber K, Fall M, Mathieu B, Venter G, et al. Phylogenetic relationships and molecular delimitation of *Culicoides* Latreille (Diptera: Ceratopogonidae) species in the Afrotropical region: interest for the subgenus *Avaritia*. *Syst Entomol.* 2018;43:355–71.
- Cetre-Sossah C, Baldet T, Delecolle JC, Mathieu B, Perrin A, Grillet C, et al. Molecular detection of *Culicoides* spp. and *Culicoides imicola*, the principal vector of bluetongue (BT) and African horse sickness (AHS) in Africa and Europe. *Vet Res.* 2004;35:325–37.
- Gomulski LM, Meiswinkel R, Delecolle JC, Goffredo M, Gasperi G. Phylogeny of the subgenus *Culicoides* and related species in Italy, inferred from internal transcribed spacer 2 ribosomal DNA sequences. *Med Vet Entomol.* 2006;20:229–38.
- Nolan DV, Carpenter S, Barber J, Mellor PS, Dallas JF, Mordue Luntz AJ, et al. Rapid diagnostic PCR assays for members of the *Culicoides obsoletus* and *Culicoides pulicaris* species complexes, implicated vectors of bluetongue virus in Europe. *Vet Microbiol.* 2007;124:82–94.
- Augot D, Sauvage F, Jouet D, Simphal E, Veuille M, Couloux A, et al. Discrimination of *Culicoides obsoletus* and *Culicoides scoticus*, potential bluetongue vectors, by morphometrical and mitochondrial cytochrome oxidase subunit I analysis. *Infect Genet Evol.* 2010;10:629–37.
- Pages N, Munoz-Munoz F, Talavera S, Sarto V, Lorca C, Nunez JI. Identification of cryptic species of *Culicoides* (Diptera: Ceratopogonidae) in the subgenus *Culicoides* and development of species-specific PCR assays based on barcode regions. *Vet Parasitol.* 2009;165:298–310.
- Schwenkenbecher JM, Mordue AJ, Piertney SB. Phylogenetic analysis indicates that *Culicoides dewulfi* should not be considered part of the *Culicoides obsoletus* complex. *Bull Entomol Res.* 2009;99:371–5.
- Sarto i Monteys V, Ventura D, Pages N, Aranda C, Escosa R. Expansion of *Culicoides imicola*, the main bluetongue virus vector in Europe, into Catalonia, Spain. *Vet Rec.* 2005;156:415–7.
- Harrup LE, Bellis GA, Balenghien T, Garros C. *Culicoides* Latreille (Diptera: Ceratopogonidae) taxonomy: current challenges and future directions. *Infect Genet Evol.* 2015;30:249–66.
- Delecolle JC. Nouvelle contribution à l'étude systématique et iconographique des espèces du genre *Culicoides* (Diptera: Ceratopogonidae) du Nord-Est de la France. PhD Thesis, University of Strasbourg, Strasbourg; 1985.
- Campbell JA, Pelham-Clinton EC. Taxonomic review of the British species of *Culicoides* Latreille (Diptera, Ceratopogonidae). *Proc R Entomol Soc.* 1960;67:181–302.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792–7.
- Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 2009;25:1451–2.
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000;17:540–52.
- Wickham H. ggplot2: Elegant graphics for data analysis. *J R Stat Soc A Stat.* 2009;174:245–50.
- Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol.* 1980;16:111–20.
- Nei M, Rooney AP. Concerted and birth-and-death evolution of multi-gene families. *Annu Rev Genet.* 2005;39:121–52.

43. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35:1547–9.
44. Reid NM, Carstens BC. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol Biol*. 2012;12:196.
45. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29:1969–73.
46. Fujisawa T, Barraclough TG. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst Biol*. 2013;62:707–24.
47. Zhang J, Kapli P, Pavlidis P, Stamatakis A. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*. 2013;29:2869–76.
48. Crozier RH, Crozier YC. The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics*. 1993;133:97–117.
49. Möhlmann TWR, Bekendam AM, van Kemenade I, Wennergren U, Favia G, Takken W, et al. Latitudinal diversity of biting midge species within the *Obsoletus* group across three habitats in Europe. *Med Vet Entomol*. 2019;33:420–6.
50. Goffredo M, Meiswinkel R, Federici V, Di Nicola F, Mancini G, Ippoliti C, et al. The '*Culicoides obsoletus* group' in Italy: relative abundance, geographic range, and role as vector for bluetongue virus. *Vet Ital*. 2016;52:235–41.
51. Ippoliti C, Goffredo M, Salini R, Bocci M, Quaglia M, Federici V, et al. Ecological indicators for three *Culicoides* species of the *Obsoletus* complex in Italy. *Vet Ital*. 2016;52:213–22.
52. Ander M, Meiswinkel R, Chirico J. Seasonal dynamics of biting midges (Diptera: Ceratopogonidae: *Culicoides*), the potential vectors of bluetongue virus, in Sweden. *Vet Parasitol*. 2012;184:59–67.
53. Allen AP, Gillooly JF. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecol Lett*. 2006;9:947–54.
54. Wenk CE, Kaufmann C, Schaffner F, Mathis A. Molecular characterization of Swiss Ceratopogonidae (Diptera) and evaluation of real-time PCR assays for the identification of *Culicoides* biting midges. *Vet Parasitol*. 2012;184:258–66.
55. Augot D, Mathieu B, Hadj-Henni L, Barriol V, Zapata Mena S, Smolis S, et al. Molecular phylogeny of 42 species of *Culicoides* (Diptera, Ceratopogonidae) from three continents. *Parasite*. 2017;24:23.
56. Kirkeby C, Dominiak P. *Culicoides (Avaritia) gornostaevae* Mirzaeva, 1984 (Diptera: Ceratopogonidae)—a possible vector species of the *Obsoletus* group new to the European fauna. *Parasit Vectors*. 2014;7:445.
57. Mathieu B, Delecolle JC, Garros C, Balenghien T, Setier-Rio ML, Candolfi E, et al. Simultaneous quantification of the relative abundance of species complex members: application to *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae), potential vectors of bluetongue virus. *Vet Parasitol*. 2011;182:297–306.
58. Pages N, Sarto IVM. Differentiation of *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae) based on mitochondrial cytochrome oxidase subunit I. *J Med Entomol*. 2005;42:1026–34.
59. Kiehl E, Walldorf V, Klimpel S, Al-Quraishy S, Mehlhorn H. The European vectors of bluetongue virus: are there species complexes, single species or races in *Culicoides obsoletus* and *C. pulicaris* detectable by sequencing ITS-1, ITS-2 and 18S-rDNA? *Parasitol Res*. 2009;105:331–6.
60. Esselstyn JA, Evans BJ, Sedlock JL, Anwarali Khan FA, Heaney LR. Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc R Entomol Soc*. 2012;279:3678–86.
61. Talavera G, Dinca V, Vila R. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods Ecol Evol*. 2013;4:1101–10.
62. Dellicour S, Flot JF. Delimiting species-poor data sets using single molecular markers: a study of barcode gaps, haplowebs and GMYC. *Syst Biol*. 2015;64:900–8.
63. Rosenberg NA. The probability of topological concordance of gene trees and species trees. *Theor Popul Biol*. 2002;61:225–47.
64. Nichols R. Gene trees and species trees are not the same. *Trends Ecol Evol*. 2001;16:358–64.
65. Pentinsaari M, Vos R, Mutanen M. Algorithmic single-locus species delimitation: effects of sampling effort, variation and nonmonophyly in four methods and 1870 species of beetles. *Mol Ecol Resour*. 2017;17:393–404.
66. Fujita MK, Leache AD, Burbrink FT, McGuire JA, Moritz C. Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol Evol*. 2012;27:480–8.
67. Leache AD, Fujita MK. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc R Entomol Soc*. 2010;277:3071–7.
68. Boykin LM, Schutze MK, Krosch MN, Chomič A, Chapman TA, Englezou A, et al. Multi-gene phylogenetic analysis of south-east Asian pest members of the *Bactrocera dorsalis* species complex (Diptera: Tephritidae) does not support current taxonomy. *J Appl Entomol*. 2014;138:235–53.
69. Dinca V, Lukhtanov VA, Talavera G, Vila R. Unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. *Nat Commun*. 2011;2:324.
70. Hsieh CH, Ko CC, Chung CH, Wang HY. Multilocus approach to clarify species status and the divergence history of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. *Mol Phylogenet Evol*. 2014;76:172–80.
71. Song J-H, Ahn K-J. Species delimitation in the *Aleochara fucicola* species complex (Coleoptera: Staphylinidae: Aleocharinae) and its phylogenetic relationships. *Zool Scr*. 2014;43:629–40.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Supplementary information

Tableau S1. Information on adult female *Culicoides* sampling sites and results of specific assignation using cox1.

Locations names	Country	Latitude	Longitude	Year of collection	Number of <i>Culicoides obsoletus</i>	Number of <i>Culicoides montanus</i>	Number of <i>Culicoides obsoletus</i> clade O2	Number of <i>Culicoides scoticus</i>	Number of <i>Culicoides obsoletus</i> clade dark	Number of <i>Culicoides scoticus</i> clade 2	Number of <i>Culicoides</i> sp.
Montana	Bulgaria	43,41	23,23	2016	146	0	0	8	0	0	0
Nikolovo	Bulgaria	43,87	26,12	2016	146	2	0	15	0	0	0
Dyssevej	Denmark	55,20	12,11	2015	110	0	21	6	1	0	0
Groddevej	Denmark	56,13	9,18	2015	43	0	0	10	0	1	0
Lægdsgårdsvej	Denmark	56,12	9,12	2015	89	0	30	20	0	0	0
Leestrupvej	Denmark	55,22	12,05	2015	91	0	4	85	0	0	0
Inari	Finland	68,67	27,56	2015	8	0	49	0	0	1	0
Pelkoseniemi	Finland	67,18	27,46	2015	0	0	90	6	0	15	0
Sotkamo	Finland	64,04	28,72	2015	19	0	8	36	4	0	0
Barjols	France	43,54	5,98	2016	13	0	0	14	0	0	0
Bonnay	France	49,93	2,51	2016	39	0	0	6	0	0	0
Les Hitres	France	45,75	4,07	2016	43	0	0	4	0	0	0
Reygnac	France	45,34	1,82	2016	11	0	7	0	12	0	0
Sucé-sur-Erdre	France	47,32	-1,54	2016	8	0	17	15	0	0	0
Peetzig	Germany	53,08	13,94	2016	8	0	0	2	0	0	0
Schleiz	Germany	50,57	11,81	2016	33	0	0	0	0	0	0
Ziegelbach	Germany	47,89	9,86	2016	31	0	0	11	0	0	0
Abercorn	Great britain	56,00	-3,46	2014	11	0	0	6	0	0	0
Abercorn	Great britain	51,18	-3,47	2014	5	0	0	10	0	0	0
Abercorn	Great britain	56,00	-3,48	2014	1	0	0	2	0	0	0
Abercorn	Great britain	51,18	-3,47	2014	4	0	0	0	0	0	0
Abercorn	Great britain	56,00	-3,47	2014	1	0	0	1	0	0	0
Abercorn	Great britain	56,00	3,47	2014	2	0	0	2	0	0	0
Blackmoor	Great britain	51,18	-3,94	2015	30	1	0	2	0	0	0
London	Great britain	51,51	-0,13	2010	27	4	0	0	0	0	0
Thorpe	Great britain	53,04	-1,78	2015	17	0	0	15	0	0	0
Varybobi	Greece	38,13	23,78	2016	50	1	0	0	0	0	0
Velestino	Greece	39,38	22,75	2016	112	0	0	0	0	0	0
Brunico	Italy	46,80	11,96	2016	46	5	2	3	4	0	0
Mammola	Italy	38,38	16,19	2016	19	3	0	20	0	5	0
Teramo	Italy	42,64	13,67	2016	14	0	0	41	0	0	0
Kalnciemi	Latvia	57,66	25,44	2016	165	0	26	46	0	0	3
Mazlimbas	Latvia	57,71	25,40	2016	75	0	16	65	2	4	0
Oubogo	Macedonia	41,63	21,99	2016	34	1	0	2	0	0	0
Ait Oumribet	Morocco	29,54	-9,59	2016	0	2	0	0	0	0	0
Bab Taza	Morocco	31,26	-5,19	2016	1	2	0	1	0	0	0
Bounaamane	Morocco	29,54	-9,80	2016	0	43	0	0	0	0	0
Chefchaouen	Morocco	35,16	-5,27	2016	18	4	0	2	0	0	0
Dardara	Morocco	34,49	-5,30	2016	2	0	0	0	0	0	0
Moulay Boussehham	Morocco	31,26	-6,25	2016	1	0	0	0	0	0	0
Ouijjane	Morocco	29,62	-9,54	2016	0	19	0	0	0	0	0
Oulad Mtaa	Morocco	31,26	-8,26	2016	1	0	0	32	0	1	0
Ourtzagh	Morocco	34,49	-5,11	2016	0	3	0	0	0	0	0
Rabat	Morocco	33,96	-6,80	2010	4	1	0	0	0	0	0
Sidi Badhaj	Morocco	31,27	-8,22	2016	0	0	0	1	0	0	0
Tizgzaouine	Morocco	31,77	-8,13	2016	1	3	0	0	0	0	0
Tnine Aglou	Morocco	30,42	-9,76	2016	0	1	0	0	0	0	0
Lomm	Netherlands	51,45	6,20	2017	18	0	0	10	0	0	0
Olst	Netherlands	52,37	6,12	2017	11	0	0	6	0	0	0
Narvik	Norway	68,48	17,51	2015	82	0	21	3	0	0	0
Storforshei	Norway	66,31	14,81	2015	4	0	91	0	1	0	0
Tiller	Norway	63,34	10,42	2015	19	2	50	4	1	2	0

Jaksamanice	Poland	49,75	22,86	2016	93	2	22	26	0	3	0
Olszewnica	Poland	51,76	22,51	2016	25	0	0	0	0	0	0
Pikulice	Poland	49,75	22,79	2016	195	1	23	29	0	2	0
Wronka	Poland	54,00	21,68	2016	0	0	35	0	0	0	0
Aboadela	Portugal	41,27	-8,00	2009	26	2	0	0	0	0	0
Bajouca	Portugal	39,87	-8,77	2009	1	0	0	1	0	0	0
Aleksandrovo	Serbia	45,63	20,58	2016	61	1	0	1	0	0	0
Cirikovac	Serbia	44,67	21,19	2016	78	0	0	0	0	0	0
Gamzigrad	Serbia	43,91	22,18	2016	71	1	0	3	0	0	0
Pacir	Serbia	45,89	19,45	2016	59	0	0	1	0	0	0
Araaut	Spain	41,97	1,76	2017	30	0	0	6	0	1	0
Arucas	Spain	28,13	-15,50	2010	37	1	0	0	0	0	0
Castril	Spain	37,73	-2,34	2017	16	0	0	1	0	0	0
Massanes	Spain	41,77	2,67	2017	0	0	0	17	0	20	0
Orgiva	Spain	36,85	-2,99	2017	6	0	0	4	0	0	0
Palma	Spain	39,56	2,65	2015	26	1	0	0	0	0	0
Torre del Campo	Spain	37,83	-3,90	2017	13	0	0	0	0	0	0
Dittingen	Switzerland	47,44	7,50	2009	18	0	0	22	0	0	0
Muhleturnen	Switzerland	46,81	7,51	2009	15	0	0	21	1	0	0
Kuzgöl	Turkey	40,91	32,10	2016	2	0	0	1	0	0	0
Şahinköy	Turkey	40,24	28,24	2016	31	0	0	0	0	0	0

Tableau S2. Primers used for PCRs and sequencing in this study.

Target Marker	Forward primer (5' to 3')	Reverse Primer (5' to 3')	Length of amplified fragment (bp)	References
Cox1	LCO1490: GGTCAACAAATCATAAAGATATTGG	HCO2198: TAAACTTCAGGGTGACCAAAAAATCA	710	Folmer et al. 1994
16S rDNA	16S-EK_F: CACGTAAGAAGACTAAATAGTCGAAC	16S-EK_R: GACCGTGCAAAGGTAGCATAATC	450	Ekrem et al. 2010
28S rDNA	28S_S3660: GAG AGT TMA ASA GTA CGT GAA AC	28S_A335: TCG GAA GGA ACC AGC TAC TA	657	Dowton & Austin 1998 Whiting et al. 1997

Tableau S3. Reference sequences used for specific assignation.

Species	Accession number	Origin
<i>Culicoides abchazicus</i>	MK760081	GenBank
<i>Culicoides abchazicus</i>	MK760082	GenBank
<i>Culicoides alachua</i>	MK760088	GenBank
<i>Culicoides chiopterus</i>	KT278212	GenBank
<i>Culicoides chiopterus</i>	KM941505	GenBank
<i>Culicoides chiopterus</i>	KM940185	GenBank
<i>Culicoides dewulfi</i>	KJ162977	GenBank
<i>Culicoides montanus</i>	MK760206	GenBank
<i>Culicoides montanus</i>	MK760202	GenBank
<i>Culicoides montanus</i>	MK760207	GenBank
<i>Culicoides montanus</i>	MK760205	GenBank
<i>Culicoides obsoletus</i>	DQ162811	GenBank
<i>Culicoides obsoletus</i>	HM022839	GenBank
<i>Culicoides obsoletus</i>	HM022842	GenBank
<i>Culicoides obsoletus</i>	HM022856	GenBank
<i>Culicoides obsoletus</i>	HQ824374	GenBank
<i>Culicoides obsoletus</i>	KJ624104	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057928	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057935	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057940	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057929	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057930	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057932	GenBank
<i>Culicoides obsoletus</i> clade O1	HQ824371	GenBank
<i>Culicoides obsoletus</i> clade O1	HQ824378	GenBank
<i>Culicoides obsoletus</i> clade O1	JQ620132	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824381	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824382	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824383	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824384	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ620144	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ978454	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ978459	GenBank
<i>Culicoides sanguisuga</i>	MK760237	GenBank
<i>Culicoides sanguisuga</i>	MK760238	GenBank
<i>Culicoides scoticus</i>	HM022865	GenBank
<i>Culicoides scoticus</i>	HM022875	GenBank
<i>Culicoides scoticus</i>	KJ624126	GenBank
<i>Culicoides scoticus</i>	KT186881	GenBank
<i>Culicoides sinanoensis</i>	MK760244	GenBank
<i>Culicoides sinanoensis</i>	MK760242	GenBank
<i>Culicoides</i> sp.	GMGRC1056-13	BOLD Systems
<i>Culicoides</i> sp.	GMGRC1000-13	BOLD Systems
<i>Culicoides</i> sp.	GMGRD2587-13	BOLD Systems

Tableau S4. Information on all *Culicoides* sequenced and GenBank accession numbers.

The table being too voluminous, it is not included in this manuscript.

DOI : <https://doi.org/10.1186/s13071-020-04114-1>

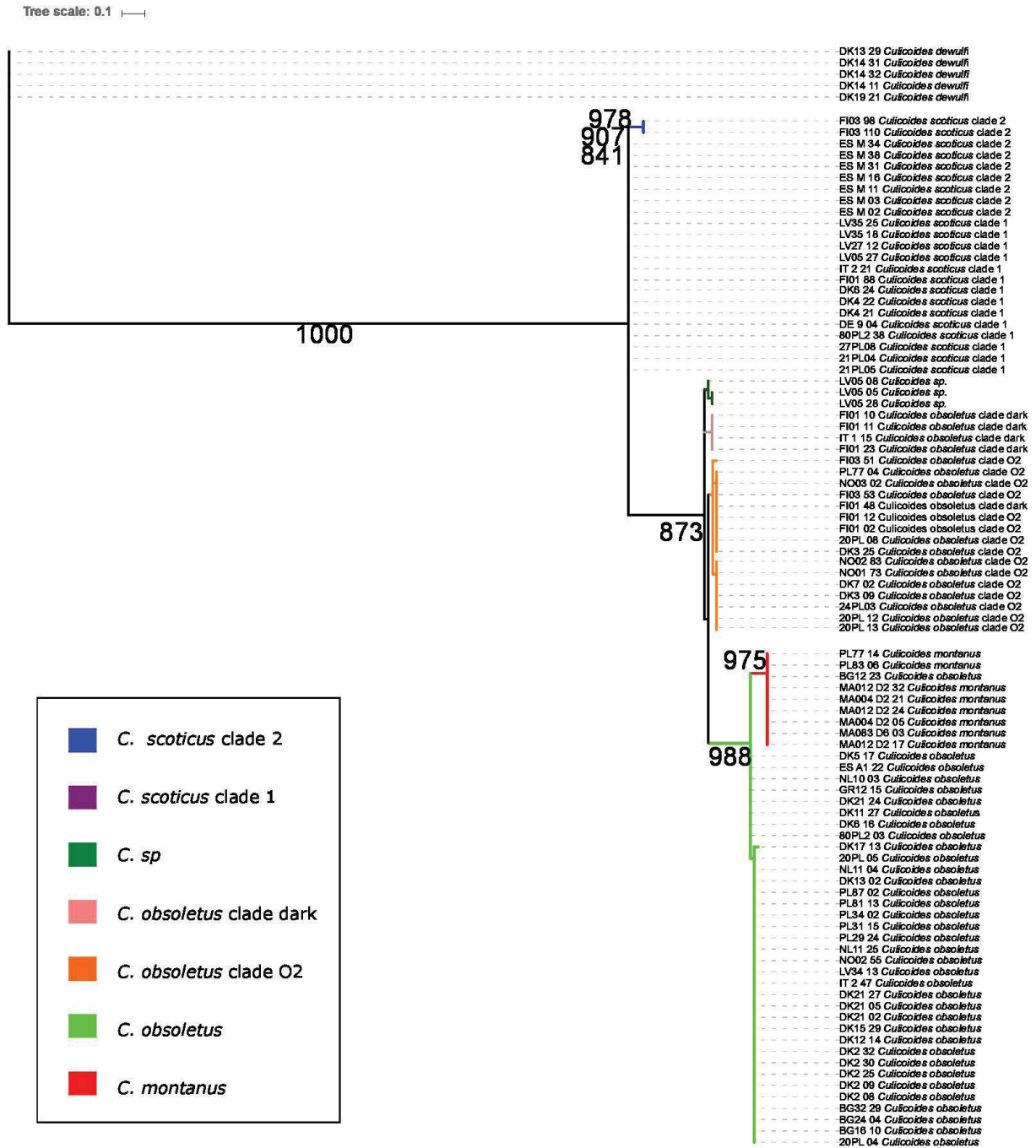


Figure S1. Maximum likelihood phylogenetic tree using *16S* rDNA. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

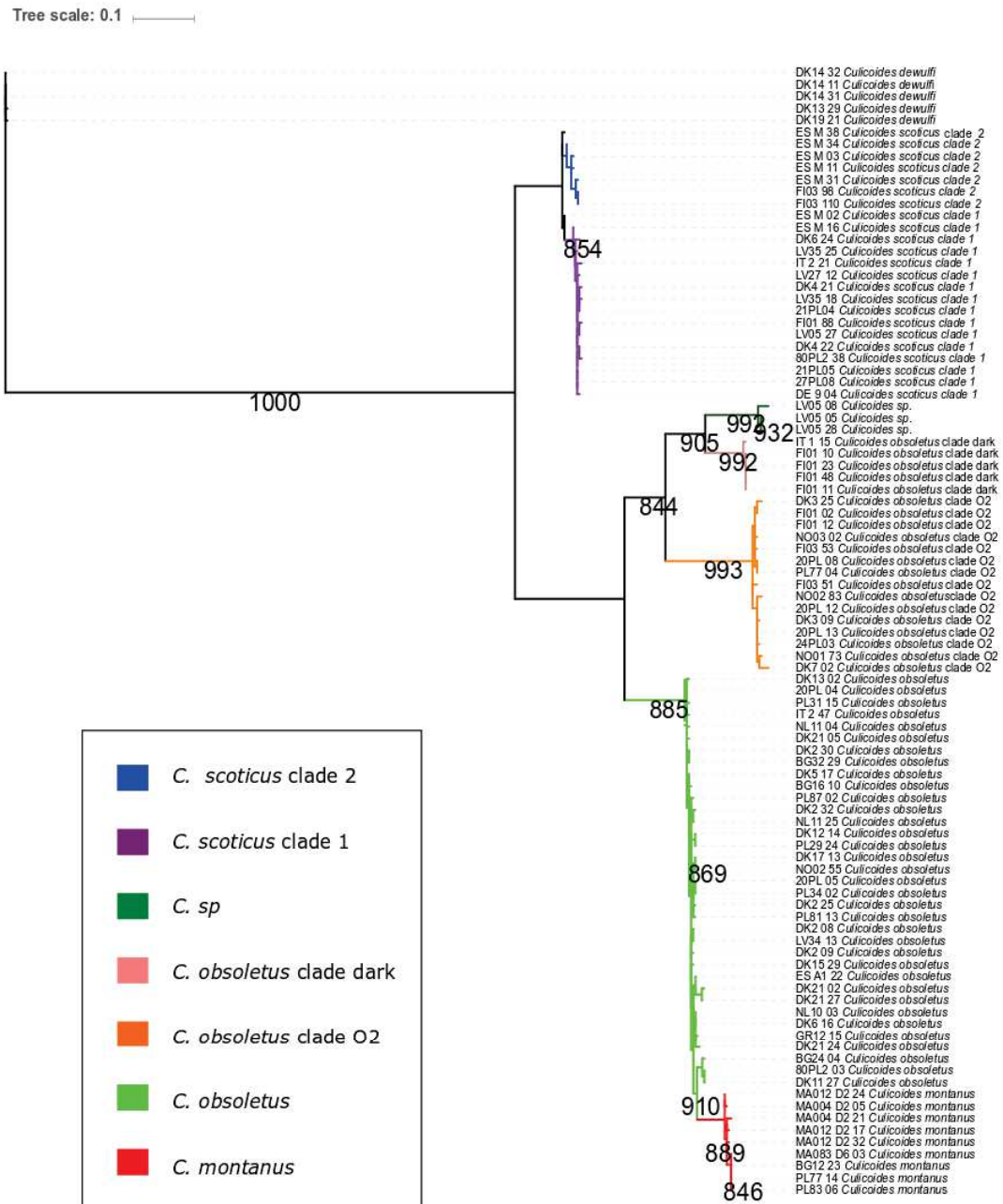


Figure S2. Maximum likelihood phylogenetic tree using *cox1*. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

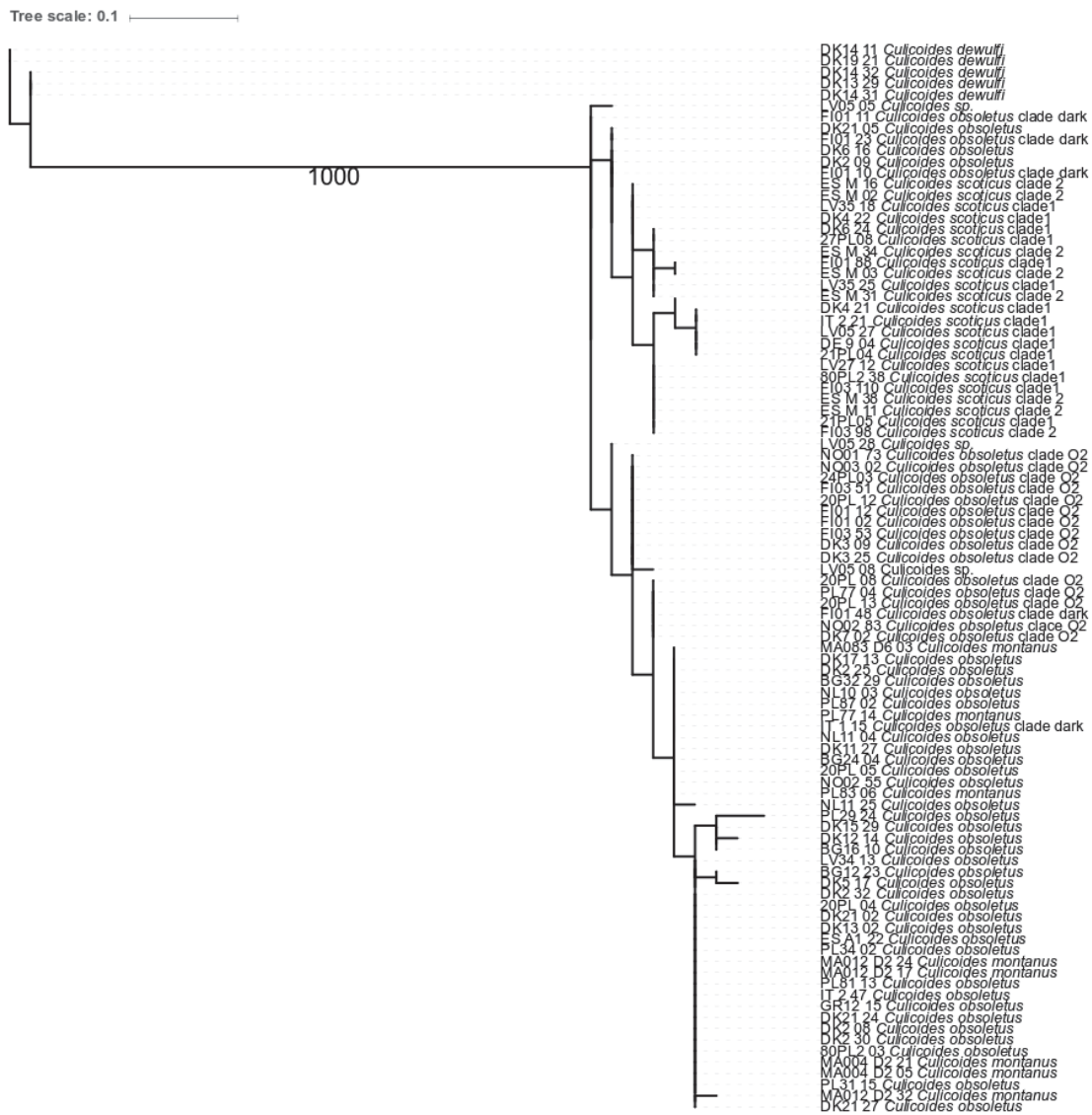


Figure S3 -Maximum likelihood phylogenetic tree using 28S rDNA. Values at the nodes represent bootstrap values greater than 800 for the most important nodes (1000 replicates).

Chapitre II : Recent phylogeographic history in western Europe of one main species vector of bluetongue and Schmallenberg viruses, *Culicoides obsoletus* (Diptera: Ceratopogonidae) demonstrates important continental connectivity

Culicoides obsoletus est une espèce autochtone et très abondante en Europe de l'ouest, mais la cartographie de la diversité cryptique a mis en évidence des variations de son abondance relative. *Culicoides obsoletus* est quasi omniprésent dans la région des Balkans. Au contraire, on le trouve en plus faible proportion dans le nord de l'Europe et beaucoup plus rarement en Afrique du nord. Les épisodes glaciaires du Pléistocène ont modifié l'aire de répartition géographique, la structure génétique et l'histoire démographique de la plupart des espèces européennes. Si l'histoire évolutive de *C. imicola* a été décrite, elle reste la seule espèce pour laquelle des hypothèses approfondies ont été faites pour expliquer la distribution de sa diversité actuelle.

Ce deuxième chapitre présente une étude multi-marqueurs de génétique des populations et de phylogéographie (11 marqueurs microsatellites nouvellement développés et le gène mitochondrial Cox1) menée à l'échelle européenne sur *C. obsoletus* pour identifier la structure génétique et phylogéographique de l'espèce et déterminer les événements historiques qui ont façonné le schéma actuel de la structure génétique de cette espèce.

Nos travaux permettent pour la première fois de déterminer, en observant les schémas de distribution des fréquences alléliques, comment la dispersion et les événements historiques ont pu façonner l'agencement spatial de la diversité génétique. Notre étude a révélé de faibles niveaux de différenciation génétique entre les populations à l'échelle européenne. Les méthodes bayésiennes de classification révèlent un cluster principalement représenté dans la zone des Balkans, un représentatif des îles Canaries, et un autre englobant le reste des populations Européennes. Nous avons montré une absence de signature phylogénétique et l'analyse du réseau d'haplotypes met également en évidence un manque de structuration spatiale. Nos résultats soutiennent une expansion démographique récente. Les méthodes bayésiennes ABC nous ont permis de mettre en évidence une divergence récente (au cours des 10 000 dernières années) des populations de *Culicoides*, avec une recolonisation post-glaciaire à partir des refuges nordiques et balkaniques, suivie d'une expansion géographique vers l'ouest de l'Europe des populations issues des refuges, créant des populations mélangées à l'ouest de l'Europe.

Ce chapitre présente des résultats préliminaires qui sont encore à affiner avant soumission pour publication dans *Scientific Reports*.

Recent phylogeographic history in western Europe of one main species vector of bluetongue and Schmallenberg viruses, *Culicoides obsoletus* (Diptera: Ceratopogonidae) demonstrates important continental connectivity

Antoine Mignotte^{1,2*}, Claire Garros^{1,2}, Laetitia Gardès^{1,3}, Thomas Balenghien^{1,2,4}, Maxime Duhayon^{1,2}, Ignace Rakotoarivony^{1,2}, Karine Huber¹.

¹ ASTRE, Univ Montpellier, Cirad, INRA, Montpellier, France.

² Cirad, UMR ASTRE, F-34398 Montpellier, France.

³ Cirad, UMR ASTRE, F- 97170 Petit-Bourg, Guadeloupe, France.

⁴ Institut Agronomique et Vétérinaire Hassan II, Unité Parasitologie et maladies parasitaires, 10100 Rabat, Morocco.

* Corresponding author: Antoine Mignotte, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, antoine.mignotte@cirad.fr; or Claire Garros, UMR 117 ASTRE, Campus international de Baillarguet, 34398 Montpellier cedex 5, claire.garros@cirad.fr

Abstract

Since 2006, Northern Europe faced massive outbreaks of bluetongue virus (BTV) which affects wild and domestic ruminants. This was followed by the emergence of Schmallenberg virus (SBV) in 2011, another *Culicoides*-borne virus, which also caused important economic losses to livestock breeders. *Culicoides obsoletus* has been identified as the main vector species in western Europe. *Culicoides obsoletus* is autochthonous and very abundant in Europe, in particular in the Balkan Peninsula. On the contrary, it is found at a lower proportion in northern Europe and much rarer in North Africa. During the Pleistocene glacial episodes, the geographical range, genetic structure and demographic history of most European species was impacted. Moreover, *Culicoides* dispersion is referred to as "stratified dispersion", which results from the combination of processes occurring at short distances and jumps over long distances. The objective of this work is to determine, by observing patterns of allelic frequency distributions, how dispersion and historic events may have shaped the spatial arrangement of genetic diversity. This study is conducted at continental scale, on two transects from Morocco to Finland and from Turkey to the United Kingdom. We used 11 microsatellite markers and a mitochondrial gene (*cox1*), to identify genetic and phylogeographic structure of *C. obsoletus* at the European scale and determine historical events that have shaped the current genetic structure pattern. Our study highlights important gene flows at the continental level and the connectivity between populations separated by a high geographical distance. The use of mitochondrial and microsatellite markers has allowed us to highlight a lack of phylogeographic and genetic structure, a recent increase in effective population size linked with post-glacial recolonization from Nordic and Balkanic refuges, then a geographical expansion towards the West of Europe of the populations coming from the refuges, creating populations admixed in the West of Europe. The recent population expansion highlights the potential role of the glacial episode and the post-glacial colonization event in shaping the geographic distribution of *C. obsoletus*.

Key words: *Culicoides obsoletus*, ABC, Dispersion, Microsatellites, Genetic structure, Palearctic region, population genetics, Phylogeography

Introduction

In 2006, western Palearctic region faced outbreaks of bluetongue virus (BTV) followed by Schmallenberg virus (SBV) in 2011 (Carpenter *et al*, 2013; Carpenter *et al*, 2009; Hoffmann *et al*, 2009; Savini *et al*, 2008). The unexpected nature of 2006 outbreaks compared to the outbreaks in the Mediterranean basin in the 1990-2000's is based on the absence of *C. imicola* in the western Palearctic region (Carpenter *et al*, 2009; Mellor *et al*, 2008; Purse *et al*, 2005). Massive spread in the western Palearctic region raises the question on the vector competence of native *Culicoides* (Carpenter *et al*, 2009; Wilson and Mellor, 2009). Among them, *Culicoides obsoletus* has been highlighted as the main vector species (Carpenter *et al*, 2009; Hoffmann *et al*, 2009; Savini *et al*, 2008).

Culicoides obsoletus is a native and abundant species in Europe and able to breed in a wide range of habitats (Mehlhorn *et al*, 2007; Purse *et al*, 2015). Initially described in Europe, many synonymous species have been described as *varius* ((Winnertz), 1852) in Germany, *yezoensis* ((Matsumura), 1911) in Russia, *lacteinervis* (Kieffer, 1919) in Ukraine, *kabyliensis* (Kieffer, 1922) in Algeria, or *intermedius* (Okada, 1941) in Japan. *Culicoides obsoletus* has a wide latitudinal distribution from North Africa (Morocco (Bourquia *et al*, 2019), Tunisia (Sghaier *et al*, 2017), Algeria (Hakima *et al*, 2020)) to northern Europe (Finland and Sweden (Mignotte *et al*, 2020)). It is also found on islands such as the Balearic Islands (Miranda *et al*, 2003), the Canary Islands (Martinez-de la Puente *et al*, 2012), and Corsica (Delecolle, 1985). It is thus mainly present in the northern hemisphere (Borkent, 2016). However, although its distribution center seems to be in Europe, it is also present on two other areas of the northern hemisphere (Borkent, 2016), America (Canada (Anderson *et al*, 1991) and USA (Borkent, 2016), and Asia (China (Gao *et al*, 2020), South Korea (Cho and Chong, 1974)). This species is therefore considered to be Palearctic and Nearctic (Borkent, 2016; Mathieu, 2011). Its trophic behaviour is flexible, allowing it to be present in a wide range of habitats. In addition, *C. obsoletus* is very generalist and is capable of blood feeding on a wide range of hosts such as goats, horses, sheep, rodents and birds (Garros and Balenghien, 2017). The abundance of *C. obsoletus* varies greatly depending on the country (Cuellar *et al*, 2018). A recent study describing the diversity within the Obsoletus/Scoticus Complex showed a pattern of decreasing proportion of *C. obsoletus* numbers captured following an East-West European gradient (Mignotte *et al*, 2020). Indeed, in the Balkan Peninsula, *C. obsoletus* is the most abundant species of the complex almost the only species caught during the sampling surveys. This has also been highlighted in Romania,

Slovakia, Albania, Macedonia, Kosovo, Montenegro, Serbia, Bosnia, Bulgaria and Croatia (Goffredo *et al*, 2004; Hristescu *et al*, 2020; Mignotte *et al*, 2020; Pudar *et al*, 2018; Sarvasova *et al*, 2014). On the contrary, it is found at a lower proportion in northern Europe (Mignotte *et al*, 2020) and much rarer in North Africa (Bourquia *et al*, 2019; Hakima *et al*, 2020; Sghaier *et al*, 2017). In addition, *C. obsoletus* has a strong intra- and interspecific diversity distributed throughout Europe, with western areas, such as France and Italy, where all the members of the Obsoletus/Scoticus Complex have recently been identified (Mignotte *et al*, 2020).

Geographic range, genetic structure and demographic history of most European species was modified during the glacial episodes of the Pleistocene (Bowler and Benton, 2005; Hewitt, 2004; Phillips *et al*, 2010; Widmer and Lexer, 2001). As a result of the rise in temperatures during interglacial episodes, new ecological niches became vacant in ice free regions, allowing the recolonization of no cold adapted species from glacial refuges (Hewitt, 2004; Petit *et al*, 2003; Taberlet *et al*, 1998). This post-glacial colonization also allows the creation of reconnection areas. In Europe, the Balkan, Italian and Iberian peninsulas have played the role of glacial refuges due to the reduction of viable areas during the glacial episodes (Hewitt, 1999; Petit *et al*, 2003; Taberlet *et al*, 1998). During the interglacial periods, these peninsula were the geographical origin of the spatial and demographic expansion of many species (Abellan and Svenning, 2014; Gómez and Lunt, 2007; Rodríguez-Sánchez *et al*, 2010). For example, the Balkan Peninsula has served as glacial refuges for butterflies (Schmitt *et al*. 2006a). In addition, recent studies show that in some areas in central and eastern Europe with stable microclimates may also have acted as glacial refuges for some cold tolerant species (Bhagwat and Willis, 2008; Schmitt and Varga, 2012; Sommer and Zachos, 2009; Stewart and Lister, 2001; Whittaker and Field, 2000).

The dispersal capacity of insect vectors is an important factor in the development of epidemiological models for predicting the evolution of epidemics. Indeed, dispersal capacity of insect vectors can be used to identify spatial pathogen transmission pattern (Kozakiewicz *et al*, 2018). The dispersion of *Culicoides* corresponds to what ecologists call 'dispersive stratified' which results from the combination of processes occurring at different spatial scales, that is to say combination of dispersal processes occurring actively at short distances, and passively or semi actively with jumps over long distances with wind-mediated dispersal patterns (Jacquet *et al*, 2016a; Murray and Kirkland, 1995; Onyango *et al*, 2015a). Important gene flows have been demonstrated between remote geographical areas in *Culicoides* species in Mediterranean basin

(Jacquet *et al*, 2016a), Africa (Onyango *et al*, 2015a) and in Australia (Pili *et al*, 2010). Dispersal capacity is an important component during periods of recolonization of newly available habitats, as is the case during inter-glacial periods. This allows for rapid spatial and demographic expansion towards ecological niches that have become vacant again.

Gene cytochrome oxidase I (cox1) has been used in phylogeographic studies on *C. imicola* in order to identify the colonization routes of this initially Afrotropical vector in the Mediterranean basin (Calvo *et al*, 2009; Nolan *et al*, 2008). Maternally inherited, the mitochondrial DNA has a relatively high rate of evolution and does not recombine between loci, which makes it a coherent marker for reconstructing the pathways of taxa during their recolonisation (Estoup and Guillemaud, 2010; Kirk *et al*, 2013). Microsatellite markers are complementary for studying population connectivity and demographic events (Hewitt, 2004). The use of a multi-local approach allows for a more refined and time-scaled estimate of connectivity between populations (Brito and Edwards, 2009; Prugnolle and de Meeus, 2002). We characterized the genetic structure using Bayesian clustering and traditional population genetics tools Approximate Bayesian Computation (ABC) (Beaumont and Rannala, 2004) were used to identify evolutionary scenarios to retrace and date demographic and historical events shaping the current population genetic structure of *C. obsoletus*. The objective of our study is to identify species connectivity on a large geographical scale of main species vector of bluetongue and Schmallenberg viruses in Europe. For this, we used 11 microsatellite markers and a mitochondrial gene (cox1), to identify genetic and phylogeographic structure of *C. obsoletus* at the European scale and determine historical events that have shaped the current genetic structure pattern. We used populations from 20 countries along a Palaeartic Mediterranean transect covering Scandinavia to Canary Islands (North to South) and Canary Islands to Turkey (West to East). These results will allow us to clarify the study of the dispersal of *C. obsoletus* on land, but also to know the phylogeographic history of one main vectors of livestock virus in Europe.

Methods

Sampling and species morphological identification

We collected a total of 2409 individual female biting midges in 19 countries, from 59 sampling sites in the Palearctic region between 2009 and 2017 using national surveillance networks for *Culicoides* populations or local collections (Table S1). Collection and morphological identification were carried out according to the same protocol as in a previous study on the phylogeny of *Culicoides* (Mignotte *et al*, 2020).

DNA extraction, amplification and sequencing

A total of 2409 individuals belonging to the *C. obsoletus* species complex were DNA-extracted, *cox1* amplified and sequenced following the procedure detailed in Mignotte *et al.* (2020). To ensure that only *C. obsoletus* individuals were included in the final data set, a DNA barcode analysis was performed according to Mignotte *et al.* (2020). All *cox1* sequences were then aligned using the MUSCLE algorithm (Edgar, 2004) implemented in the software package GENEIOUS v.6.0.5 (Biomatters, <http://www.geneious.com>). In order to estimate unbiased gene flows, only individuals identified as *C. obsoletus* were used for further analyses.

Among these individuals, 829 were amplified and genotypes (Tab S6) according to (Mignotte *et al*, Submitted in 2020). Allelic size allocations for all individuals and microsatellite markers were performed using the program GeneMapper® version 5 (Life Technologies, Applied Biosystem) with double blind reading to limit the potential interpretation bias from the reader.

Mitochondrial analyses

We calculated nucleotide diversity (π), haplotype number (H) and diversity (Hd) using the program DNASP v.5.10 (Librado and Rozas, 2009). In order to detect a past demographic change we performed neutrality tests based on Tajima's D and Fu's Fs statistics with DNASP v.5.10. (Fu, 1997; Ramos-Onsins and Rozas, 2002; Tajima, 1989). In order to detect sites under adaptative selection, we run fixed-effects likelihood (FEL) and single-likelihood ancestor counting (SLAC) on the *cox1* data set. These analyses were run in the DataMonkey server (<http://www.datamonkey.org>) (Kosakovsky Pond and Frost, 2005; Murrell *et al*, 2012). We inferred demographic history using mismatch distribution analysis (Rogers and Harpending, 1992) with ARLEQUIN 3.5 program. Due to the lack of spatial phylogeographic structuring, it was not possible to define a similar population grouping so all populations were therefore treated as a single population.

Genetic structure, diversity and dissimilarity

Matrix of pairwise Φ_{ST} among populations (Excoffier *et al*, 1992) was calculated with SPADS, and used to investigate mitochondrial isolation-by-distance. In order to visualize genealogical relationships among European populations, a median-joining haplotype network of *cox1* mitochondrial sequences of *C. obsoletus* was constructed with POPART (REF). Mitochondrial data set was used to evaluate population differentiation G_{ST} (Pons and Petit, 1995) that makes use only of the allelic frequencies, N_{ST} (Pons and Petit, 1996) for which similarities between the haplotypes are considered. The phylogeographic structure is evaluated by testing $N_{ST} - G_{ST}$ (Pons and Petit, 1996) with 10,000 permutations using SPADS (Dellicour and Mardulyn, 2014). Program MAPI was used based on 1,000 permutations and an α value of 0.05, to map inter-population genetic distances, Φ_{ST} , to visualize the spatial variation in genetic dissimilarity and define areas where genetic dissimilarity is significantly lower or higher than expected by chance (Piry *et al*, 2016)

Microsatellites analyses

Population structure and genetic diversity

We calculated number of alleles per locus, the allele size range, the observed and expected heterozygosity were estimated per locus, and per population over the entire dataset using the R packages *poppr* (Kamvar *et al*, 2014) and *mmod* (Winter, 2012). We calculated in R the Spearman's rank correlation coefficient between the allele sizes and their frequencies (Wattier *et al*, 1998) to identify a potential bias related to short allele dominance (Manangwa *et al*, 2019; Wattier *et al*, 1998). Null allele frequencies were calculated for each pair of loci using the methods of Chakraborty *et al* (1994) and Brookfield (1996) (Brookfield, 1996; Chakraborty *et al*, 1994) implemented in R package *PopGenReport* (Adamack *et al*, 2014). Linkage disequilibrium between each pair of loci were performed by the calculation of index of association (Agapow and Burt, 2001) implemented in R package *poppr* (Kamvar *et al*, 2014). Bonferroni's correction was applied to linkage disequilibrium analyses to take into account potential biases associated with multiple comparisons (Rice, 1989). We calculated the fixation index (F_{IS}) and tested the deviation from the Hardy-Weinberg equilibrium, (Weir and Cockerham, 1984) using FSTAT v2.9.3.2 (Goudet *et al*, 1996). Correlation between F_{IS} and null allele frequencies was calculated to try to explain a potential deviation from the Hardy-Weinberg equilibrium.

We used the program STRUCTURE (version 2.3.4) (Pritchard *et al*, 2000) to estimate the most probable number of genetic clusters (K) and individual probability of assignment to each genetic cluster. In order to assume identical allele frequencies among populations, Bayesian clustering analyse was run with correlated allele frequency. Admixture model was used to estimate for each individual the belonging proportion to each K clusters. Particularly adapted when there is a weak genetic structuring (Hubisz *et al*, 2009), the “locprior” model was used to use sampling locations as prior information to assist the clustering. Value of tested K varying from 1 to 10, with 10 independent runs, with a burn-in of 100,000 iterations in the Markov Monte Carlo chain (MCMC) and a stationary phase of 1,000,000 iterations in the MCMC. We used the STRUCTURE Harvester web platform (Earl and vonHoldt, 2011) to identified most probable number of clusters with the ΔK method (Evanno *et al*, 2005). In order to visualize the partition in genetic differentiation among all samples a neighbor-joining (NJ) tree (Takezaki and Nei, 1996) was constructed with the software POPULATIONS v.1.2.30 (<http://bioinformatics.org/~tryphon/populations/>) from a Cavalli-Sforza and Edwards genetic distance (1967) matrix. The trees were then subjected to 1000 bootstrap replications to confirm topological accuracy. In order to investigate isolation-by-distance, we used the R package *fields* (Nychka *et al*, 2017) to perform Mantel tests with the R package *vegan* (Oksanen *et al*, 2013) and linear regressions between inter-population ($F_{ST}/(1-F_{ST})$) genetic distances and log-transformed geographic distances between sampling sites with R package *fields* (Nychka *et al*, 2017). Allelic richness, classically defined as the number of alleles present in a population at a given locus, was calculated and corrected for variations in sample sizes between populations using the rarefaction method of the R package *Hierfstat* (Weir and Goudet, 2017; Yang, 1998). We used NE ESTIMATOR software and linkage disequilibrium method to evaluate the effective population size of *Culicoides* populations using the microsatellite dataset (Do *et al*, 2014). The program MAPI was used, based on 1,000 permutations and an α value of 0.05, to map inter-population genetic distances (pairwise F_{ST}) to visualize the spatial variation in the genetic dissimilarity and to define areas where the genetic dissimilarity is significantly lower or higher than expected by chance (Piry *et al*, 2016).

Scenarios inference with Approximate Bayesian computation

In order to infer evolutionary scenarios of *C. obsoletus* range expansion in Europe, we used Approximate Bayesian computation (ABC) (Beaumont *et al*, 2002; Bertorelle *et al*, 2010) approach using DIYABC software v.2.0.4 (Cornuet *et al*, 2014). ABC analyses were run on our microsatellite dataset. We used the population independently without grouping them together

as representative of the geographical area. The list of the mutational, demographic and historical prior of model parameters set implemented in DIYABC are present in table S3. We used mean number of alleles, mean allelic size variance, mean expected heterozygosity (Nei *et al*, 1975), pairwise F_{ST} values (Weir and Cockerham, 1984), mean ratio of the number of alleles over the range of allele size (Garza-Williamson's M) (Garza and Williamson, 2001), and the classification index (Pascual *et al*, 2007) as summary statistics. We assumed 4 generations per year for *C. obsoletus* (Mathieu, 2011). Microsatellite mutation rate use for ABC analysis ranging from 6.10^{-6} to 10^{-4} . Each scenario was tested using 1 million of simulated data sets. In order to identify the most probable scenario, we calculated the posterior probabilities associated with each scenario using a polychotomous logistic regression (Cornuet *et al*, 2014). This regression was performed on the 1% of the simulated data sets closest to the observed data set.

In order to reconstruct evolutionary history of *C. obsoletus* in Europe, different ABC scenarios were tested. This set of scenarios is aimed at identifying the pattern of divergence and the genealogy between the populations. For this purpose, the populations of Velestino (Greece), Narvik (Norway), Les Hitres (France) and Torre del Campo (Spain) were used. The populations included in this analysis are some of the most genetically differentiated, populations and are located in distant geographical areas, making it possible to trace phylogeographical history of *C. obsoletus* in Europe and test the hypothesis of European glacial refuges. All the tested scenarios are presented in table S2. In view of the main European glacial refuges and the genetic structure observed in the Mediterranean basin, the ABC scenarios were constructed to identify the geographical origin of post-glacial expansion at different dates of divergence (old: 20,000-100,000 years, recent: 5,000-25, 000 years, or very recent: 2500 - 10,000 years). Four coalescence-based population evolution hypotheses on the demographic and phylogeographic history of *C. obsoletus* in Europe were tested with second set of scenario and different divergence time: (i) post-glacial recolonization from Nordic and Balkanic refuges, then a geographical expansion towards the West of Europe of the populations coming from the refuges, creating populations admixed in the West of Europe (ii) post-glacial recolonization from Iberic and Balkanic refuges, then a geographical expansion towards the North of Europe of the populations coming from the refuges, creating populations admixed in the North of Europe (iii) an origin from the Iberian Peninsula with a settlement from West to East (iiii) an origin from the Balkan Peninsula with a settlement from East to West (Figure 1).

Results

Mitochondrial analyses

We sequenced a fragment of 627 bp of the *cox1* gene from 2409 *C. obsoletus* from 59 sampling sites (Table S1). We identified 106 different haplotypes in the *cox1* dataset. The haplotype diversity (H_d) was 0,772 for all mitochondrial data set and a total of 67 variable sites out of 627 bp was observed. Haplotype diversity ranging from 0,895 (Dittingen, Switzerland) to 0,107 (Arucas, Spain). All the metrics showing the genetic diversity for each population of the *cox1* dataset are shown in Table S4.

Genealogical relationships, genetic structure and genetic diversity

The median-joining network based on *cox1* sequences (Figure 2) shows, four widespread dominant haplotypes (H26, H86, H44, H5) grouping together individuals collected in distant geographical areas at the European scale. They are surrounded by several rare haplotypes distant from a few mutations. The median-joining network shows a star-like pattern signature characteristic of populations that have undergone a demographic expansion. The AMOVA Φ_{ST} estimator shows a significant genetic differentiation over all sampling sites ($\Phi_{ST} = 0.083$, p-value < 0.05). No significant phylogeographic signal was observed ($N_{ST} - G_{ST} = -1.357$, p-value = 1.0).

The mitochondrial genetic dissimilarity results from MAPI analyses show a relatively homogeneous pattern of dissimilarity, except for populations from Canary and Morocco regions. The maximum inter-individual genetic dissimilarity is of 0.495 and the lowest is of -0.240. The Canary Islands is the most genetically dissimilar region. On the contrary, the continental Europe together with British islands showed a strong genetic homogeneity (Figure 3.a). A significant but low isolation-by-distance was detected with the population genetic distance Φ_{ST} for the mitochondrial dataset ($r=0.020$, $p=0.002$, Figure 3.b). Mitochondrial nucleotide diversity (P_i) shows a heterogeneous distribution of diversity (Fig S3).

Demographic history of C. obsoletus in western Palearctic Europe

The mismatch distributions plot showed a unimodal shape indicating recent demographic expansions in Europe (Rogers and Harpending, 1992) (Figure 4). We rejected neutrality for the neutrality analyses on the whole *cox1* dataset with significant negative values observed for

Tajima's D ($D = -2.17897$, $p\text{-value} < 0.05$) and Fu's ($D = -25.81725$, $p\text{-value} < 0.05$) indicating a strong demographic expansion or natural positive selection events. No sites under positive selection were detected by SLAC and FEL analyses, and only 22 sites of 627 sites of *cox1* were under negative selection ($P\text{ value} < 0.1$). The effective population sizes were estimated for each population with Ne Estimator and range from 17.1 to infinite value (Table S7).

Microsatellites analyses

Microsatellites clustering analyses, mapping of genetic distances and genetic diversity

We genotyped 829 *C. obsoletus* from 27 sampling sites located in the western Palearctic region (Table S6) at 11 microsatellite loci (Table S5). The number of alleles per locus varied from 14 (OBSms29) to 47 (OBSms5) for the whole dataset (Supplementary Table 8). All populations are at Hardy-Weinberg equilibrium (Supplementary Table 7). All population genetic diversity metrics are shown in the table S7. F_{IS} values per locus are significantly correlated with null allele values ($r=0.953$, $p\text{-value}>0.001$). No significant correlation between allele size and allele frequency was found for any of the 11 loci (Table S7). The index of association does not revealed any linkage disequilibrium between all loci after Bonferroni correction for multiple comparisons.

The Bayesian clustering analysis performed with STRUCTURE on microsatellite data identified four genetic groups (ΔK maximum for $K=4$) (Figure S1). Cluster 1 is the majority in almost all the populations of Continental Europe and the British Isles (Figure 5). The populations of Continental Europe revealed a low genetic structure of *C. obsoletus*. The second cluster is the majority in two populations of Greece and Turkey. The third cluster is only represented in majority in the Canary Islands and in Serbia. The results for the second optimal K value ($K=2$) of the STRUCTURE program gave a similar result, reinforcing the conclusion that the populations are genetically weakly structured, with only two populations from Greece and Turkey standing out (Figure S2). The rarefied allelic richness of the microsatellite dataset varies from 5,903 (Canary Islands) to 8,513 alleles (Turkey). The Balkan region is the region with the greatest allelic richness (Macedonia and Turkey). On the contrary, the populations of Northern Europe and Western Europe are the ones with the lowest allelic richness. No clear geographical structure was identified between all sample using neighbor-joining tree based on Cavalli-Sforza and Edwards genetic distance (FigS4).

The microsatellite genetic dissimilarity results from MAPI analyses showed a non-homogeneous pattern of dissimilarity. The maximum inter-individual genetic dissimilarity was of 0.042 and the lowest was of 0.011. The southern Europe was the most genetically dissimilar region. On the contrary, northern Europe showed low genetic dissimilarity (Figure 6.a). However, almost none of the areas was significantly genetically dissimilar in MAPI analysis with the exception of a few areas in northern Europe showing significant genetic dissimilarity. A significant isolation-by-distance was detected with the population genetic distance $F_{ST}/(1-F_{ST})$ for microsatellite dataset ($R^2=0.084$, $p=0.003$, Figure 6.b).

Scenarios inference with Approximate Bayesian Computation

The posterior probabilities calculated for each scenario provided higher statistical support to the scenarios 1 ($p = 0.6350$) (Table 1). This scenario assumes an origin from the Balkan peninsula and from Eastern Europe followed by an admixture event, with a recent divergence time (10,000-2500 years) (Table S2). All the tested scenarios and results are presented in table S2. Type I and II errors were calculated in order to evaluate to what extent this scenario could be trusted. Type I error rate (proportion of simulations in which the scenario considered as true is not true) for true scenario was 0.432. Type II errors (proportion of simulations in which the true scenario is not considered true) are 0.142 for scenarios 1. These relatively high values of error could reflect that our data are not enough informative or that our scenario could be more complex. The real and simulated data are relatively well grouped by Principal Component analysis (PCA) (Fig. S6), indicating that some priorities could be refined, but that our scenario is still consistent in explaining the real dataset. We can conclude that scenario 1 is the most likely scenario among the proposed scenarios provide relatively good description of observed dataset, but it will have to be completed.

Discussion

Combined approaches of population genetics and phylogeography, based on microsatellite and mitochondrial markers have allowed us to explore the genetic structure but also the evolutionary and demographic history of *C. obsoletus* at the European scale. Our study revealed low levels of genetic differentiation between populations at the European scale. Bayesian clustering methods reveal four genetic clusters distributed all over the study area but one of the clusters is mainly represented in the Balkan area and another one encompassed the Canary Islands. We have shown an absence of phylogenetic signature using the difference between N_{ST} and G_{ST} . The haplotype network analysis highlights also a lack of spatial structuring. The significant negative values of Fu's and Tajima's D support a strong demographic expansion. Approximate Bayesian computation methods allowed us to highlight a recent divergence (10,000-2,500 years) of *Culicoides* populations, with a post-glacial recolonization from Nordic and Balkanic refuges. Then, a geographical expansion towards the West of Europe of the populations coming from the refuges, creating populations admixed in the West of Europe.

A strong dispersal capacity that blurs the genetic structure

The study of mitochondrial polymorphism does not reveal any genetic structure in our study area. Microsatellite markers do not make it possible to highlight a very marked genetic structure although a few slightly differentiated Mediterranean populations are observed. Mitochondrial-nuclear discordances are regularly reported in studies using a multi-markers approach (Kolleck *et al*, 2013; Rasic *et al*, 2015). This type of discordance is expected because the mitochondrial genome is haploid and uniparentally inherited in most animal species, and therefore has a fourfold smaller effective population size. Except incomplete lineage sorting, mitonuclear discordance can also arise if there are differences in how selection acts on the mitochondrial genome as compared to the nuclear genome or if there is biased movement of either marker type driven by demographic asymmetries, such as sex-biased dispersal (Frank *et al*, 2011). If selection for mtDNA variants varies geographically, then discordant patterns between mtDNA and nuDNA can arise (Irwin *et al*, 2005). Our results do not show a positive selection signal, but only a few sites under negative selection on mtDNA. However, these negatively selected sites are found in all sampling areas, which rules out selection as an explanatory factor of the observed spatial genetic patterns (Murrell *et al*, 2012). Female-biased dispersal can also promote the dispersal of mtDNA in the absence of concordant movement of nuDNA (Funk and Omland, 2003). If the dispersion of *Culicoides* by the wind on long distance is independent of

the sex it is not the case of the active dispersion on shorter distances. Indeed, in blood-feeding insects, females generally travel longer distances than males in search of egg-laying sites and blood meals (Kirkeby *et al*, 2013). The combination of these different mechanisms could therefore explain the slight signal differences between the two types of markers.

Only the population of the Canary Islands is well genetically differentiated. This population also exhibits low allelic richness and a very low nucleotide diversity. Island species are likely to have experienced population bottlenecks at some point in their evolutionary history due to founder events during the initial island colonization. As only a fraction of individuals from the original population found an island population, only a fraction of the original genetic diversity of the population will be maintained, and effective population sizes (N_e) will be small (James *et al*, 2016; Nei *et al*, 1975). As a result, this can lead to a reduction in genetic diversity and rapid genetic differentiation between these species and their native populations.

In many species of flying insects, there are important dispersal phenomena marked by a weak genetic structure (Lecocq *et al*, 2013; Lecocq *et al*, 2016; Mayer *et al*, 2015). For example, in Hymenoptera species of the genus *Bombus* a genetic differentiation is only observed between continental and island populations whereas no significant differentiation was found among populations from the European continent (Ellis *et al*, 2006; Estoup *et al*, 1996; Goulson *et al*, 2011; Lecocq *et al*, 2013; Shao *et al*, 2004; Widmer *et al*, 1998). This high level of gene flow is consistent with what we have observed for *C. obsoletus* at the scale of France (Mignotte *et al*, Submitted in 2020). This has also been reported for other *Culicoides* species in other geographical areas, in Australia for *C. brevitarsis* (Onyango *et al*, 2015a), in America for *C. stellifer* (Shults *et al*, 2020) and in Africa for *C. imicola* (Jacquet *et al*, 2016a).

We observed a significant but low level of isolation by distance in continental Europe that suggests a stepping-stone model of migration compatible with the combination of passive and active dispersal capacity of *Culicoides* species as it has already been demonstrated for *C. imicola* (Jacquet *et al*, 2016b). The passive dispersion of the *Culicoides* by winds over long distances certainly contributes to blurring the signal of isolation by distance. The intensification of grazing practices could allow to increase the size of *Culicoides* populations and facilitated the spread of BTV. The combined effects of active dispersion of *Culicoides* at short distances, passive dispersion by winds allowing long-distance jumping, and the recent intensification of breeding activities has certainly contributed to genetic homogenization, masking the past phylogeographic structure. *Culicoides obsoletus* is a species that feeds on a wide range of hosts and can lay eggs in a wide range of breeding sites (Garros and Balenghien, 2017). The generalist

and opportunistic behavior of this species, allowing them to adapt to a wide range of habitats (Harrup *et al*, 2016), has also contributed to this genetic homogenization.

A glacial refuge in the Balkan Peninsula during the Pleistocene

The low species diversity within the *Obsoletus/Scoticus* Complex observed in southeastern Europe and the Balkans (Mignotte *et al*, 2020), combined with the high allelic richness observed for *C. obsoletus* in this region, led us to hypothesize that this area could have served as a refuge for this species during the Pleistocene. The Balkan peninsula have played the role of glacial refuge for many species due to the reduction of viable areas during the glacial episodes (Hewitt, 1999; Petit *et al*, 2003; Taberlet *et al*, 1998). The lack of gene flow between isolated populations due to the glaciation has allowed genetic differentiation of populations due to genetic drift and the selection process, leading to potential speciation events (Hewitt, 1999; Hewitt, 2004). The spatial expansion of populations from glacial refuge has been accompanied by a gradient of reduction in the genetic diversity of the population during the colonization process (Comes and Kadereit, 1998), caused by serial founder events, creating genetic bottlenecks (Dlugosch and Parker, 2008; White *et al*, 2014). The genetic diversity is on average higher in areas that have long functioned as refuges than in areas of post-glacial redeployment (Prugnolle *et al*, 2005; White *et al*, 2013). In the same way for species of eastern Palearctic origin, the existence of refuges located within the limits of the ice caps where many plant species survived the last glacial maximum, have been shown (Abbott *et al*, 2003; Lagercrantz and Ryman, 1990). The Approximate Bayesian computation methods allowed us to identify as the most probable scenario a scenario identifying the Balkans and the arctic region as the post-glacial origin of ancestral populations, 5,000 to 10,000 years ago, which after more recent admixture events (2,500 years ago) gave rise to the populations of Western Europe.

The DIYABC software used to conduct these analyses does not allow to include recurrent migration phenomena in the tested scenarios (Cornuet *et al*, 2014), which can be a hindrance to study the evolutionary history of species with high dispersal capacities such as *C. obsoletus*. A high dispersal capacity can sometimes even prevent the identification of invasion routes, as was the case during the rapid and accidental invasion in Europe of *Leptoglossus occidentalis*, an insect native to North America (Lesieur *et al*, 2018). Moreover, some studies have suggested the existence of refugia not only in southern Europe, but also in western, central and eastern Europe (Schmitt and Varga, 2012; Stewart and Lister, 2001). *Culicoides obsoletus* has a very wide ecological niche allowing it to adapt to a wide variety of environments. In consequence, during episodes of glaciation, more scattered and northerly glacial refuges were able to allow

the persistence of coniferous and deciduous forests the animal species associated with them and among it potentially *C. obsoletus* (Provan and Bennett, 2008; Tzedakis *et al*, 2013). *Culicoides obsoletus* can be found from sea level to high altitudes. This wide altitudinal distribution has potentially allowed it, as it has already been shown for many species of boreal-mountain, a broader distribution during glaciations than in interglacial periods, limiting the genetic and phylogeographic structure of their populations (Schmitt and Varga, 2012; Taberlet *et al*, 1998). Thus, to have a global vision of the phylogeographic history of the species it would be necessary to integrate individuals from Asia and America. More locally, to dress a complete view of the evolutionary history of *C. obsoletus* in Europe, it would be relevant to integrate future ABC analysis of genotyped populations from the Italian peninsula, one of the three European glacial refuges, and more oriental population.

Several recent studies have highlighted cryptic diversity within the *Obsoletus/Scoticus* Complex (*C. obsoletus* clade O2, *C. obsoletus* clade dark, *C. montanus*) and raised questions about the taxonomic status of *C. montanus*, a species phylogenetically close to *C. obsoletus*. The taxonomic status of these clades is not yet fully established although there are strong indications that these groups constitute species in their own right (Ander *et al*, 2013; Meiswinkel *et al*, 2007; Mignotte *et al*, 2020). However, most of these studies were conducted using mitochondrial markers. These markers, by responding differently than nuclear markers to demographic fluctuations may have diverged more strongly while nuclear genes may have retained variability. It is thus possible that several highly divergent mtDNA lineages coexist over large geographical range despite the absence of complete reproductive isolation (Despres, 2019). Therefore, it is possible that by concentrating our study on *C. obsoletus sensu stricto*, we may have missed part of the evolutionary history of this species.

The haplotype mismatch distribution, the lack of matrilineal subdivision, the star-like pattern of median-joining haplotype network centred around few frequent haplotypes combined to significant negative values of Fu's and Tajima's D suggest a rapid demographic expansion of *C. obsoletus* in Europe. Recent demographic expansion is consistent with demographic history of most European species modified during the glacial episodes of the Pleistocene (Bowler and Benton, 2005; Hewitt, 2004; Phillips *et al*, 2010; Widmer and Lexer, 2001) but inconsistent with our ABC results which report more recent events. A demographic expansion has been also been detected in *C. imicola* with an increase of the effective population size in the western Mediterranean basin and sub-Saharan Africa during the last 17,000 years and 80,000 years

respectively (Jacquet *et al*, 2015). Given the dependence of *Culicoides* on bovine hosts, the development of modern human agriculture may have played an important role in the population explosion of *C. obsoletus* in Europe. However, the domestication of ungulates has only taken place in the Middle East and Europe during the last 10,000 years. It thus appears important to be able to specify the dating of these expansions in order to be able to refine our hypotheses.

Conclusion

Our study highlights important gene flows at the continental level and the connectivity between populations separated by a high geographical distance, possible thanks to the stratified dispersion of *Culicoides* combining short and long-distance displacements. The use of mitochondrial and microsatellite markers has allowed us to highlight a lack of phylogeographic and genetic structure, a recent increase in effective population size, a recent divergence of Western European *Culicoides* populations (5,000 to 10,000 years ago), and post-glacial origin located in the Balkans and the arctic region of ancestral populations. The population expansion highlights the potential role of the glacial episode and the post-glacial colonization event in shaping the geographic distribution of *C. obsoletus*. The observed difference in structure between nuclear and mitochondrial markers suggests that a coupled effect of stratified *Culicoides* dispersal, displacement by human activities, and the recent increase in reproductive activity may have led to genetic homogenization, erasing the past phylogeographic structure of the population. However, a more precise time of expansion and divergence would help to better understand the genetic homogeneity of *C. obsoletus* in the eastern Palearctic region. Our result on the demographic and evolutionary history of *C. obsoletus* leads us to reassess the importance of this vector in the dispersal of the BTV.

Author contributions

A.M., K.H. and C.G. designed the study. I.G., M.D. contributed the identification of *Culicoides*. A.M., L.G., have performed molecular biology manipulations. A.M. analysed the data. K.H. and C.G contributed to the manuscript firstly written by A.M. All authors read and commented the final manuscript version.

Acknowledgement

This study was partially funded by the EU grant H2020-727393 PALE-Blu and by the VectorNet project (OC/EFSA/AHAW/2013/02-FWC1) funded by the European Centre for Disease Prevention and Control (ECDC) and the European Food Safety Authority (EFSA). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission, of the ECDC or of the EFSA. The authors are grateful to all European partners who gave assistance in *Culicoides* sampling and shipping.

Figures and tables

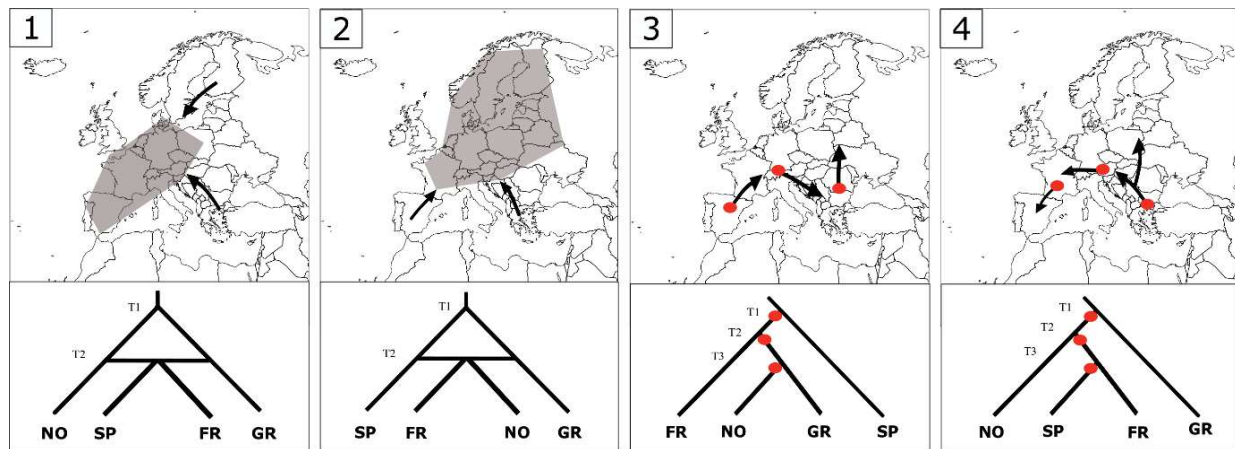


Figure 1. Graphical representation of the 4 scenarios simulated using an approximate Bayesian computation (ABC) approach on microsatellite data. Details of all scenarios and parameters are shown in Supplementary Table S2. Red circles correspond to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence times. NO: population from Norway (Narvik), SP: population from Spain (Torre del campo), FR: population from France (Les Hitres), GR: population from Greece (Velestino).

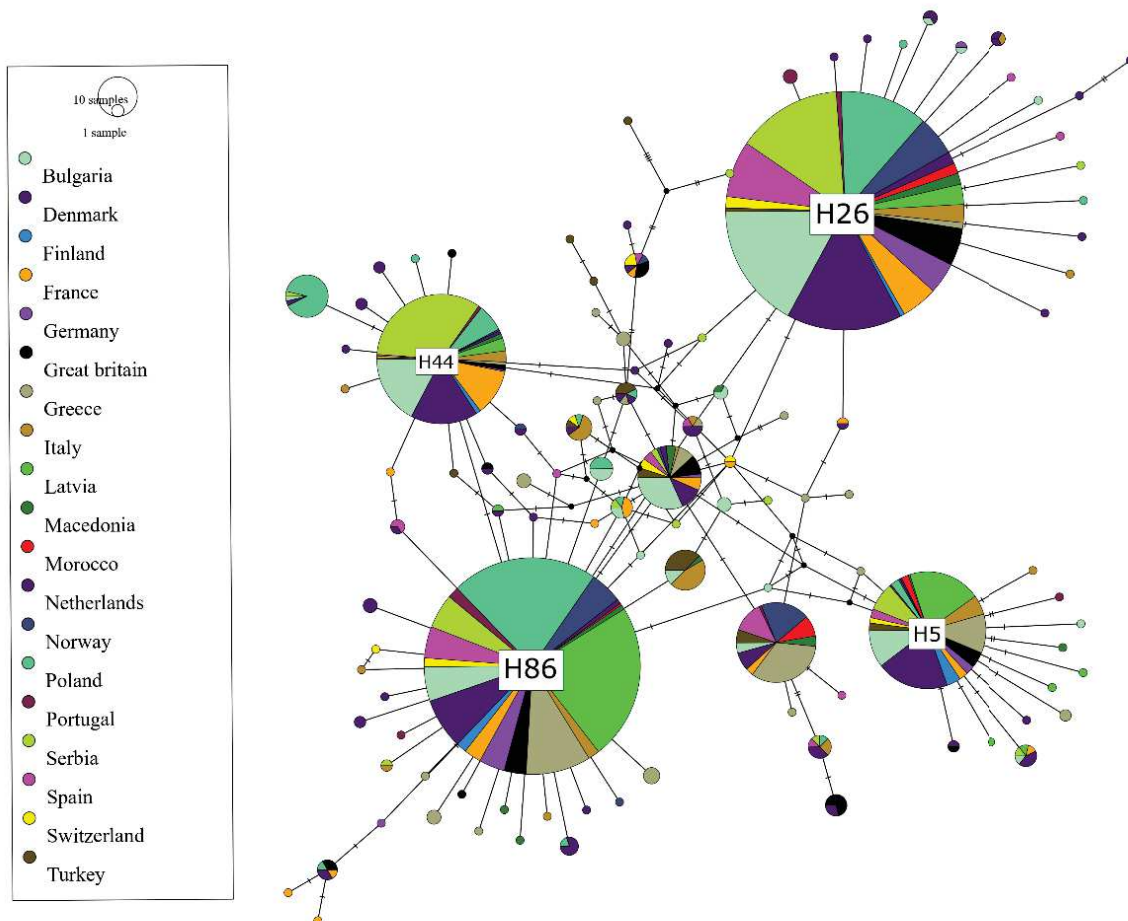


Figure 2. Median-joining haplotype networks of *cox1* mitochondrial sequences of *C. obsoletus*. The size of the circles is proportional to the number of individuals with that haplotype. Colors represent the country of sampled individuals. The length of the branches separating haplotypes is proportional to the number of mutational steps between them.

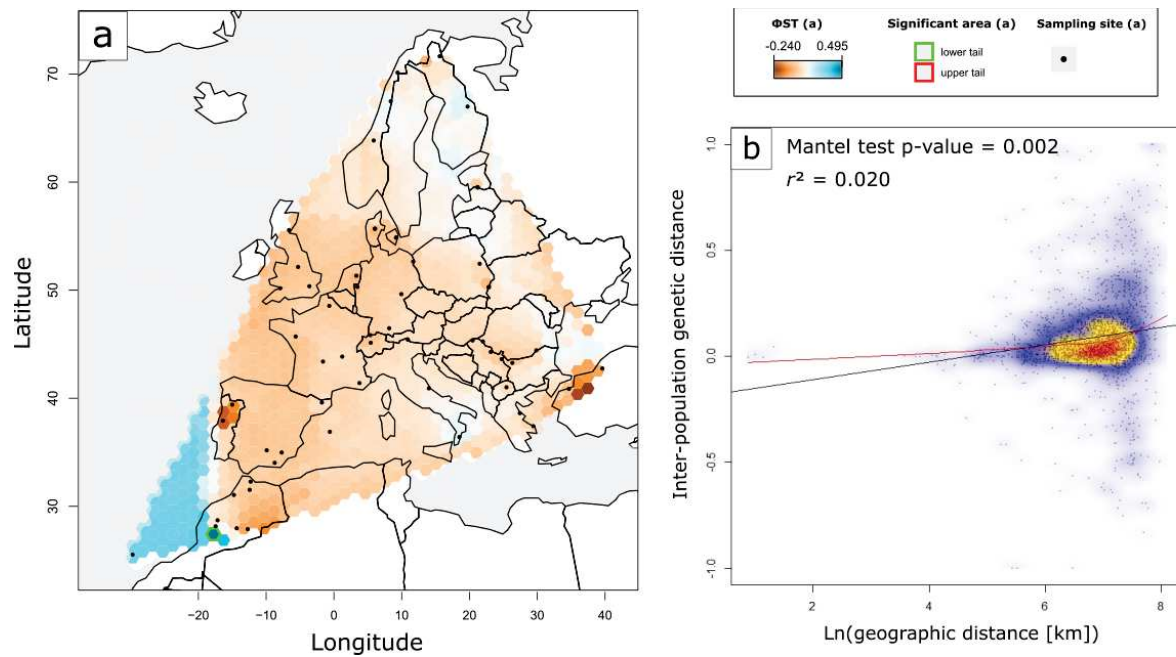


Figure 3. Mitochondrial genetic differentiation and isolation-by-distance analyses of *C. obsoletus*. Results of the smoothing of pairwise measures performed with MAPI and based on (a) $F_{ST}/(1-F_{ST})$ and (b) Isolation-by-distance analysis, Mantel test and a kernel density estimate of pairwise correlations from each sample between ($F_{ST}/(1-F_{ST})$ and geographic distance. In panels (a), genetic dissimilarity is represented by a color scale ranging from brown (lower genetic dissimilarity) to blue (higher genetic dissimilarity). The Canary Islands populations are not considered in these analyses.

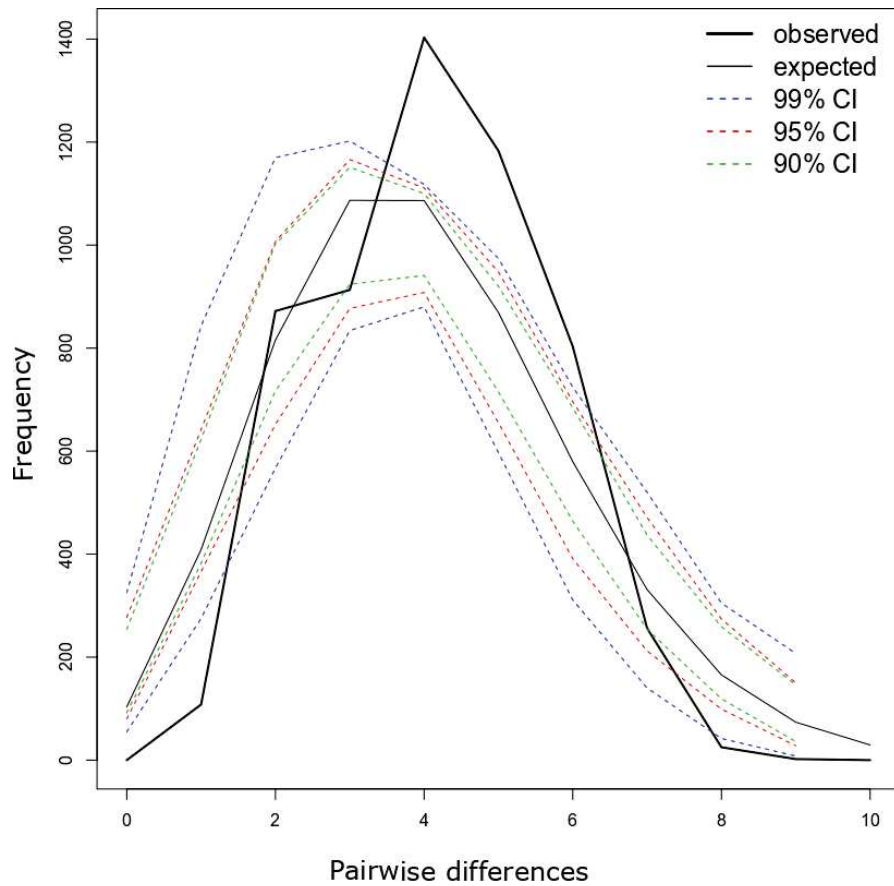


Figure 4. Mismatch distribution plots for 106 haplotypes of *cox1* of *C. obsoletus*. (a) The mean estimate and the 90, 95 and 99% highest posterior density (HPD) limits are indicated.

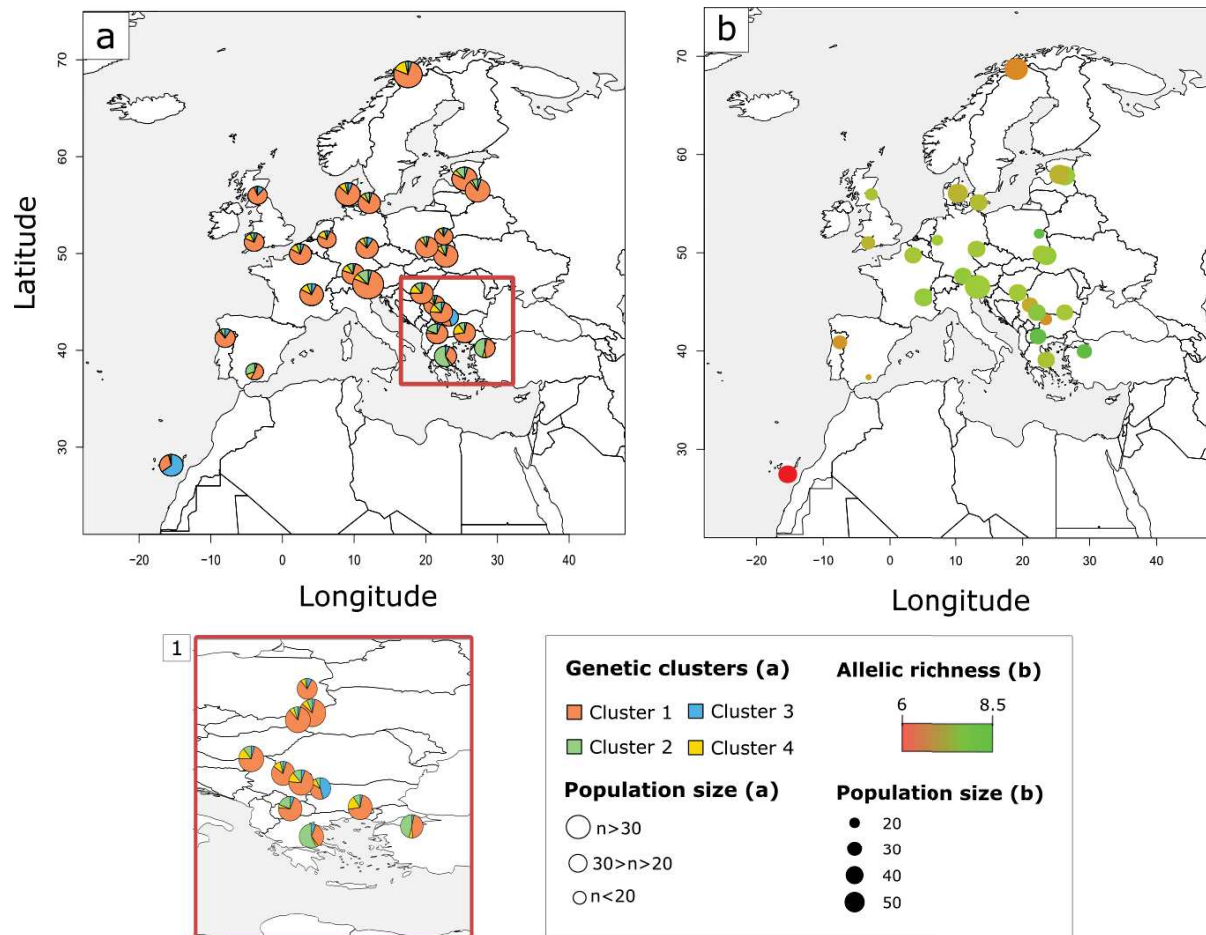


Figure 5. Microsatellite genetic clustering and genetic diversity of *C. obsoletus* sampled populations. Results of the genetic clustering analyses performed with STRUCTURE (a) and rarefied allelic richness of populations (b). Panel (1) corresponds to an enlargement of the geographical area (a) of the panel circled in red. One specific color has been assigned to each genetic cluster in panels (a) and (1).

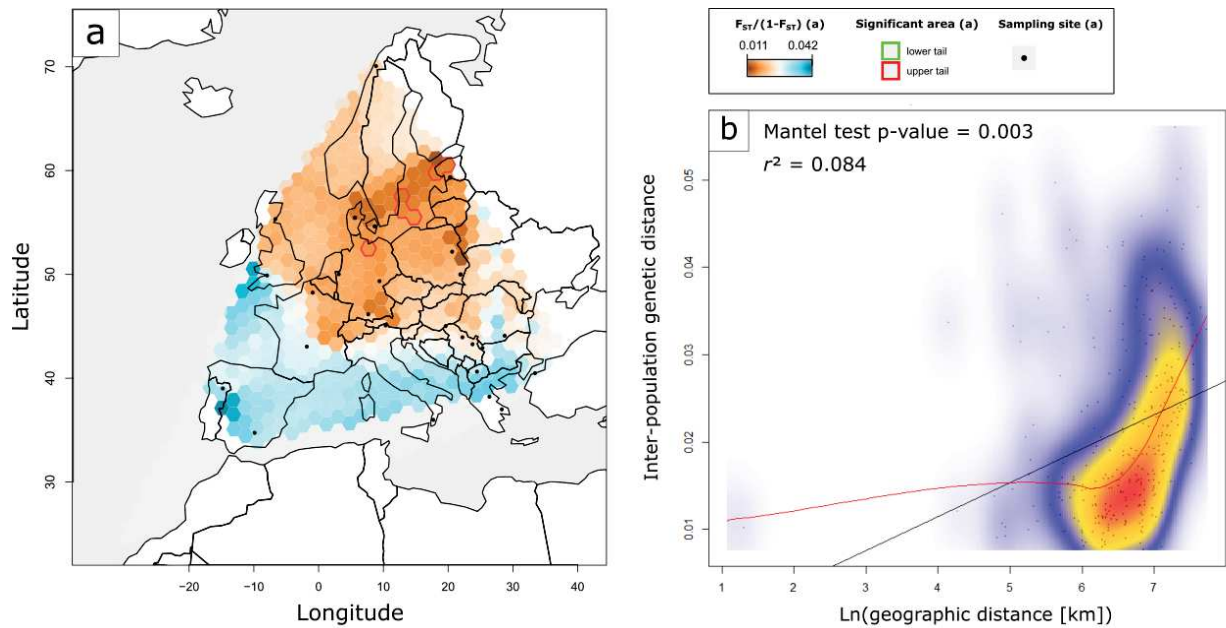
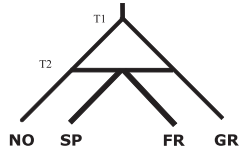
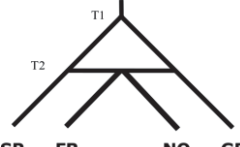
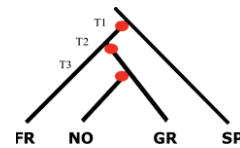
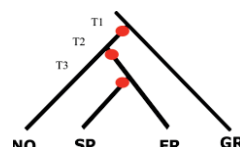


Figure 6. Microsatellite genetic differentiation and isolation-by-distance analyses of *C. obsoletus*. Results of the smoothing of pairwise measures performed with MAPI and based on (a) $F_{ST}/(1-F_{ST})$ and (b) Isolation-by-distance analysis, Mantel test and a kernel density estimate of pairwise correlations from each sample between ($F_{ST}/(1-F_{ST})$ and geographic distance. In panels (a), genetic dissimilarity is represented by a color scale ranging from brown (lower genetic dissimilarity) to blue (higher genetic dissimilarity). The Canary Islands population is not considered in these analyses because of its insularity.

Tableau 1. Result of confidence in scenario selection based on posterior probabilities to retrace evolutionary history of *C. obsoletus* in Europe by approximate Bayesian Computation (ABC) on microsatellite data. 95% confidence intervals and type I and II errors were presented. Red circles correspond to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence between two taxa. NO: Norwegian population (Narvik), SP: Spanish population (Torre del campo), FR: French population (Les Hitres), GR: Greek population (Velesino).

Scenario	Description of tested scenarios	Posterior probability [95% CI]	Error of scenario choice Type I	Error of scenario choice Type II
1		0.6350 [0.5795,0.6905]	0.432	0.142
2		0.2672 [0.2146,0.3198]		
3		0.0033 [0.0000,0.0230]		
4		0.0002 [0.0000,0.0201]		

Supplementary information

Tableau S1. Metadata of the 2409 sequences of *cox1 C. obsoletus*.

The table being too voluminous it is not included in this manuscript.

Tableau S2. Result of most probable scenario for all ABC simulation. 95% confidence intervals and type I and II errors were presented to. Red circles corresponds to a bottleneck event. T1, T2 and T3 correspond to successive dates of divergence between two taxa. NO: Norway population (Narvik), SP: Spain population (Torre del campo), FR: French population (Les Hitres), GR: Greece populations (Velestino).

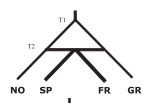
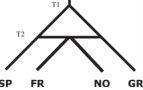
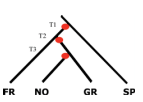
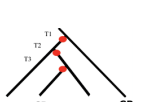

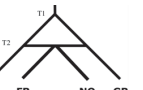
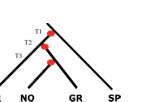




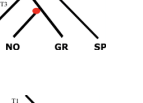
Scenario	Description of selected scenarios	Divergence date	Posterior probability [95% CI]	Error of scenario choice Type I	Error of scenario choice Type II
1		T1: 10 000-5000 years T2: 5000-2500 years	0.6350 [0.5795,0.6905]	0.432	0.142
2		T1: 10 000-5000 years T2: 5000-2500 years	0.2672 [0.2146,0.3198]		
3		T1: 10 000-5000 years T2: 5000-2500 years T3: 2500-1 years	0.0033 [0.0000,0.0230]		
4		T1: 10 000-5000 years T2: 5000-2500 years T3: 2500-1 years	0.0002 [0.0000,0.0201]		
5		T1: 25 000-12 500 years T2: 12 500-7500 years	0.0587 [0.0391,0.0783]		
6		T1: 25 000-12 500 years T2: 12 500-7500 years	0.0348 [0.0170,0.0527]		
7		T1: 25 000-12 500 years T2: 12 500-7500 years T3: 7500-5000 years	0.0004 [0.0000,0.0203]		
8		T1: 25 000-12 500 years T2: 12 500-7500 years T3: 7500-5000 years	0.0000 [0.0000,0.0200]		
9		T1: 10 000-5000 years T2: 5000-2500 years	0.0000 [0.0000,0.0200]		
10		T1: 100 000-50 000 years T2: 50 000-25 000 years	0.0000 [0.0000,0.0200]		
11		T1: 100 000-50 000 years T2: 50 000-25 000 years T3: 25 000-20 000 years	0.0000 [0.0000,0.0200]		
12		T1: 100 000-50 000 years T2: 50 000-25 000 years T3: 25 000-20 000 years	0.0000 [0.0000,0.0200]		

Tableau S3. Demographic, historical and mutational prior parameters for ABC analyses.

Effective population sizes (N) are expressed in number of diploid individuals and times of events (t) in number of generations going back to the past.

Parameter name	Distribution Prior	Parameters abbreviation
Mean mutation rate for nucleotide instability	Uniform [10 ⁻⁸ ; 10 ⁻⁴]	SNI
Mean parameter of the geometric distribution of length	Uniform [10 ⁻¹ ; 3.10 ⁻¹]	P
Mean microsatellite mutation model	Loguniform [6.10 ⁻⁶ ; 10 ⁻⁴]	μ
Very recent ancestral divergence time	Uniform [40 000; 20 000]	T1
Time of second event of divergence	Uniform [20 000; 10 000]	T2
Time of third event of divergence	Uniform [10 000; 4]	T3
Recent ancestral divergence time	Uniform [100 000; 50 000]	T1
Time of second event of divergence	Uniform [50 000; 30 000]	T2
Time of third event of divergence	Uniform [30 000; 20 000]	T3
Old ancestral divergence time	Uniform [400 000; 200 000]	T1
Time of second event of divergence	Uniform [200 000; 100 000]	T2
Time of third event of divergence	Uniform [100 000; 80 000]	T3
Duration of bottleneck	Uniform [0; 100]	DB
Effective population size during bottleneck	Loguniform [2; 1000]	N _{bott}
Stable effective population size in ancestral (a) and European (e) populations	Loguniform [100; 200000]	N _a , N _e

Tableau S4. Population genetic diversity of 2409 cox1 sequence of *C. obsoletus* in Europe.

Pi: Nucleotide diversity per site, Nh: haplotype number.

Locality	Country	Longitude	Latitude	Collection date	Abbreviation	Nseq	Number of variable sites	Nh	Haplotype diversity	Standard Deviation of Haplotype diversity	Pi	Standard deviation of Pi	Ar	Fu's Fs statistic	Tajima's D
Montana	Bulgaria	23,23	43,41	2016	Mon	143	22	21	0,737	0,034	0,004	0,000	2,846	-7,127	-1,100
Nikolovo	Bulgaria	26,115334	43,87123	2016	Nik	146	20	18	0,697	0,037	0,004	0,000	2,849	-5,454	-1,147
Dyssevej	Denmark	12,11	55,2	2015	Dys	83	30	24	0,873	0,026	0,007	0,000	2,036	-7,930	-1,113
Groddevej	Denmark	9,18	56,13	2015	Grw	50	22	17	0,759	0,062	0,005	0,001	2,540	-5,983	-1,308
Leestrupvej	Denmark	12,05	55,22	2015	Lee	92	23	20	0,752	0,042	0,005	0,001	2,348	-5,543	-1,124
Lægdsgårdsvej	Denmark	9,12	56,12	2015	Leg	108	30	26	0,823	0,029	0,006	0,000	2,259	-10,434	-1,338
Inari	Finland	27,56	68,67	2015	Ina	8	3,000	3	0,750	0,096	0,003	0,000	5,750	0,971	1,220
Sotkamo	Finland	28,72	64,04	2015	Sot	19	9	5	0,737	0,069	0,007	0,001	3,211	2,480	0,962
Barjols	France	5,98	43,54	2016	Bar	13	5	5	0,859	0,050	0,003	0,000	2,000	0,043	0,298
Bonnay	France	2,51	49,93	2016	Bon	38	16	12	0,774	0,056	0,005	0,001	2,368	-2,711	-0,992
Les Hitres	France	4,07	45,75	2016	Les	43	11	10	0,759	0,053	0,003	0,000	2,698	-2,223	-0,830
Reynac	France	1,8176424	45,3429047	2016	Rey	11	14	6	0,800	0,114	0,008	0,002	1,909	0,494	-0,483
Sucé-sur-Erdre	France	-1,54	47,32	2016	Suc	8	8	4	0,750	0,139	0,005	0,002	2,125	0,896	-0,866
Peetzig	Germany	13,94	53,08	2016	Pee	8	3	3	0,607	0,164	0,002	0,001	3,375	0,723	0,585
Schleiz	Germany	11,81	50,57	2016	Sch	33	11	8	0,680	0,054	0,003	0,001	3,121	-1,655	-1,349
Ziegelbach	Germany	9,86	47,89	2016	Zie	31	9	5	0,652	0,063	0,004	0,001	3,194	2,018	-0,116
Abercorn	Great britain	-3,455284	55,995067	2014	Abe	24	11	7	0,779	0,057	0,004	0,001	2,625	0,073	-0,553
Blackmoor	Great britain	-3,94	51,18	2015	Bla	30	13	7	0,756	0,050	0,005	0,001	2,567	0,992	-0,400
London	Great britain	-0,13	51,51	2010	Lon	27	16	8	0,604	0,107	0,005	0,001	4,778	-0,452	-1,328
Thorpe	Great britain	-1,78	53,04	2015	Tho	17	15	9	0,831	0,085	0,006	0,001	2,471	-1,744	-0,992
Varybobi	Greece	23,78	38,13	2016	Var	50	20	19	0,863	0,035	0,006	0,001	3,000	-9,043	-1,242
Velestino	Greece	22,75	39,38	2016	Vel	112	19	17	0,731	0,035	0,004	0,000	4,536	-5,840	-1,448
Brunico	Italy	11,96	46,8	2016	Bru	46	20	12	0,850	0,029	0,006	0,001	1,957	-1,095	-0,813
Mammola	Italy	16,19	38,38	2016	Mam	19	10	6	0,760	0,070	0,007	0,001	1,842	1,917	1,255
Teramo	Italy	13,67	42,64	2016	Ter	14	5	4	0,571	0,132	0,003	0,001	3,143	0,531	-0,099
Kalnciems	Latvia	25,44	57,66	2016	Kal	165	13	11	0,431	0,044	0,003	0,000	4,691	-0,284	-0,840
Mazlimbas	Latvia	25,4	57,71	2016	Maz	75	11	7	0,646	0,042	0,006	0,000	4,627	2,692	0,645
Oubogo	Macedonia	21,99037	41,625636	2016	Oub	34	18	12	0,832	0,054	0,005	0,001	2,206	-2,931	-1,272
Bab Taza	Morocco	-5,189793	31,264597	2016	Bab	1		1					5,000		
Chefchaouen	Morocco	-5,273638	35,161626	2016	Che	16	8	4	0,625	0,093	0,006	0,001	2,688	4,400	0,993
Dardara	Morocco	-5,301693	34,489081	2016	Dar	2	2	2	1,000	0,500	0,003	0,002	2,000	0,693	
Moulay Bouselham	Morocco	-6,245313	31,264597	2016	Mou	1		1					1,000		
Oulad Mtaa	Morocco	-8,256511	31,264597	2016	Oul	1		1					1,000		
Rabat	Morocco	-6,8	33,96	2010	Rab	4	4	3	0,833	0,222	0,004	0,001	1,000	0,461	1,365
Tizgzaouine	Morocco	-8,130402	31,767321	2016	Tiz	1		1					3,000		
Lomm	Netherlands	6,2	51,45	2017	Lom	18	12	8	0,797	0,088	0,004	0,001	2,444	-1,860	-1,363
Olst	Netherlands	6,12	52,37	2017	Ols	11	11	5	0,709	0,137	0,004	0,002	2,909	0,315	-1,471
Narvik	Norway	17,51	68,48	2015	Nar	82	13	8	0,684	0,029	0,003	0,000	5,098	-0,180	-1,042
Storforshei	Norway	14,81	66,31	2015	Sto	4	0	1					2,750		
Tiller	Norway	10,42	63,34	2015	Til	19	3	4	0,626	0,067	0,002	0,000	6,053	0,462	0,275
Jaksamanice	Poland	22,86	49,751	2016	Jak	93	15	13	0,678	0,035	0,003	0,000	2,882	-4,184	-1,326
Olszewnica	Poland	22,51	51,76	2016	Olz	25	8	6	0,690	0,065	0,003	0,001	3,400	0,671	-0,849
Pikulice	Poland	22,79	49,75	2016	Pik	195	18	15	0,671	0,024	0,003	0,000	3,815	-4,611	-1,347

Aboadela	Portugal	-8	41,27	2009	Abo	26	14	12	0,880	0,034	0,005	0,001	1,962	-2,211	-1,048
Bajouca	Portugal	-8,77	39,87	2009	Baj	1		1					1,000		
Aleksandrovo	Serbia	20,58	45,63	2016	Ale	61	12	9	0,649	0,041	0,004	0,000	4,525	-0,884	-0,834
Cirikovac	Serbia	21,19	44,67	2016	Cir	78	12	12	0,680	0,034	0,003	0,000	2,423	-2,552	-0,639
Gamzigrad	Serbia	22,18	43,91	2016	Gam	71	14	12	0,743	0,036	0,004	0,000	2,732	-2,199	-0,780
Pacir	Serbia	19,45	45,89	2016	Pac	59	12	6	0,704	0,036	0,005	0,001	5,729	2,542	-0,153
Araut	Spain	1,760802	41,965161	2017	Ara	30	12	14	0,864	0,048	0,004	0,001	1,700	-4,367	-0,869
Aruca	Spain	-15,5	28,13	2010	Aru	37	2	3	0,107	0,068	0,000	0,000	4,054	-2,626	-1,493
Castril	Spain	-2,34	37,73	2017	Cas	16	9	5	0,683	0,091	0,003	0,001	3,063	0,237	-1,327
Orgiva	Spain	-2,99	36,85	2017	Org	6	2	2	0,600	0,129	0,002	0,000	3,667	1,938	1,753
Palma Torre del Campo	Spain	2,65	39,56	2015	Pal	26	9	8	0,723	0,062	0,004	0,001	3,308	-1,469	-0,694
	Spain	-3,90	37,83	2017	Tor	13	8	4	0,731	0,079	0,005	0,001	2,923	1,948	-0,161
Dittingen	Switzerland	7,5	47,44	2009	Dit	18	13	9	0,895	0,048	0,005	0,001	2,056	-1,904	-0,832
Muhleturmen	Switzerland	7,51	46,81	2009	Muh	15	12	5	0,676	0,105	0,004	0,001	3,333	0,894	-1,366
Şahinköy	Turkey	28,237	40,235	2016	gah	31	23	12	0,882	0,039	0,007	0,001	1,355	-1,650	-1,116
Kuzgöl	Turkey	32,09938	40,91156	2016	kuz	1		1					1,000		
All data Set						2409	67	106	0,772	0,005	0,005			-30,374	-1,985

Tableau S5. Primers of the 11 microsatellite markers used to genotype *C. obsoletus* populations. “DYE”: fluorochrome name, “Tm”: half denaturation temperature in degree, “bp”: base pairs.

Multiplex	Locus	DYE	Repeat motif	Forward primer (5'-3')	Tm	Reverse primer (3'-5')	Tm	Size range (bp)	Accession numbers
Plex1	OBSms5	FAM	(TG)15	GTCATTTATGTAATTTGTGCC	56,4	CGAGACGGCATGTTTATCT	56,3	140-204	MT828835
	OBSms11	FAM	(CT)12	CATCCATCCACTTTGCTCT	56,1	GGACTGTCATTCAAACAAACA	56,1	262-293	MT828838
	OBSms2	NED	(TG)14	TGATAACAGATAACACGGAAGG	56,8	TGATGCAATAATGTCCTCGT	58,6	135-171	MT828832
	OBSms28	NED	(CA)8	GCCGCAATTTGTGTTTT	55,5	TTCCTTGACCTGCGAAC	55,5	239-264	MT828842
	OBSms3	VIC	(CA)13	ATTTACGCGACGATCATT	55,8	AAAGCTCGAATACAGCACAA	56,2	140-184	MT828833
	OBSms4	PET	(GA)12	AATGTAGGCCAAACAGAAA	53,3	TTGCATAACAACATCATCAC	52,7	160-180	MT828834
Plex2	OBSms29	FAM	(CA)5	GTTGACCGGGACACTTTT	55,8	CATGACAAGCACGAACG	55,4	252-268	MT828843
	OBSms6	NED	(TA)12	GACCTAATTCATAATCGTCCAA	55,8	TTAAAGCACGTTTCGATTTGT	56,1	175-201	MT828836
	OBSms13	VIC	(CA)12	CGTTCCATCGACTCATCTT	56,1	GCTTTTACTCCTTCGTGATGT	56,1	249-287	MT828839
	OBSms14	VIC	(AT)8	TCAAATCATTCCAAGAGCA	55,1	CCTCTCTGTGGTCTTTCGT	55,2	139-164	MT828840
	OBSms10	PET	(AC)17	GACATGGAATCCTTAATTGG	54,5	GATTGCTTCAGACAAACACA	54,9	224-260	MT828837

Tableau S6. Microsatellite dataset of 829 *C. obsoletus* in Europe.

The table being too voluminous it is not included in this manuscript.

Tableau S7. Sampling sites and associated genetic diversity for microsatellite data set. N: sample size, Ho: observed heterozygosity, Hs: expected heterozygosity, Fis: Fixation index, Ar : allelic richness, Ne : sample size. P-values obtained for the Hardy-Weinberg test were adjusted for multiple comparisons at the nominal level (5%) of 0.00017.

Population name	Country	Longitude	Latitude	N	Ho	Hs	Fis	Rarefied Ar	Ne (0.05)	95% CIs for Ne (Parametric)	95% CIs for Ne (JackKnife on Samples)
Les Hitres	France	4,070	45,750	35	0,729	0,796	0,085	7,792	196.3	82,0 - Infinite	60,8 - Infinite
Bonnay	France	2,510	49,930	30	0,703	0,800	0,121	7,647	263.4	83,4 - Infinite	56,9 - Infinite
Nikolovo	Bulgaria	26,115	43,871	29	0,752	0,790	0,047	7,648	75.2	44,8 - 189,1	32,3 - Infinite
Montana	Bulgaria	23,230	43,410	20	0,750	0,777	0,035	6,873	Infinite	76,5 - Infinite	58,0 - Infinite
Arucas	Spain	-15,500	28,130	34	0,722	0,743	0,029	5,903	89.9	48,3 - 342,2	37,2 - Infinite
Dyssevej	Denmark	12,110	55,200	31	0,718	0,777	0,075	7,464	256.6	73,7 - Infinite	51,5 - Infinite
Grøddevej	Denmark	9,180	56,130	38	0,740	0,771	0,041	7,359	183.2	80,0 - Infinite	80,0 - Infinite
Blackmoor	Great Britain	-3,940	51,180	23	0,688	0,773	0,109	7,385	25.2	16,9 - 43,1	9,6 - 1368,4
Torre del Campo	Spain	-3,904	37,835	16	0,664	0,793	0,163	7,139	17.1	11,2 - 29,2	8,2 - 59,2
Ziegelbach	Germany	9,860	47,890	31	0,757	0,794	0,047	7,836	234.4	82,1 - Infinite	79,7 - Infinite
Velestino	Greece	22,750	39,380	30	0,718	0,819	0,124	7,546	42.7	29,6 - 69,7	24,4 - 107,9
Brunico	Italy	11,960	46,800	58	0,686	0,789	0,131	7,692	200.0	108,1 - 823,2	72,4 - Infinite
Mazlimbas	Latvia	25,400	57,710	40	0,747	0,797	0,062	7,811	80.7	52,7 - 152,3	34,6 - Infinite
Kalnciemi	Latvia	25,440	57,660	39	0,744	0,781	0,047	7,343	Infinite	362,9 - Infinite	211,9 - Infinite
Narvik	Norway	17,510	68,480	50	0,757	0,784	0,034	6,773	76.8	51,4 - 135,2 -	40,0 - 277,4
Aboadela	Portugal	-8,000	41,270	24	0,659	0,744	0,115	6,951	Infinite	85,6 - Infinite	77,8 - Infinite
Lomm	Netherlands	6,200	51,450	19	0,736	0,785	0,063	7,657	139.0	48,1 - Infinite	44,7 - Infinite
Pikulice	Poland	22,790	49,750	37	0,734	0,809	0,093	7,858	Infinite	250,1 - Infinite	187,1 - Infinite
Jaksamanice	Poland	22,860	49,751	32	0,730	0,799	0,086	7,662	Infinite	146,8 - Infinite	79,9 - Infinite
Cirikovac	Serbia	21,190	44,670	27	0,722	0,771	0,064	7,203	Infinite	98,9 - Infinite	78,9 - Infinite
Pacir	Serbia	19,450	45,890	32	0,728	0,791	0,08	7,671	82.9	45,6 - 276,9	33,1 - Infinite
Gamzigrad	Serbia	22,180	43,910	31	0,742	0,801	0,074	7,953	103.2	49,6 - 1263,9	44,5 - Infinite
Olszewnica	Poland	22,510	51,760	18	0,747	0,834	0,104	8,493	13492.6	75,9 - Infinite	48,4 - Infinite
Oubogo	Macedonia	21,990	41,626	28	0,747	0,824	0,093	8,449	73.7	43,0 - 199,1	33,8 - 2310,1
Abercorn	Great Britain	-3,466	55,997	21	0,723	0,809	0,107	7,590	918.7	53,9 - Infinite	36,4 - Infinite
Schleiz	Germany	11,810	50,570	31	0,736	0,807	0,088	7,649	190.3	69,7 - Infinite	61,5 - Infinite
Şahinköy	Turkey	28,237	40,235	25	0,716	0,829	0,136	8,513	123.5	55,7 - Infinite	39,2 - Infinite

Tableau S8. Genetic diversity by locus. Ho: observed heterozygosity, Hs: expected heterozygosity, Fis: Fixation index, SAD: short alleles dominance, CI: confidence indice.

Locus name	Alleles number	Ho	Hs	Fis	SAD r ²	SAD pvalue	Null alleles frequency [95% CI]
OBSms11	29	0,810	0,872	0,071	0,375	0,054	0,043 [0,028-0,058]
OBSms2	29	0,781	0,849	0,080	-0,064	0,755	0,056 [0,041-0,074]
OBSms28	21	0,702	0,729	0,036	0,194	0,413	0,024 [0,004-0,046]
OBSms3	34	0,734	0,788	0,068	-0,322	0,068	0,051 [0,031-0,069]
OBSms4	20	0,778	0,816	0,047	0,111	0,661	0,031 [0,013-0,049]
OBSms5	47	0,783	0,890	0,120	-0,098	0,528	0,076 [0,058-0,094]
OBSms10	26	0,750	0,876	0,145	0,036	0,865	0,083 [0,063-0,102]
OBSms13	29	0,872	0,890	0,021	-0,256	0,198	0,018 [0,005-0,029]
OBSms14	19	0,700	0,760	0,079	-0,298	0,244	0,046 [0,025-0,066]
OBSms29	14	0,350	0,416	0,158	-0,081	0,794	0,114 [0,078-0,151]
OBSms6	40	0,724	0,827	0,125	-0,074	0,650	0,089 [0,069-0,109]

Tableau S9. Posterior checking using all the summary statistics for ABC inferences of scenario 1, the most probable scenarios.

Summary statistics	Observed value	Proportion (simulated<observed)
NAL_1_1	10.3636	0.7180
NAL_1_2	12.0909	0.8390
NAL_1_3	11.2727	0.8200
NAL_1_4	8.2727	0.6945
HET_1_1	0.7835	0.6565
HET_1_2	0.7951	0.5030
HET_1_3	0.8178	0.7185
HET_1_4	0.7855	0.5625
VAR_1_1	10.9102	0.8370
VAR_1_2	12.1763	0.8530
VAR_1_3	8.6675	0.7030
VAR_1_4	9.4842	0.7630
MGW_1_1	0.8261	0.2085
MGW_1_2	0.8750	0.4405
MGW_1_3	0.8435	0.3410
MGW_1_4	0.6547	0.0510
N2P_1_1&2	14.0000	0.8130
N2P_1_1&3	14.4545	0.8465
N2P_1_1&4	12.9091	0.8395
N2P_1_2&3	15.0909	0.8690
N2P_1_2&4	14.0909	0.8870
N2P_1_3&4	12.7273	0.7815
H2P_1_1&2	0.8038	0.6180
H2P_1_1&3	0.8284	0.7590
H2P_1_1&4	0.8024	0.6750
H2P_1_2&3	0.8265	0.6840
H2P_1_2&4	0.8064	0.5595
H2P_1_3&4	0.8246	0.7040
V2P_1_1&2	11.5671	0.8410
V2P_1_1&3	10.4636	0.7870
V2P_1_1&4	10.7841	0.8240
V2P_1_2&3	10.9031	0.7920
V2P_1_2&4	11.6752	0.8350
V2P_1_3&4	8.9590	0.7120
FST_1_1&2	0.0380	0.6125
FST_1_1&3	0.0774	0.8305
FST_1_1&4	0.0603	0.7725
FST_1_2&3	0.0476	0.8865
FST_1_2&4	0.0377	0.7600
FST_1_3&4	0.0428	0.5855
LIK_1_1&2	1.8603	0.7095
LIK_1_1&3	2.2983	0.8785
LIK_1_1&4	2.3294	0.9120
LIK_1_2&1	1.9866	0.6005
LIK_1_2&3	2.1175	0.7920
LIK_1_2&4	2.2919	0.8830
LIK_1_3&1	2.6196	0.9205
LIK_1_3&2	2.3803	0.9510*
LIK_1_3&4	2.3464	0.8340
LIK_1_4&1	2.2651	0.8720
LIK_1_4&2	2.1221	0.9065
LIK_1_4&3	1.9438	0.5865

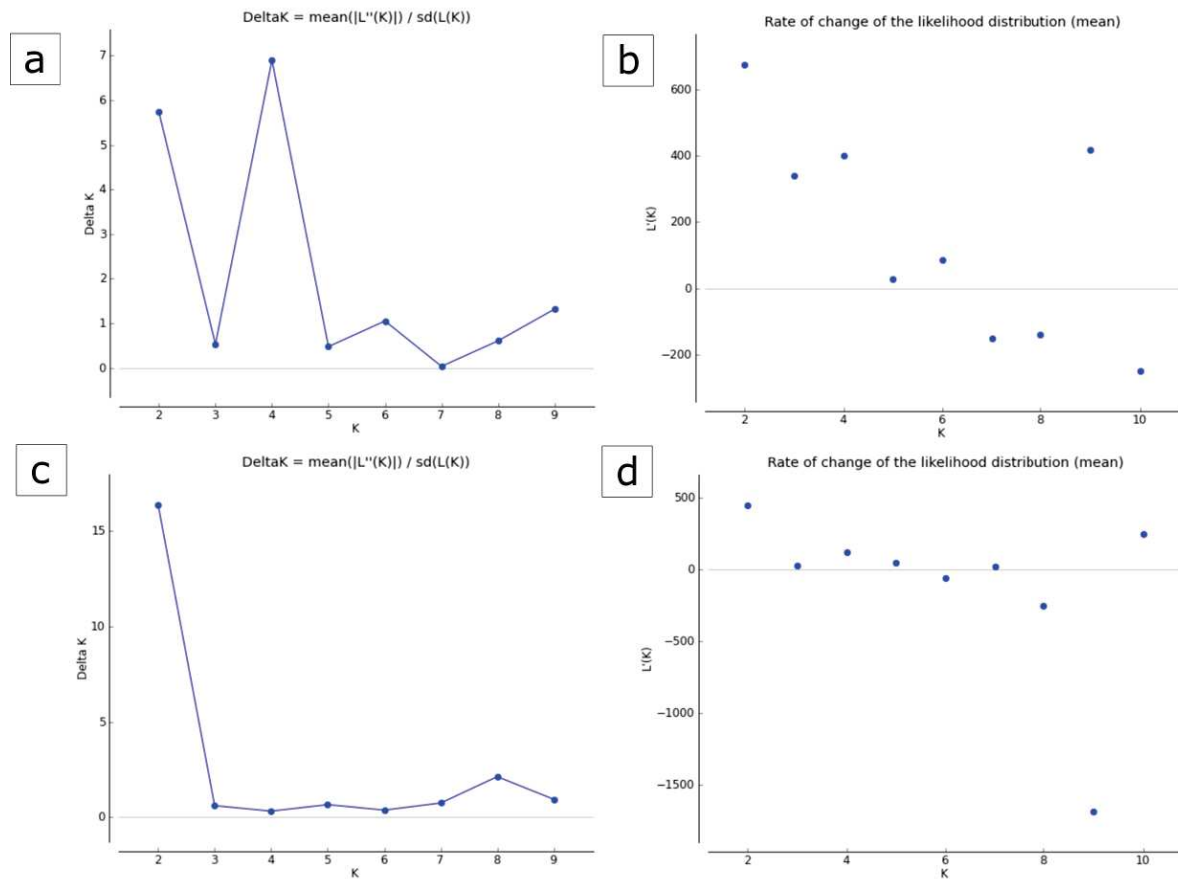


Figure S1. Identification of the optimal number of genetic clusters (K) inferred by STRUCTURE using the delta(K) and L'(K) methods. (a) and (b) panels corresponds to analyses with all 829 individuals, and (c) and (d) panels corresponds to STRUCTURE analyses on central Europe dataset.

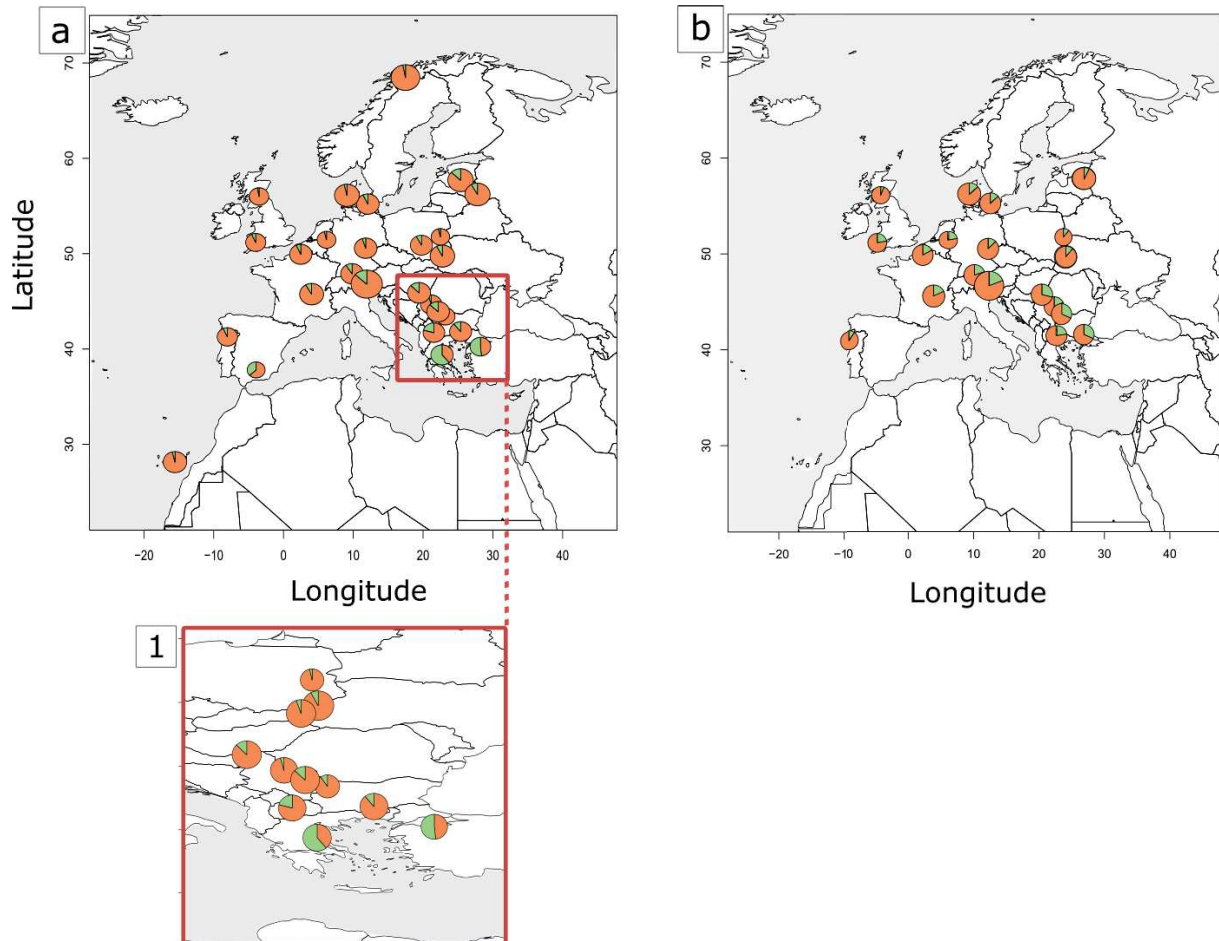


Figure S2. Genetic clustering of *C. obsoletus*. (a) Results of the genetic clustering analyses performed with STRUCTURE on complete dataset for K=2 (second most probable K), and (b) Results of the genetic clustering analyses performed with STRUCTURE on dataset of central Europe for K=2 (most probable K). Panel (1) corresponds to an enlargement of the geographical area (a) of the panel circled in red. One specific color has been assigned to each genetic cluster in panels (a) and (b).

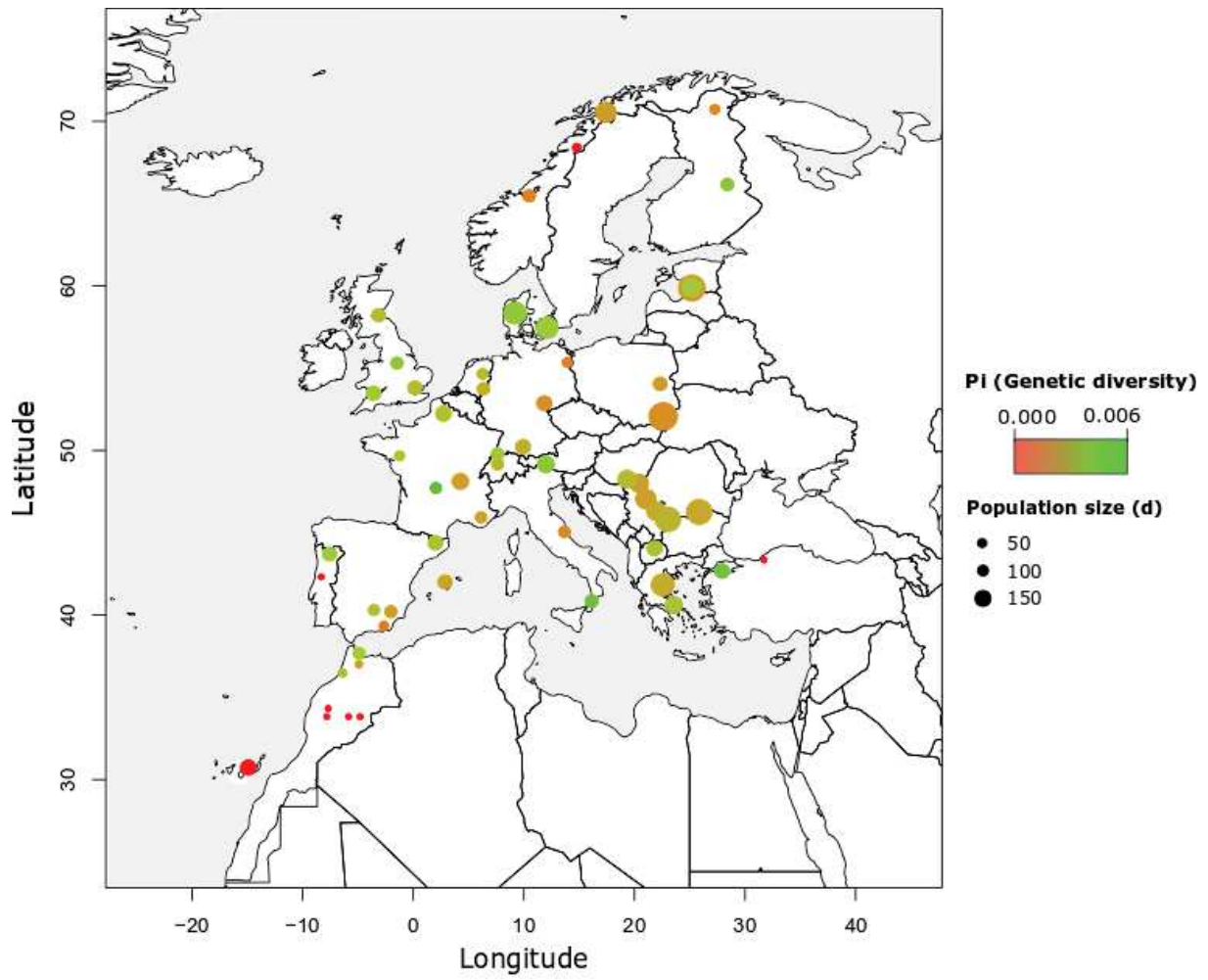


Figure S3. Map of genetic diversity of 2409 *cox1* sequence of *C. obsoletus*

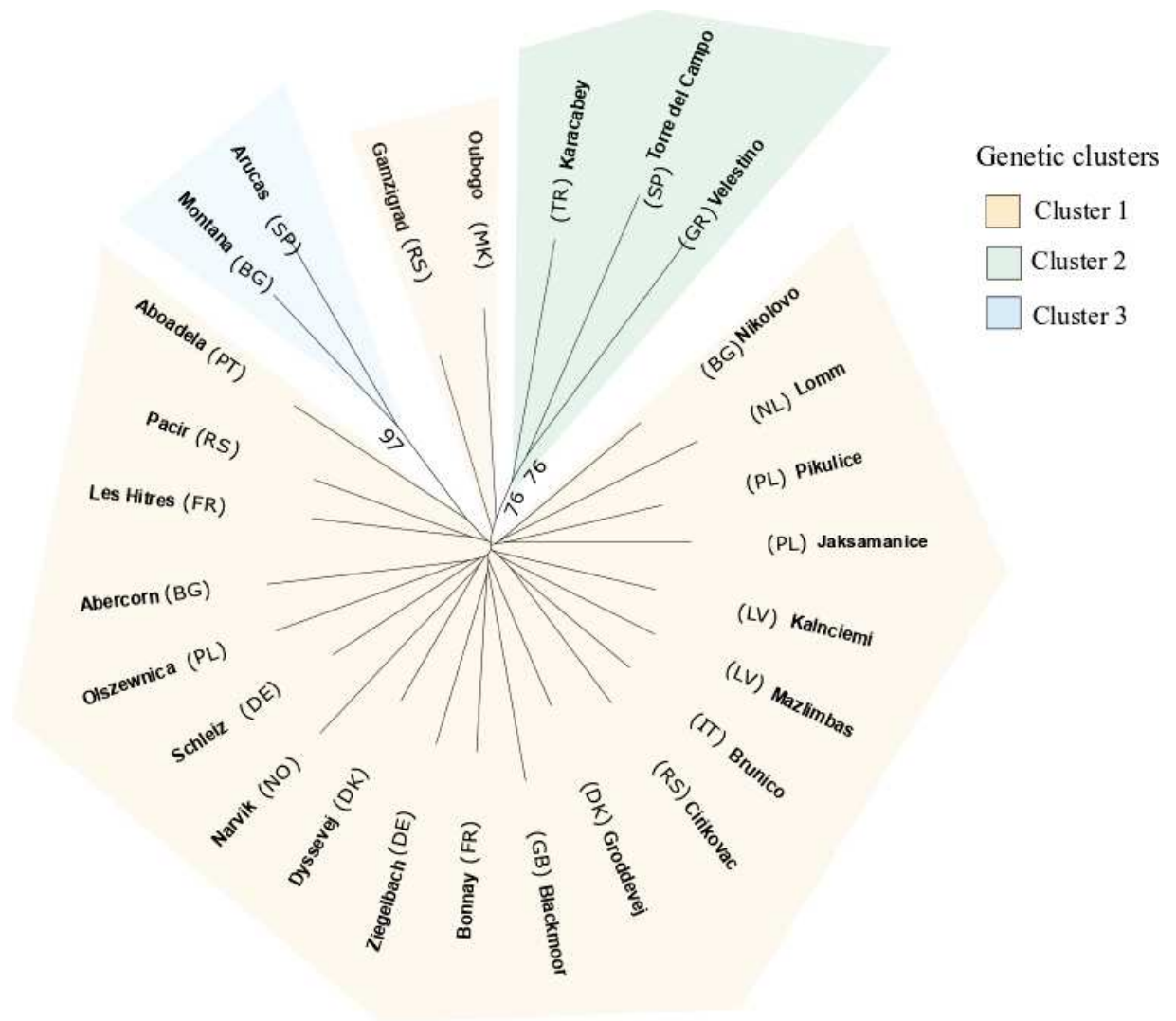


Figure S4. Neighbor-joining tree of *C. obsoletus* population based on the distance of Cavalli-Sforza & Edwards (1967) computed on microsatellite dataset. The belonging of populations to a genetic cluster defined by the STRUCTURE Analysis is represented by the colors, orange, green and blue for cluster 1, 2 and 3 respectively. Bootstraps values were calculated over 1000 replicates and are represented as percentage and only values superior to 60% are reported.

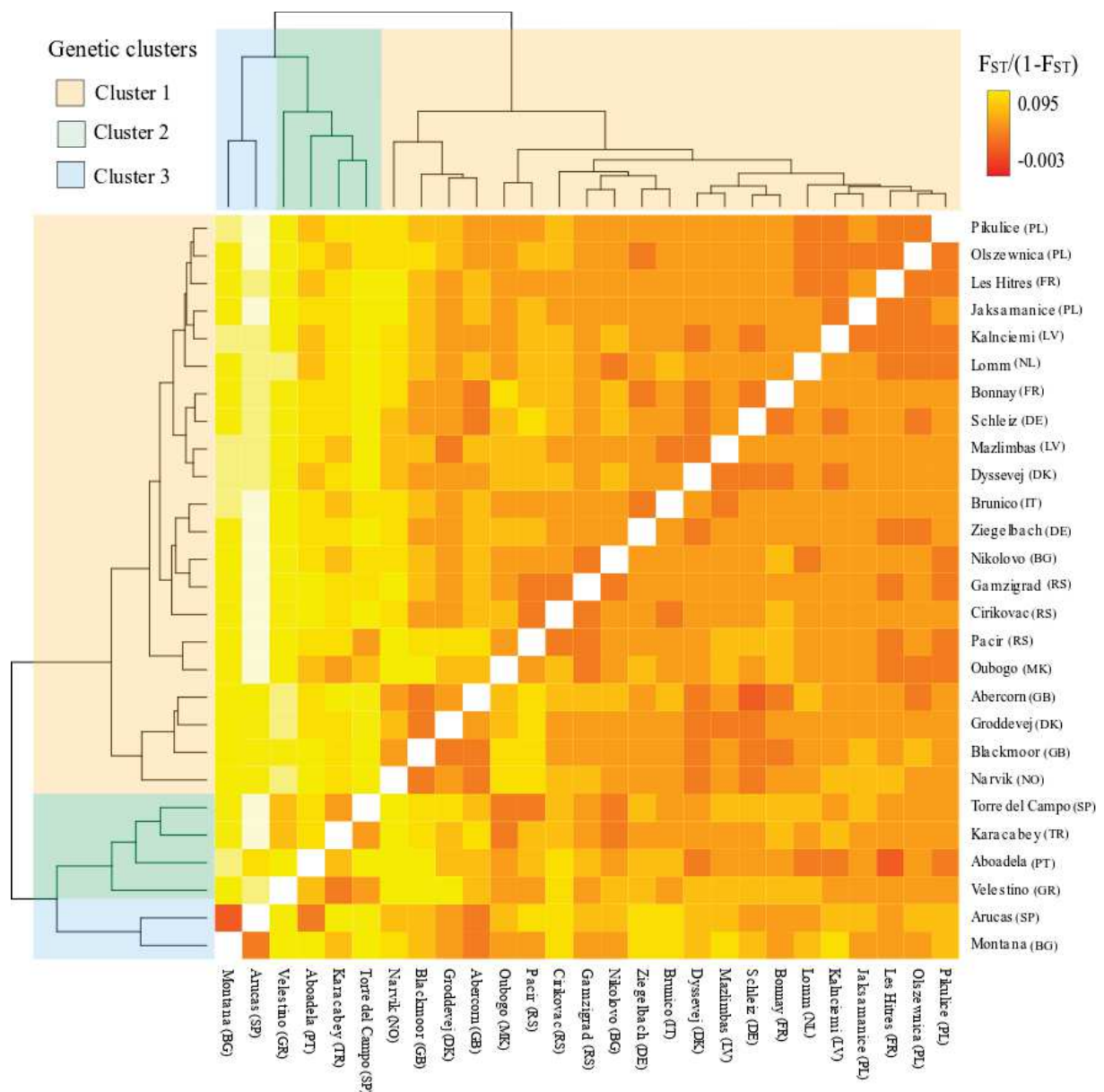


Figure S5. Heatmap of $F_{ST}/(1-F_{ST})$ of all population for microsatellite dataset of *C. obsoletus* in Europe. Levels of genetic distance is represented by color gradient, red mean low level of differentiation and yellow mean high levels of differentiation. The belonging of populations to a genetic cluster defined by the STRUCTURE Analysis is represented by the colors, orange, green and blue for cluster 1, 2 and 3 respectively.

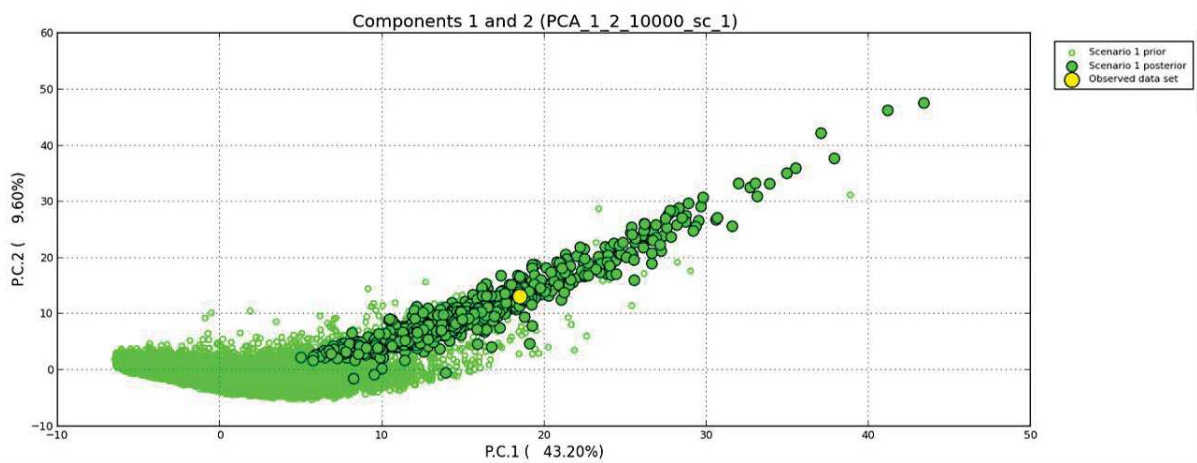


Figure S6. PCA summary statistics simulated under scenario 1, the most probable scenarios. Yellow dot represents the real data set. Large dots correspond to simulated posteriors and the small circles correspond to simulated statistics from the prior statistics

References

- Abbott RJ, James JK, Milne RI, Gillies AC (2003). Plant introductions, hybridization and gene flow. *Philos Trans R Soc Lond B Biol Sci* **358**(1434): 1123-1132.
- Abellan P, Svenning JC (2014). Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biol J Linn Soc* **113**: 13–28.
- Adamack AT, Gruber B, Dray S (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* **5**(4): 384-387.
- Agapow P-M, Burt A (2001). Indices of multilocus linkage disequilibrium. *Mol Ecol Notes* **1**(1-2): 101-102.
- Ander M, Troell K, Chirico J (2013). Barcoding of biting midges in the genus *Culicoides*: a tool for species determination. *Med Vet Entomol* **27**(3): 323-331.
- Anderson GS, Belton P, Kleider N (1991). *Culicoides obsoletus* (Diptera: Ceratopogonidae) as a causal agent of *Culicoides* hypersensitivity (sweet itch) in British Columbia. *J Med Entomol* **28**(5): 685-693.
- Beaumont MA, Rannala B (2004). The Bayesian revolution in genetics. *Nat Rev Genet* **5**(4): 251-261.
- Beaumont MA, Zhang W, Balding DJ (2002). Approximate Bayesian computation in population genetics. *Genetics* **162**(4): 2025-2035.
- Bertorelle G, Benazzo A, Mona S (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Mol Ecol* **19**(13): 2609-2625.
- Bhagwat SA, Willis KJ (2008). Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *J Biogeogr* **35**(3): 464-482.
- Borkent A (2016). World Species of Biting Midges (Diptera: Ceratopogonidae).
- Bourquia M, Garros C, Rakotoarivony I, Gardes L, Huber K, Boukhari I *et al* (2019). Update of the species checklist of *Culicoides* Latreille, 1809 biting midges (Diptera: Ceratopogonidae) of Morocco. *Parasit Vectors* **12**(1): 459.
- Bowler DE, Benton TG (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biol Rev Camb Philos Soc* **80**(2): 205-225.
- Brito PH, Edwards SV (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135**(3): 439-455.
- Brookfield JF (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* **5**(3): 453-455.

- Calvo JH, Calvete C, Martinez-Royo A, Estrada R, Miranda MA, Borrás D *et al* (2009). Variations in the mitochondrial cytochrome c oxidase subunit I gene indicate northward expanding populations of *Culicoides imicola* in Spain. *Bull Entomol Res* **99**(6): 583-591.
- Carpenter S, Groschup MH, Garros C, Felipe-Bauer ML, Purse BV (2013). *Culicoides* biting midges, arboviruses and public health in Europe. *Antiviral Res* **100**(1): 102-113.
- Carpenter S, Wilson A, Mellor PS (2009). *Culicoides* and the emergence of bluetongue virus in northern Europe. *Trends Microbiol* **17**(4): 172-178.
- Chakraborty R, Zhong Y, Jin L, Budowle B (1994). Nondetectability of restriction fragments and independence of DNA fragment sizes within and between loci in RFLP typing of DNA. *Am J Hum Genet* **55**(2): 391-401.
- Cho HC, Chong CS (1974). Notes On Biting Midges Of The Genus *Culicoides* From South Korea: With Special Reference To Unrecorded Species And Distribution. *Kisaengchunghak Chapchi* **12**(1): 45-75.
- Comes HP, Kadereit JW (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends Plant Sci* **3**(11): 432-438.
- Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R *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**(8): 1187-1189.
- Cuellar AC, Kjaer LJ, Kirkeby C, Skovgaard H, Nielsen SA, Stockmarr A *et al* (2018). Spatial and temporal variation in the abundance of *Culicoides* biting midges (Diptera: Ceratopogonidae) in nine European countries. *Parasit Vectors* **11**(1): 112.
- Delecolle JC (1985). Nouvelle contribution à l'étude systématique et iconographique des espèces du genre *Culicoides*, (Diptera) : (Ceratopogonidae) du Nord-Est de la France. Strasbourg.
- Dellicour S, Mardulyn P (2014). spads 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. *Mol Ecol Resour* **14**(3): 647-651.
- Despres L (2019). One, two or more species? Mitonuclear discordance and species delimitation. *Mol Ecol* **28**(17): 3845-3847.
- Dlugosch KM, Parker IM (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* **17**(1): 431-449.
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* **14**(1): 209-214.

- Earl DA, vonHoldt BM (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**(2): 359-361.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**(5): 1792-1797.
- Ellis JS, Knight ME, Darvill B, Goulson D (2006). Extremely low effective population sizes, genetic structuring and reduced genetic diversity in a threatened bumblebee species, *Bombus sylvarum* (Hymenoptera: Apidae). *Mol Ecol* **15**(14): 4375-4386.
- Estoup A, Guillemaud T (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Mol Ecol* **19**(19): 4113-4130.
- Estoup A, Solignac M, Cornuet JM, Goudet J, Scholl A (1996). Genetic differentiation of continental and island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. *Mol Ecol* **5**(1): 19-31.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**(8): 2611-2620.
- Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **491**(1985): 479-491.
- Frank BM, Piccolo JJ, Baret PV (2011). A review of ecological models for brown trout: towards a new demogenetic model. *Ecology of Freshwater Fish* **20**(2): 167-198.
- Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**(2): 915-925.
- Funk DJ, Omland KE (2003). Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annu Rev Ecol Evol Syst* **34**(1): 397-423.
- Gao X, Wang L, Liu T, Xiao J, Wang H (2020). Effect of agro-ecological landscape on the distribution of *Culicoides obsoletus* in northeast China. *Pest Manag Sci*.
- Garros C, Balenghien T (2017). Chapitre 14. Les *Culicoides* (Diptera : Ceratopogonidae). In: Éditions I (ed) *Entomologie médicale et vétérinaire*, pp 345-365.
- Garza JC, Williamson EG (2001). Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* **10**(2): 305-318.
- Goffredo M, Delecolle JC, Semproni G, Lika A (2004). *Culicoides* (Diptera: Ceratopogonidae) in Albania: results of the 2002 entomological survey for bluetongue. *Vet Ital* **40**(3): 209-211.

- Gómez A, Lunt DH (2007). Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Springer (ed) *Phylogeography of Southern European Refugia*
- Goudet J, Raymond M, de Meeus T, Rousset F (1996). Testing differentiation in diploid populations. *Genetics* **144**(4): 1933-1940.
- Goulson D, Kaden JC, Lepais O, Lye GC, Darvill B (2011). Population structure, dispersal and colonization history of the garden bumblebee *Bombus hortorum* in the Western Isles of Scotland. *Conserv Genet* **12**(4): 867-879.
- Hakima B, Hwang HS, Lee KY (2020). Molecular identification of *Culicoides* (Diptera: Ceratopogonidae) species in Algeria. *Acta Trop* **202**: 105261.
- Harrup LE, Laban S, Purse BV, Reddy YK, Reddy YN, Byregowda SM *et al* (2016). DNA barcoding and surveillance sampling strategies for *Culicoides* biting midges (Diptera: Ceratopogonidae) in southern India. *Parasit Vectors* **9**: 461.
- Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biol J Linn Soc* **68**(1-2): 87-112.
- Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* **359**(1442): 183-195; discussion 195.
- Hoffmann B, Eschbaumer M, Beer M (2009). Real-time quantitative reverse transcription-PCR assays specifically detecting bluetongue virus serotypes 1, 6, and 8. *J Clin Microbiol* **47**(9): 2992-2994.
- Hristescu D, Barbuceanu F, Dascalu L, Nitescu C, Goffredo M, Santilli A *et al* (2020). Species composition and relative abundance of the genus *Culicoides* (Diptera: Ceratopogonidae) in Romania. *Parasit Vectors* **13**(1): 393.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* **9**(5): 1322-1332.
- Irwin DE, Bensch S, Irwin JH, Price TD (2005). Speciation by distance in a ring species. *Science* **307**(5708): 414-416.
- Jacquet S, Garros C, Lombaert E, Walton C, Restrepo J, Allene X *et al* (2015). Colonization of the Mediterranean basin by the vector biting midge species *Culicoides imicola*: an old story. *Mol Ecol* **24**(22): 5707-5725.
- Jacquet S, Huber K, Guis H, Setier-Rio ML, Goffredo M, Allene X *et al* (2016a). Spatio-temporal genetic variation of the biting midge vector species *Culicoides imicola* (Ceratopogonidae) Kieffer in France. *Parasit Vectors* **9**: 141.
- Jacquet S, Huber K, Pages N, Talavera S, Burgin LE, Carpenter S *et al* (2016b). Range expansion of the Bluetongue vector, *Culicoides imicola*, in continental France likely due to rare wind-transport events. *Sci Rep* **6**: 27247.

- James JE, Lanfear R, Eyre-Walker A (2016). Molecular Evolutionary Consequences of Island Colonization. *Genome Biol Evol* **8**(6): 1876-1888.
- Kamvar ZN, Tabima JF, Grunwald NJ (2014). Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*.
- Kirk H, Dorn S, Mazzi D (2013). Worldwide population genetic structure of the oriental fruit moth (*Grapholita molesta*), a globally invasive pest. *BMC Ecol* **13**: 12.
- Kirkeby C, Bodker R, Stockmarr A, Lind P, Heegaard PM (2013). Quantifying dispersal of european *Culicoides* (Diptera: Ceratopogonidae) vectors between farms using a novel mark-release-recapture technique. *PLoS One* **8**(4): e61269.
- Kolleck J, Yang M, Zinner D, Roos C (2013). Genetic diversity in endangered Guizhou snub-nosed monkeys (*Rhinopithecus brelichi*): contrasting results from microsatellite and mitochondrial DNA data. *PLoS One* **8**(8): e73647.
- Kosakovsky Pond SL, Frost SD (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* **22**(5): 1208-1222.
- Kozakiewicz CP, Burrige CP, Funk WC, VandeWoude S, Craft ME, Crooks KR *et al* (2018). Pathogens in space: Advancing understanding of pathogen dynamics and disease ecology through landscape genetics. *Evol Appl* **11**(10): 1763-1778.
- Lagercrantz U, Ryman N (1990). Genetic Structure of Norway Spruce (*Picea Abies*): Concordance of Morphological and Allozymic Variation. *Evolution* **44**(1): 38-53.
- Lecocq T, Dellicour S, Michez D, Lhomme P, Vanderplanck M, Valterova I *et al* (2013). Scent of a break-up: phylogeography and reproductive trait divergences in the red-tailed bumblebee (*Bombus lapidarius*). *BMC Evol Biol* **13**: 263.
- Lecocq T, Rasmont P, Harpke A, Schweiger O (2016). Improving International Trade Regulation by Considering Intraspecific Variation for Invasion Risk Assessment of Commercially Traded Species: The *Bombus terrestris* Case. *Conservation Letters* **9**(4): 281-289.
- Lesieur V, Lombaert E, Guillemaud T, Courtial B, Strong W, Roques A *et al* (2018). The rapid spread of *Leptoglossus occidentalis* in Europe: a bridgehead invasion. *J Pest Sci* **92**(1): 189-200.
- Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**(11): 1451-1452.
- Manangwa O, De Meeus T, Grebaut P, Segard A, Byamungu M, Ravel S (2019). Detecting Wahlund effects together with amplification problems: Cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. *Mol Ecol Resour* **19**(3): 757-772.
- Martinez-de la Puente J, Martinez J, Ferraguti M, Morales-de la Nuez A, Castro N, Figuerola J (2012). Genetic characterization and molecular identification of the bloodmeal sources

- of the potential bluetongue vector *Culicoides obsoletus* in the Canary Islands, Spain. *Parasit Vectors* **5**: 147.
- Mathieu B (2011). Les espèces de *Culicoides* du sous-genre *Avaritia* (Diptera : Ceratopogonidae) dans le monde : révision systématique et taxonomique des espèces d'intérêt dans la transmission d'Orbivirus. 253.
- Mayer F, Piel FB, Cassel-Lundhagen A, Kirichenko N, Grumiau L, Okland B *et al* (2015). Comparative multilocus phylogeography of two Palearctic spruce bark beetles: influence of contrasting ecological strategies on genetic variation. *Mol Ecol* **24**(6): 1292-1310.
- Mehlhorn H, Walldorf V, Klimpel S, Jahn B, Jaeger F, Eschweiler J *et al* (2007). First occurrence of *Culicoides obsoletus*-transmitted Bluetongue virus epidemic in Central Europe. *Parasitol Res* **101**(1): 219-228.
- Meiswinkel R, van Rijn P, Leijts P, Goffredo M (2007). Potential new *Culicoides* vector of bluetongue virus in northern Europe. *Vet Rec* **161**(16): 564-565.
- Mellor PS, Carpenter S, Harrup L, Baylis M, Mertens PP (2008). Bluetongue in Europe and the Mediterranean Basin: history of occurrence prior to 2006. *Prev Vet Med* **87**(1-2): 4-20.
- Mignotte A, Garros C, Dellicour S, Gilbert M, Gardès L, Balenghien T *et al* (Submitted in 2020). *Important Dispersal Capacity of One Main Species Vector of Bluetongue and Schmallenberg Viruses Culicoides Obsoletus (Diptera: Ceratopogonidae) Revealed by Landscape Genetic Analyses.*
- Mignotte A, Garros C, Gardes L, Balenghien T, Duhayon M, Rakotoarivony I *et al* (2020). The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palearctic vector *Obsoletus/Scoticus* Complex (Diptera: Ceratopogonidae) at the European level. *Parasites & Vectors* **13**(1): 265.
- Miranda MA, Borrás D, Rincon C, Alemany A (2003). Presence in the Balearic Islands (Spain) of the midges *Culicoides imicola* and *Culicoides obsoletus* group. *Medical and Veterinary Entomology* **17**(1): 52-54.
- Murray MD, Kirkland PD (1995). Bluetongue and Douglas virus activity in New South Wales in 1989 : further evidence for long-distance dispersal of the biting midge *Culicoides brevitarsis*. *Australian Veterinary Journal* **Vol. 72**.
- Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* **8**(7): e1002764.
- Nei M, Maruyama T, Chakraborty R (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution* **29**(1): 1-10.
- Nolan DV, Dallas JF, Pierrney SB, Mordue Luntz AJ (2008). Incursion and range expansion in the bluetongue vector *Culicoides imicola* in the Mediterranean basin: a phylogeographic analysis. *Med Vet Entomol* **22**(4): 340-351.

- Nychka D, Furrer R, Paige J, Sain S (2017). *Fields: Tools for spatial data*.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB *et al* (2013). *Vegan: Community Ecology Package*. R Package Version. 2.0-10. *CRAN*.
- Onyango MG, Beebe NW, Gopurenko D, Bellis G, Nicholas A, Ogugo M *et al* (2015a). Assessment of population genetic structure in the arbovirus vector midge, *Culicoides brevitarsis* (Diptera: Ceratopogonidae), using multi-locus DNA microsatellites. *Vet Res* **46**: 108.
- Pascual M, Chapuis MP, Mestres F, Balanya J, Huey RB, Gilchrist GW *et al* (2007). Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Mol Ecol* **16**(15): 3069-3083.
- Petit R, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R *et al* (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**(5625): 1563-1565.
- Phillips BL, Brown GP, Shine R (2010). Life-history evolution in range-shifting populations. *Ecology* **91**(6): 1617-1627.
- Pili E, Carcangiu L, Oppo M, Marchi A (2010). Genetic structure and population dynamics of the biting midges *Culicoides obsoletus* and *Culicoides scoticus*: implications for the transmission and maintenance of bluetongue. *Med Vet Entomol* **24**(4): 441-448.
- Piry S, Chapuis MP, Gauffre B, Papaix J, Cruaud A, Berthier K (2016). Mapping Averaged Pairwise Information (MAPI): a new exploratory tool to uncover spatial structure. *Methods Ecol Evol* **7**(12): 1463-1475.
- Pons O, Petit R (1996). Measuring and testing genetic differentiation with ordered vs. unordered alleles. *Genetics* **144**: 1237-1245.
- Pons O, Petit RJ (1995). Estimation, variance and optimal sampling of gene diversity : I. Haploid locus. *Theor Appl Genet* **90**(3-4): 462-470.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2): 945-959.
- Provan J, Bennett KD (2008). Phylogeographic insights into cryptic glacial refugia. *Trends Ecol Evol* **23**(10): 564-571.
- Prugnolle F, de Meeus T (2002). Inferring sex-biased dispersal from population genetic tools: a review. *Heredity (Edinb)* **88**(3): 161-165.
- Prugnolle F, Liu H, de Meeus T, Balloux F (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *Int J Parasitol* **35**(3): 255-263.

- Pudar D, Petric D, Allene X, Alten B, Ayhan N, Cvetkovikj A *et al* (2018). An update of the *Culicoides* (Diptera: Ceratopogonidae) checklist for the Balkans. *Parasit Vectors* **11**(1): 462.
- Purse BV, Carpenter S, Venter GJ, Bellis G, Mullens BA (2015). Bionomics of temperate and tropical *Culicoides* midges: knowledge gaps and consequences for transmission of *Culicoides*-borne viruses. *Annu Rev Entomol* **60**: 373-392.
- Purse BV, Mellor PS, Rogers DJ, Samuel AR, Mertens PP, Baylis M (2005). Climate change and the recent emergence of bluetongue in Europe. *Nat Rev Microbiol* **3**(2): 171-181.
- Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**(12): 2092-2100.
- Rasic G, Schama R, Powell R, Maciel-de Freitas R, Endersby-Harshman NM, Filipovic I *et al* (2015). Contrasting genetic structure between mitochondrial and nuclear markers in the dengue fever mosquito from Rio de Janeiro: implications for vector control. *Evol Appl* **8**(9): 901-915.
- Rice WR (1989). Analysing tables of statistic tests. *Evolution* **43**(1): 223-225.
- Rodríguez-Sánchez F, Hampe A, Jordano P, Arroyo J (2010). Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: A review. *Rev Palaeobot Palynol* **162**(3): 507-521.
- Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* **9**(3): 552-569.
- Sarvasova A, Goffredo M, Sopoliga I, Savini G, Kocisova A (2014). *Culicoides* midges (diptera: ceratopogonidae) as vectors of orbiviruses in Slovakia. *Vet Ital* **50**(3): 203-212.
- Savini G, MacLachlan NJ, Sanchez-Vizcaino JM, Zientara S (2008). Vaccines against bluetongue in Europe. *Comp Immunol Microbiol Infect Dis* **31**(2-3): 101-120.
- Schmitt T, Varga Z (2012). Extra-Mediterranean refugia: The rule and not the exception? *Front Zool* **9**(1): 22.
- Sghaier S, Hammami S, Goffredo M, Hammami M, Portanti O, Lorusso A *et al* (2017). New species of the genus *Culicoides* (Diptera Ceratopogonidae) for Tunisia, with detection of Bluetongue viruses in vectors. *Vet Ital* **53**(4): 357-366.
- Shao ZY, Mao HX, Fu WJ, Ono M, Wang DS, Bonizzoni M *et al* (2004). Genetic structure of Asian populations of *Bombus ignitus* (Hymenoptera: Apidae). *J Hered* **95**(1): 46-52.
- Shults P, Ho A, Martin EM, McGregor BL, Vargo EL (2020). Genetic Diversity of *Culicoides stellifer* (Diptera: Ceratopogonidae) in the Southeastern United States Compared With Sequences From Ontario, Canada. *J Med Entomol* **57**(4): 1324-1327.

- Sommer RS, Zachos FE (2009). Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. *J Biogeogr* **36**(11): 2013-2020.
- Stewart JR, Lister AM (2001). Cryptic northern refugia and the origins of the modern biota. *Trends Ecol Evol* **16**(11): 608-613.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* **7**(4): 453-464.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**(3): 585-595.
- Takezaki N, Nei M (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**(1): 389-399.
- Tzedakis PC, Emerson BC, Hewitt GM (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends Ecol Evol* **28**(12): 696-704.
- Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol Ecol* **7**(11): 1569-1573.
- Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**(6): 1358-1370.
- Weir BS, Goudet J (2017). A Unified Characterization of Population Structure and Relatedness. *Genetics* **206**(4): 2085-2103.
- White BP, Pilgrim EM, Boykin LM, Stein ED, Mazor RD (2014). Comparison of four species-delimitation methods applied to a DNA barcode data set of insect larvae for use in routine bioassessment. *Freshwater Science* **33**(1): 338-348.
- White TA, Perkins SE, Heckel G, Searle JB (2013). Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. *Mol Ecol* **22**(11): 2971-2985.
- Whittaker RJ, Field R (2000). Tree species richness modelling: an approach of global applicability? *Oikos* **89**(2): 399-402.
- Widmer A, Lexer C (2001). Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends Ecol Evol* **16**(6).
- Widmer A, Schmid-Hempel P, Estoup A, Scholl A (1998). Population genetic structure and colonization history of *Bombus terrestris* s.l. (Hymenoptera : Apidae) from the Canary Islands and Madeira. *Heredity* **81**: 563-572.
- Wilson AJ, Mellor PS (2009). Bluetongue in Europe: past, present and future. *Philos Trans R Soc Lond B Biol Sci* **364**(1530): 2669-2681.

Winter DJ (2012). MMOD: an R library for the calculation of population differentiation statistics. *Mol Ecol Resour* **12**(6): 1158-1160.

Yang RC (1998). Estimating Hierarchical F-Statistics. *Evolution* **52**(4): 950-956.

Chapitre III : Important dispersal capacity of one main species vector of bluetongue and Schmallenberg viruses *Culicoides obsoletus* (Diptera: Ceratopogonidae) revealed by landscape genetic analyses at French scale

A l'échelle continentale, une faible structuration est observée, témoignage d'une histoire évolutive récente combinée à des flux de gènes importants sur l'aire de distribution ouest paléarctique. Mais quels sont les facteurs environnementaux qui conditionnent la connectivité paysagère des *Culicoides* ?

La dispersion des *Culicoides* est connue pour être stratifiée, c'est-à-dire combinant des processus de dispersion actifs, à courte distance, et passifs ou semi-actifs à longue distance, permettant aux individus de faire des bonds de plusieurs centaines de kilomètres. La dispersion active, principalement motivée par la recherche d'hôtes vertébrés pour un repas de sang ou d'un lieu d'oviposition, est conditionnée par les conditions climatiques et météorologiques qui peuvent diminuer l'activité et la survie des *Culicoides*. La dispersion passive par le vent est, elle, influencée par la direction et la vitesse du vent.

Ce troisième chapitre présente pour la première fois une étude de génétique du paysage appliquée aux *Culicoides*. Elle vise à identifier les facteurs environnementaux qui favorisent ou limitent les flux génétiques de *C. obsoletus*. Nos travaux permettent de quantifier l'impact d'un large éventail de facteurs environnementaux et météorologiques classiquement cités comme impactant la dispersion des *Culicoides*. Un plan d'échantillonnage individuel a permis de génotyper 348 individus issus de 46 sites en France à l'aide de 11 marqueurs microsatellites. Une approche de régression multiple sur les matrices de distance nous a permis d'étudier la corrélation entre les distances génétiques et environnementales. L'utilisation d'un raster nul et d'une approche de commonalité nous a permis de nous défaire de la multicollinéarité de ce type d'analyse et d'identifier des contributions uniques et communes des prédicteurs à la variance de la distance environnementale.

Nos travaux permettent de renforcer nos conclusions quant à une dispersion importante des *Culicoides* au-dessus des terres, comme en témoigne l'absence totale d'isolement par la distance à l'échelle française. Une faible différenciation et structure génétique des populations sont observées dans tout le pays. Alors que la contribution unique de tous les facteurs environnementaux et météorologiques testés est faible, celle de la densité du bétail est la plus forte que nous ayons détectée. L'isolement anisotropique significatif sur un axe nord-sud peut être le signe de la dispersion par le vent des *Culicoides*. Une telle capacité de dispersion du *C. obsoletus* sur les territoires terrestres appelle à une ré-évaluation de l'impact des *Culicoides* sur la dispersion du virus. Ce chapitre a été soumis et est sous révision dans le journal *Parasites and Vectors*. <https://dx.doi.org/10.21203%2Frs.3.rs-75603%2Fv1>

Important dispersal capacity of one main species vector of bluetongue and Schmallenberg viruses *Culicoides obsoletus* (Diptera: Ceratopogonidae) revealed by landscape genetic analyses at French scale

Antoine Mignotte^{1,2*}, Claire Garros^{1,2}, Simon Dellicour^{3,4}, Maude Jacquot^{3,5}, Marius Gilbert³, Laetitia Gardès^{1,6}, Thomas Balenghien^{1,7,8}, Maxime Duhayon^{1,2}, Ignace Rakotoarivony^{1,2}, Maïa de Wavrechin^{1,2}, Karine Huber¹.

¹ ASTRE, Univ Montpellier, Cirad, INRAE, Montpellier, France.

² Cirad, UMR ASTRE, F-34398 Montpellier, France.

³ Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, CP160/12, 50, av. FD Roosevelt, 1050 Bruxelles, Belgium

⁴ Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

⁵ Université Clermont Auvergne, INRAE, VetAgro Sup, UMR EPIA, F-63122 Saint-Genès-Champanelle, France

⁶ Cirad, UMR ASTRE, F- 97170 Petit-Bourg, Guadeloupe, France.

⁷ Cirad, UMR ASTRE, M- 10100 Rabat, Morocco.

⁸ Institut Agronomique et Vétérinaire Hassan II, Unité Microbiologie, immunologie et maladies contagieuses, 10100 Rabat-Instituts, Morocco.

* Correspondence: antoine.mignotte@cirad.fr, claire.garros@cirad.fr and karine.huber@cirad.fr

Emails:

AM: antoine.mignotte@cirad.fr

CG: claire.garros@cirad.fr

SD : sdellico@ulb.ac.be

MJ : maude.jacquot@inrae.fr

MG : marius.gilbert@ulb.ac.be

LG: laetitia.gardes@cirad.fr

TB: thomas.balenghien@cirad.fr

MD: maxime.duhayon@cirad.fr

IR: ignace.rakotoarivony@cirad.fr

MW: maiadewav@gmail.com

KH: karine.huber@cirad.fr

Abstract

Background: In the last two decades, recurrent epizootics of bluetongue virus and Schmallenberg virus have been reported in the western Palearctic region. These viruses affect domestic cattle, sheep, goats and wild ruminants and are transmitted by native hematophagous midges of the genus *Culicoides* (Diptera: Ceratopogonidae). *Culicoides* dispersal is known to be stratified, *i.e.* due to a combination of dispersal processes occurring actively at short distances and passively or semi-actively at long distances, allowing individuals to jump hundreds of kilometers.

Methods: Here, we aim to identify the environmental factors that promote or limit gene flows of *Culicoides obsoletus*, an abundant and widespread vector species in Europe using an innovative framework integrating spatial, population genetics and statistical approaches. A total of 348 individuals were sampled in 46 sites in France and have been genotyped using 11 newly-designed microsatellite markers.

Results: We found a low genetic differentiation and population structure for *C. obsoletus* across the country. Using three complementary inter-individual genetic distances, we did not detect any significant isolation-by-distance, but a significant anisotropic isolation-by-distance on a North-South axis. We have employed a multiple regression approach on distance matrices to investigate the correlation between genetic and environmental distances. Among all the environmental factors that have been tested, only cattle density seems to have an impact on *C. obsoletus* gene flow.

Conclusions: Such a high dispersal capacity of *C. obsoletus* over land territories calls for a re-evaluation of the impact of *Culicoides* on virus dispersal and highlights the urgent need to better integrate molecular, spatial and statistical information to guide vector-borne diseases control.

Keywords: *Culicoides obsoletus*, landscape genetics, microsatellites, dispersal, Palearctic region

Background

During a vector-borne disease outbreak, a precise understanding of the dispersal capacity of the vector species is a key aspect to implement appropriate control strategies in order to limit the spread of the disease. Dipteran dispersal is known to be impacted by the composition and configuration of the landscape. When the dispersal of an organism is favored by a landscape characteristic (“environmental factor”) it can be categorized as a conductance factor (Manel and Holderegger, 2013). On the contrary, when the dispersal is limited by a landscape characteristic, it can be categorized as a resistance factor (Manel *et al*, 2007; Manel and Holderegger, 2013). However, some types of dispersal allow overcoming resistance factor, as it is the case for “stratified dispersion”. *Culicoides* dispersal is described as “stratified”, due to a combination of dispersal processes occurring actively at short distances, and passively or semi actively at long distances (Murray and Kirkland, 1995).

Northern Europe experienced very sudden and rapid outbreaks of bluetongue virus (BTV) in 2006-2008 and of Schmallenberg virus (SBV) in 2011-2012 (Carpenter *et al*, 2013). Both viruses have spread very quickly and massively across the whole western Palearctic region transmitted by native *Culicoides* species. *Culicoides obsoletus* has been identified as the main vector species responsible for the transmission of BTV and SBV to wild and domestic ruminants in western Palearctic region (Carpenter *et al*, 2009; Hoffmann *et al*, 2009; Savini *et al*, 2008). *Culicoides obsoletus* is a very widespread species able to breed in a wide range of habitats (Mehlhorn *et al*, 2007; Purse *et al*, 2015). BTV spread is facilitated by favorable conditions for midge activity and viral replication and the environmental and climatic drivers of BTV transmission in Europe have been suggested (Cuellar *et al*, 2020; Guis *et al*, 2012; Jacquot *et al*, 2017; Maclachlan and Guthrie, 2010; Purse *et al*, 2005). For example, high ambient temperature reduce incubation period of the virus (Wittmann *et al*, 2002) whereas high precipitation and wind speeds can reduced the *Culicoides* flight activity (Baylis *et al*, 2010; Carpenter *et al*, 2008c; Gloster *et al*, 2008).

Previous studies of mark-release-recapture on *Culicoides* species showed that the post-release dispersal distance travelled, during two nights, ranges from 1 to 2.5 km, and is linked to the gradual search for hosts or oviposition sites (Kluiters *et al*, 2015b). The maximum recapture travelled distance recorded was 6 km for *Culicoides mohave* in a particular desert landscape (Brenner *et al*, 1984). Yet, the potentially high dispersal capacity of *Culicoides* and high mortality of adults when manipulating represents a practical limitation in the context of a mark-

release-recapture procedure. Dispersal phenomena driven by wind currents on a much larger scale have also been detected during outbreaks related to *Culicoides*-borne diseases (Alba *et al*, 2004; Ducheyne *et al*, 2007; Eagles *et al*, 2012; Gloster *et al*, 2008; Hendrickx *et al*, 2008; Jacquet *et al*, 2016b; Murray, 1987; Murray and Kirkland, 1995). Introduction of BTV serotypes by wind-borne infected *Culicoides* has been demonstrated from Northern Africa to Southern Europe (Gubbins *et al*, 2008), from Kenya to South-Western Indian Ocean Islands (Onyango *et al*, 2015a), from Sardinia to the Balearic Islands (Alba *et al*, 2004), from France to Corsica (Jacquet *et al*, 2016a), in Ireland (McGrath *et al*, 2018), or in UK (Burgin *et al*, 2017). Most of these studies are supported by modelling analyses in which dispersal trajectories have been evaluated by atmospheric dispersion models over water bodies (Agren *et al*, 2010; Aguilar-Vega *et al*, 2019; Andrew Jones, 2007; Durr *et al*, 2017; Eagles *et al*, 2014; McGrath *et al*, 2018). *Culicoides* dispersal over land territories has been less investigated (Burgin *et al*, 2013) and little attempt has been made to link it to environmental factors. Given this high dispersal capacity and the stratified dispersal pattern of *Culicoides*, it is crucial to determine inland connectivity among populations and to identify the potential environmental factors that promote or limit gene flows between them.

Classic approaches to study the impact of landscape on gene flow generally use population-based sampling and are often based on a limited number of sampled populations (Jacquet *et al*, 2015; Onyango *et al*, 2015b). However, *Culicoides* such as other dipterans species, are not spatially structured into separate populations and must be considered as a continuum of individuals heterogeneously distributed across landscape. In order to identify the environmental factors having an impact on the dispersal of *Culicoides*, an individual approach is more relevant as it avoids inter-population genetic distance misinterpretations (Shirk *et al*, 2017). Indeed, the individual approach in landscape genetics aims at maximizing the number of sampling sites, and thus brings a much greater statistical power to detect spatial patterns of genetic differentiation and the environmental factors that cause them (Landguth *et al*, 2010). In addition to the analysis of inter-individual genetic distances, population genetic structure can also be investigated, *e.g.* using gold-standard Bayesian clustering methods. However, if isolation-by-distance occurs, these clustering analyses could overestimate the actual number of genetic clusters (Frantz *et al*, 2009). It is therefore important to also incorporate a visualization approach into the analytical workflow (Piry *et al*, 2016). MAPI allow to visually compare genetic dissimilarity with some environmental factors but also to develop working hypotheses.

On the other side, it is also necessary to use as a complement statistical analyses of genetic and environmental distances.

The present study here aims at determining population inland connectivity at large geographical scales of a main vector species and to identify the environmental factors that promote or limit gene flows between *Culicoides* populations in France. For this purpose, we characterized 13 microsatellite markers dedicated to *C. obsoletus*. We are proposing a complementary framework integrating multiple approaches which can be applied more generally to the study of gene flow and links with environmental factors.

Methods

***Culicoides obsoletus* sampling and morphological identification**

The data set analysed in the present study is composed of 368 individual females biting midges collected in 46 sites in France in April 2011, using the national surveillance network for *Culicoides* populations or local complementary collections (Table S1). Collections were made overnight using Onderstepoort Veterinary Institute (OVI) light traps set up at farms near cattle or sheep. All insects were stored in 70% ethanol. Morphological identification to species level has been carried out in the same way as described in a previous study (Mignotte *et al.*, 2020) under a binocular microscope using the available identification keys (Delecolle, 1985).

DNA extraction, amplification, sequencing and sequence analyses

A total of 368 individuals belonging to the *C. obsoletus* species complex were DNA-extracted, *cox1* amplified and sequenced following the procedure detailed in Mignotte *et al.* (2020). To ensure that only *C. obsoletus* individuals were included in the final data set (Garros *et al.*, 2014), sequences assignment was run according to Mignotte *et al.* (2020). The reference *cox1* sequences used to make a specific assignment of all individuals are available in Table S2. All *cox1* sequences were then aligned using the MUSCLE algorithm (Edgar, 2004) implemented in the software package GENEIOUS v.6.0.5 (Biomatters, <http://www.geneious.com>). In order to estimate unbiased gene flows of *C. obsoletus*, only individuals identified as such were used for further analyses.

Development of microsatellite markers

Microsatellite markers were defined by next-generation sequencing following a similar protocol to the one described by Morillo *et al.* (2016). *Culicoides obsoletus* samples from France and United Kingdom (Table S3) were used as biological material for the DNA library. Among the 82 276 simple sequence repeats (SSRs) identified, primers were designed for 43 676 SSRs, of which 15 888 were dinucleotides, 23 656 were trinucleotides, and 4,132 contained complex SSR motifs. For the SSR screening, we selected 3 135 primer pairs flanking dinucleotide SSR motifs with a minimum of 12 repetitions and amplifying fragments between 150 and 300 base pairs (bp) in length and a half denaturation temperature (T_m) close to 55 °C. We screened a set of 30 primer pairs and optimized both polymerase chain reaction (PCR) and multiplex PCR conditions.

Microsatellite genotyping and loci filtering

Microsatellite makers were amplified by multiplex PCR with the Type-it-Microsatellites kit (Qiagen, Valencia, CA, USA) according to the protocol described in the manufacturer's manual and the annealing temperature given in Table S4. Standard conditions for PCR amplification included an initial denaturation step at 95 °C for 5 min, 35 cycles of denaturation for 30s at 95 °C, annealing for 1 min at variable temperature (Table S4) and elongation for 1 min at 72 °C, followed by a final elongation of 5 min at 72°C. Fragments were separated on an Applied Biosystems 3500xL Genetic Analyzer. Allelic size allocations for all individuals and microsatellite markers were performed using the program GeneMapper® version 5 (Life Technologies, Applied Biosystem) with double blind reading to limit the potential interpretation bias from the reader.

A first validation of the polymorphism of these 30 markers was carried out by amplifying and genotyping 48 individuals from various locations within the species geographic range. Then 13 loci showing good genetic profiles and clear allelic size variability were characterized as polymorphic markers and were selected for further analyses (Table S4).

The 368 *C. obsoletus* previously extracted were then amplified and genotyped at these 13 microsatellite loci. Genetic variability parameters, such as the number of alleles per locus, the allele size range, the observed and expected heterozygosity were estimated per locus, and per population over the entire dataset using the R packages *ggplot* (Ginestet, 2011), *poppr* (Kamvar *et al*, 2014) and *mmod* (Winter, 2012). Short allele dominance could be a source of heterozygote deficiency in a microsatellites data set (Manangwa *et al*, 2019; Wattier *et al*, 1998). In order to avoid such bias, Spearman's correlations rank between the size of the alleles and their frequencies were calculated in R (Wattier *et al*, 1998). Null allele frequencies were estimated using the R package *PopGenReport* (Adamack *et al*, 2014), and a Bonferroni's correction was applied to all matched tests to take into account potential biases associated with multiple comparisons (Rice, 1989). Correlation between F_{IS} and null allele was calculated in R, and locus with significative correlation were removed. The presence of an imbalance in genetic binding between each pair of loci has been tested by accurate Fisher tests performed with the R package *poppr* (Kamvar *et al*, 2014). The same R package was used along with the *pegas* R package to identify a potential gap in the panmictic matching regime. For this purpose, we performed a chi-square tests to compare the observed heterozygosity deficits (H_o) to the expected heterozygosity (H_e) under the Hardy-Weinberg (HW) equilibrium with a significance threshold of 5%. We also calculated the fixation index (F_{IS}) (Weir and Cockerham, 1984).

Genetic clustering analysis

In order to determine the most likely number of genetic clusters (K) and the probability of assignment of each individual to these clusters, we used the Bayesian clustering methods implemented in the program STRUCTURE (version 2.3.4) (Pritchard *et al.*, 2000) and GENELAND (version 4.5.0) (Guillot *et al.*, 2005). The algorithm implemented in STRUCTURE infers genetic clusters that minimize deviations from Hardy-Weinberg equilibrium. In STRUCTURE, we specified the following settings: correlated allele frequency, an admixture model, and a “locprior” model. These options respectively allow to: assume allele frequencies are similar among populations, estimate for each individual the belonging proportion to each K clusters, and use sampling locations as prior information to assist the clustering, which is particularly adapted when there is a weak genetic structuring (Hubisz *et al.*, 2009). We performed 10 independent runs for each value of K varying from 1 to 10. Each run consisted in a burn-in of 100,000 iterations in the Markov Monte Carlo chain (MCMC) and a stationary phase of 1,000,000 iterations in the MCMC. The most probable number of clusters was inferred with the ΔK method (Evanno *et al.* 2005), *via* the STRUCTURE Harvester web platform (Earl and vonHoldt, 2011). This ΔK value is a way to determine the inflection in the [Ln P(D)] curve. The clustering resulting from the Bayesian inference was transposed into percentages of assignment of each individual to the K inferred clusters and plotted on a map.

GENELAND also uses a Bayesian algorithm to infer population genetic clusters while taking into account the spatial position of individuals, making it a spatially-explicit clustering method. The most probable number of clusters was also determined by running the algorithm with K ranging from 1 to 10. The analysis was based on 1,000,000 MCMC iterations with a thinning of 1000, maximum rate of the Poisson process fixed to 100, maximum number of nuclei in the Poisson-Voronoi tessellation fixed to 300 and a burn-in of 100. We used the R package *graphics* to produce a distribution map of genetic structure resulting of STRUCTURE and GENELAND analyses.

Computation of genetic distances and mapping genetic dissimilarity

We computed three complementary inter-individual genetic distances. The first inter-individual genetic distance is based on a Factorial Correspondence Analysis (FCA) performed with the program GENETIX v. 4.05.2 (Belkhir, 2004). The first ten FCA axes were used to calculate an Euclidean distance between all individuals using the R package *Ecodist* (Goslee and Urban,

2007), hereafter called the “FCA” distance (Kimmig *et al*, 2020). The second inter-individual genetic distance is the Rousset’s distance (a_R), (Rousset, 2000), corresponding to the $F_{ST}/(1 - F_{ST})$ but for pairs of individuals. We used the SPAGeDi 1.5 program (Hardy and Vekemans, 2002) to compute a_R genetic distance. The third inter-individual genetic metric is the Loiselle’s kinship coefficient (“LKC”, (Loiselle *et al*, 1995)), which, unlike FCA and a_R that are measures of genetic dissimilarity, is an index of similarity (Hardy, 2003). We used the R package *fields* (Nychka *et al*, 2017) to investigate isolation-by-distance by performing linear regressions and Mantel tests between inter-individual genetic distances and log-transformed geographic distances with R package *fields* (Nychka *et al*, 2017). Mantel tests (Mantel, 1967) were performed with the R package *vegan* (Oksanen *et al*, 2013). Inter-individual genetic distances were mapped using the program MAPI (Piry *et al*, 2016). The method implemented in MAPI allows to visualize spatial variation in genetic dissimilarity. It can also define areas where genetic dissimilarity is significantly lower or higher than expected by chance (Piry *et al*, 2016). MAPI analyses were based on 1,000 permutations and an α value of 0.05.

Impact of environmental factors on genetic differentiation

We investigated the potential impact of several environmental factors on the genetic differentiation of *C. obsoletus*. In practice, we tested the association between matrices of pairwise genetic distances (see above) and matrices of environmental distances. These environmental distances were computed using the program CIRCUITSCAPE 4.0.5 (McRae and Shah, 2008; McRae, 2006) and were based on distinct environmental rasters (*i.e.* geo-referenced grids of environmental values) either treated as potential resistance (R) or conductance (C) factors:

- i) host densities: European index of distribution of roe deer (C) (Neil *et al*, 2014), European index of distribution of red deer (C) (Neil *et al*, 2014), global cattle, sheep and goat distributions in 2010 (C) (Gilbert *et al*, 2018).
- ii) terrestrial habitat: mean global elevation in 2010 (R) (Carabajal *et al*, 2011) and type of landscape cover (urban area, grassland, forest areas and croplands) (R). The four distinct land cover rasters (resolution: ~1000 m) were generated from the Corine Land Cover 2012 raster (CLC; www.eea.europa.eu; resolution: ~100 m) (Fig. S1).
- iii) meteorological and climate data: mean surface temperature (R), mean precipitation (R), and mean wind speed and direction (C). We used two

alternative rasters for each variable: April 2011 mean and 2000-2010 mean. Surface temperatures were obtained from monthly day and night land surface temperature mean (Wan, 2014). Precipitation and wind rasters were obtained from monthly means of precipitation, wind speed and the U and V components of wind (components of the horizontal wind towards the east and North) from the European Centre for Medium-range Weather Forecasts (Era5 Interim, <http://www.ecmwf.int>) over 10 years (2001-2010). Wind U and V components were used to compute prevailing meteorological wind direction (in degrees). Map and arrowheads indicating the mean wind direction over the 10 years at each raster pixel were generated with the R package *Rwind* (Fig. S2) (Fernández-López and Schliep, 2018).

A so-called “null raster” was also created and used as a negative control, *i.e.* a raster on which environmental distances computations correspond to a proxy of geographical distances. The null raster is a uniform raster with all cell values equal to 1. As in Dellicour *et al.* (2019), three values of k (10, 100 and 1000) were used to modify the potential impact of the environment on the resistance/conductance value. We used multiple regressions on distance matrices (MRDM) coupled with commonality analyses (CA; (Newton and Spurrell, 1967)) to identify unique and common contributions of predictors to the variance in the environmental distance (response variable). All the analyses were carried out with the R packages *ecodist* (Goslee and Urban, 2007) and *yhat* (Nimon *et al.*, 2008). In addition, we also performed univariate analyses by comparing (i) the determination coefficient R^2 obtained from the linear regression of genetic distances on distances computed on the environmental raster and (ii) the determination coefficient R^2 obtained from the linear regression of genetic distances on distances computed on the null raster. Only the environmental factors associated with a R^2 higher than that the one obtained from the linear regression based on environmental distances computed on the null raster were selected for the multivariate analyses (MRDM-CA) (Dellicour *et al.*, 2019). The same criterions to identify a suppressors were used for the multivariate analysis as in the commonality analyses descriptions [79, 80]. These suppressors allow to select environmental variables that are more explanatory than the geographical distance alone. An environmental variable is considered to be a suppressor if regression coefficient and correlation, or when the unique contribution and the common contribution are of opposite sign (Prunier *et al.*, 2017).

Anisotropic isolation-by-distance

In order to identify directional gene flows, we must calculate the projected distance between two sampling sites. The angle between sites was calculated using the R package *geosphere* (Hijmans, 2019) and transformed from degrees to radians. The set of angles between 0 and 360 were then tested as angles that maximize isolation-by-distance. The great-circle geographic distance between sampling sites was computed with the R package *fields* and then log-transformed (Nychka *et al*, 2017). The projected distance matrix was calculated for each angle between all sampling sites, using the formula below. In calculation of the projected geographical distance as a function of the angle between populations, d_{AB} is the geographical distance between population A and B and a_{AB} is the angle between populations A and B.

$$\log (| d_{AB} * \cos(a_{AB}) |)$$

A linear regression of the genetic distances on the projected geographical distances obtained for each angle was then performed. The angle that will maximize the R^2 of this regression with a positive regression coefficient was considered as the angle maximizing the isolation-by-distance signal (Jay *et al*, 2013). In order to confirm this analysis with a similar but complementary approach, we evaluated the direction that maximizes the correlation between geographic and genetic distance using the PASSAGE program and a "Bearing" analysis. The workflow summarizing the genetic analyses carried-out is described in Figure S3.

Results

Characterization of microsatellite markers and genetic diversity

A total of 368 *cox1* sequences of *C. obsoletus* were obtained after the DNA barcoding step (Table S3) and 368 *C. obsoletus* from 46 populations were genotyped with 13 microsatellite loci (Table S4). A filter has also been applied to exclude missing data. Individuals with more than 10% missing data were excluded from the analysis, i.e. 20 individuals. Loci OBSms25 and OBSms26 had the highest proportion of null alleles with 56% and 22% respectively, but also the highest F_{IS} values of 0.691 and 0.357 respectively (Table S5). The expected heterozygosity, which varies from 0.918 (OBSms5) to 0.404 (OBSms29) is always higher than that observed, varying from 0.872 (OBSms13) to 0.197 (OBSms25), reflecting a heterozygous deficit (Table S5). To ensure that all markers provide independent information, linkage disequilibrium between each pair of loci was tested. After Bonferroni correction, no significant linkage disequilibrium was observed between all loci. The OBSms25 locus was also the only locus with more than 5% missing data. Thus, only markers with a null percentage inferior to 10% were kept for the final analysis. In view of the high proportion of null alleles and the strong correlation between F_{IS} and the rate of null alleles ($r=0.947$, $p\text{-value}>0.001$), the OBSms25 and OBSms26 loci were excluded from the analysis. No significant correlation between allele size and allele frequency was found for any of the 11 remaining loci (Table S5). The genetic resolution of microsatellite markers depends on the number of markers and their polymorphism. All selected 11 loci were therefore polymorphic (Table S5) and the number of alleles present at a given locus varied from 40 for OBSms5 to 9 for OBSms29. No significant isolation-by-distance was detected with the different inter-individual genetic distances, a_R ($R^2<0.001$, $p=0.439$, Fig. S4), LKC ($R^2<0.001$, $p=0.992$, Fig. S4), and FCA ($R^2<0.001$, $p=0.361$, Fig. S4).

Clustering analyses and mapping of genetic distances

The Bayesian analysis performed with STRUCTURE revealed a low genetic structuring of *C. obsoletus*. Meanwhile, the exact number of genetic clusters was not clearly defined, the optimal value of K was 2 (Fig. S5); i.e. the co-existence of two genetic clusters was the most likely statistically (Fig. 1). The assignment of individuals to the different clusters revealed a weak genetic structuring of the populations. Although the assignment of a total of 284 out of the 348 *Culicoides* individuals to a genetic cluster is well supported ($Q>0.90$), no clear association between spatial clustering and distribution can be observed, i.e. there is no association between

Culicoides sampling location and the genetic clusters distributions. However, a West-East transect seems to emerge as samples collected on this transect were more frequently assigned to cluster 1, contrary to the North-South transect more represented by cluster 2. Genetic clusters inferred by GENELAND and STRUCTURE (K=2) are consistent with these results, both clusters being found at all the sampling sites (Fig. 1 ab). Results for the second optimal K value (K=4) of the STRUCTURE program gave a similar view, reinforcing the conclusion that populations are genetically weakly structured (Fig. S6). Additional STRUCTURE analysis has been carried out with only loci OBSms 4, 11 and 28 which are at Hardy Weinberg's equilibrium and no additional structure is detected using only this loci (data not shown). Genetic dissimilarity result from MAPI analyses with the a_R genetic distance showed a non-homogeneous pattern of dissimilarity. The maximum inter-individual genetic dissimilarity was of 0.182 and the lowest was of 0.002. The South of France was the most genetically dissimilar region. The center of France was also a zone of strong inter-individual dissimilarity, while eastern France showed low genetic dissimilarity. Another area spanning from the Paris region to Alsace (North east of France) showed low genetic dissimilarity (Fig. 1 cd). MAPI analyses with FCA inter-individual genetic distance, showed a homogeneous genetic dissimilarity. However, almost none of the areas was significantly genetically dissimilar in MAPI analysis with the three genetic distances used.

Investigating the impact of environmental factors

Univariate regressions revealed a significant positive association of FCA genetic distances with cattle distributions for $k=100$ (Q=0.009, p-value=0.001) and $k=1000$ (Q=0.029, p-value=0.001), with sheep distributions for $k=100$ (Q=0.002, p-value=0.002) and for $k=1000$ (Q=0.006, p-value=0.001), with elevation for $k=100$ (Q=0.001, p-value=0.005) and for $k=1000$ (Q=0.001, p-value=0.006), and with grassland for $k=100$ (Q=0.009, p-value=0.001) and for $k=1000$ (Q=0.011, p-value=0.001) distances (Table S6). Similarly, a_R inter-individual genetic distances were significantly associated with distances computed from elevation for $k=100$ (Q=0.00018, p-value=0.006) and for $k=1000$ (Q=0.00038, p-value=0.007), with grassland for $k=100$ (Q=0.001, p-value=0.001) and for $k=1000$ (Q=0.002, p-value=0.001) and with urban areas for $k=100$ (Q=0.004, p-value=0.001) and for $k=1000$ (Q=0.007, p-value=0.001) rasters (Table S6). However, no significant association was found between “LKC” genetic distance and the environmental distances (Q values always <0).

Multivariate regressions revealed a significant positive association of FCA inter-individual genetic distance with cattle distributions for $k=1000$ ($r=0.1733$, $p\text{-value} < 0.001$) and with grassland for $k=1000$ ($r=0.1080$, $p\text{-value} < 0.001$) (Table S6). Also, multivariate regressions did not reveal any significant positive association of a_R inter-individual genetic distances with environmental distances. All environmental factors were considered as suppressors when using a_R to measure inter-individual genetic distances. Cattle distribution ($k=1000$) with FCA inter-individual genetic distance was the only environmental factor that showed a unique contribution to the variance in the dependent variable higher than 1% ($U=0.02$) (Table S6).

Anisotropic isolation-by-distance

Globally, correlation between the *genetic distance* and the distance along the bearing θ , d_θ , changed as a function of bearing θ . Anisotropic isolation-by-distance analyses revealed a significant positive correlation between a_R and FCA inter-individual genetic distances and orientational distances computed along North–South orientation and a significant negative correlation between genetic distances and orientational distances computed along a West-East orientation (Fig. 2 a,b). For the analysis with LKC, which is a variable of genetic similarity, the North–South was the orientation that maximized the isolation-by-distance (Fig. 2 c). Bearing analyses with PASSAGE software revealed a significant distance isolation on the North/South axis with the genetic distances a_R and FCA (Fig. 3 a,b). A negative correlation between LKC genetic similarity and geographical distance was observed on the North/South axis (Fig. 3 c).

Discussion

This study presented for the first time an extensive landscape genetic study carried out on a main vector species of interest for livestock diseases, *C. obsoletus*. Bayesian clustering analysis revealed a very weak genetic structure of *C. obsoletus* in France and a low level of inter-individual genetic differentiation was observed. We assessed the impact of a wide range of environmental factors on this pattern. Univariate and multivariate analyses highlighted an absence, or a weak impact of most of the tested environmental factors on inter-individual measures of pairwise genetic differentiation. No isolation-by-distance pattern was detected at the French scale with the set of inter-individual genetic distances used. However, an anisotropic isolation-by-distance analysis revealed a significant distance isolation on a North-South axis.

Important gene flows between *C. obsoletus* populations

The low overall inter-individual genetic dissimilarity highlighted by our study reflected the important level of gene flow for *C. obsoletus* and comes to reinforce what has been described previously for other *Culicoides* species. In particular, important level of gene flows have already been observed in France for *C. imicola* (Jacquet *et al*, 2016a) and in Australia for *C. brevitarsis* (Onyango *et al*, 2015a). Genetic studies on *C. imicola* in Europe revealed an important gene flow as reflected by the inference of two large genetic clusters: a “Central Mediterranean cluster” including Algeria, Sardinia, Corsica, Pyrénées-Orientales and Var French departments of France, and a “Western Mediterranean cluster” including Morocco, Spain, Portugal and Majorca (Jacquet *et al*, 2016b). In North America, similar results were also found in *C. stellifer*. No barriers to gene flow could be identified in this species in the southeast United States (Shults *et al*, 2020). In addition, no isolation-by-distance has been observed at the French scale. Identifying a genetic structure pattern is a real challenge when inter-individual genetic dissimilarity is low. Although STRUCTURE performs well at low levels of population differentiation ($0.02 < F_{ST} < 0.10$) by using prior population and correlated allele frequency models (Latch *et al*, 2006), when the differentiation is weaker, such as in the case of highly dispersed organisms, the latter may encounter difficulties (Duchesne and Turgeon, 2012). This underlines the importance of using complementary tools of visualization approaches like MAPI. In addition, the use of an individual approach avoids to artificially consider individuals grouped into populations solely due to the effect of the trapping sites (Manel *et al*, 2007; Prunier *et al*, 2013a). Although our method of analysis seems complementary and coherent to detect a genetic structure, *C. obsoletus* is not genetically or geographically structured at the French scale.

Livestock densities as potential drivers of *C. obsoletus* gene flow

It is undeniable that large gene flows homogenize the genetic diversity of *Culicoides*. This can be explained by active dispersal and host-seeking movements, as suggested by our results. While the unique contribution of all environmental factors tested was very small, the one of cattle density is the strongest we have detected. Host density as a conductance factor for *Culicoides* is obviously a factor consistent with the biology of the species. These results are in line with those obtained in landscape genetics studies of the BTV, which had identified distributions of cattle and sheep as a key factors in BTV dispersal (Jacquot *et al*, 2017). In addition, previous studies showed that dairy cattle density is negatively correlated with BTV spread. Although paradoxical with the previous conclusions, this could be explained by the fact that dairy cattle are clustered around the milking parlour and move little, forming a fixed feed source, limiting the spread of *Culicoides*. On the contrary, beef cattle disperse much more in the pastures and by this phenomenon could encourage the diffusion of *Culicoides* in search of a source of blood (Pioz *et al*, 2012). In view of these converging results, of the marked preference of certain species of *Culicoides* vector for cattle (Viennet *et al*, 2013), and of BTV emergence or reemergence events in cattle in the Netherlands in 2006 and in France in 2015 (Elbers *et al*, 2008; Sailleau *et al*, 2017), it seems crucial to closely monitor the surveillance of *Culicoides* in the vicinity of beef cattle farms (Viennet *et al*, 2013).

In addition, it has been shown that BTV spread is facilitated at low elevation, up to 300 m (Jacquet *et al*, 2015; Pioz *et al*, 2012). Altitude, an environmental factor tested in this study, failed to explain the inter-individual genetic differentiation of *C. obsoletus*. The altitude therefore does not act as a barrier between sampling sites located at relatively low altitude. It would thus be interesting to integrate in future studies sampling sites with a high altitude above 1,000 m, which is possible for *C. obsoletus* highlighting a large altitudinal range (Diaz-Sanchez *et al*, 2018). However, it is possible that the altitude does not directly impact the dispersal activity of *Culicoides*, but only the replication or viral infection of BTV due to low temperature. *Culicoides obsoletus* is extremely generalist and can take a blood meal on a wide range of hosts (Elbers and Meiswinkel, 2014). The study of phylogenetically close but ecologically very different species of *Culicoides*, such as *C. chiopterus* (Ayllon *et al*, 2014), might then allow the identification of very different dispersal patterns and an importance of bovine density, the latter being exclusively dependent on it for egg-laying (Dijkstra *et al*, 2008).

Long-dispersal of *C. obsoletus*

The most genetically dissimilar individuals are mainly located in the southernmost populations of the sampling area. Multiple non-exclusive lines of argument might explain the significant anisotropic isolation-by-distance observed on the North/South axis in France.

First, the significant anisotropic isolation-by-distance could be explained by the wind dispersal. Dispersion phenomena caused by wind currents are already established for *Culicoides*, mainly over the seas (Alba *et al*, 2004; Ducheyne *et al*, 2007; Eagles *et al*, 2012; Gloster *et al*, 2008; Hendrickx *et al*, 2008; Jacquet *et al*, 2016a; Murray, 1987; Murray and Kirkland, 1995). The reflecting dispersal events can be described as passive and active because recapture beam has been carried out downwind and upwind of the prevailing wind direction (Kluiters *et al*, 2015b). The map of the average wind direction in France over the last 10 years shows potential answers. It can be seen that the southernmost sampling sites (with the most dissimilar individuals) are in an area where the wind direction is different from the rest of France. It should be noted that the diffusion of the BTV, and thus the dispersion of the *Culicoides*, has already been associated with the wind direction. For example, 2% of BTV infections occurred at distances greater than 31 km (Sedda *et al*, 2012; Sedda *et al*, 2015) during the 2006 epizootic. The South of France was initially sampled, however a poor conservation of *Culicoides* DNA did not allow us to perform barcoding and microsatellite genotyping of these populations. The study of this geographical area could thus complete this study and potentially identify a stronger genetic structure. This will make it possible to decide whether this premise of differentiation observed in this study is due to an older phylogeographic structure, a different wind dispersion in this geographical area, or a random pattern. It is also questionable whether the very low indications of genetic structure could be due to differences in genetic diversity between the south and the north. A lower genetic diversity in the north could then be a sign of an expansion process that would contrast with the interpretation of a large-scale dispersion (Dlugosch and Parker, 2008; White *et al*, 2014). Indeed the spatial expansion of populations is generally accompanied by a gradients of reduction in the genetic diversity of the population during the expansion process, caused by serial founder events, creating a genetic bottleneck through a founding effect (Bock *et al*, 2015; Girod *et al*, 2011). Population sampling at the same scale could allow comparison and estimation of genetic diversity.

Second, anisotropic isolation-by-distance may be due to an artifact due to sampling methods. Indeed, if the extent of the sampling varies depending on the directions, the distances projected from the angles may represent different distance distributions and lead to the over-

representation of strong genetic metrics value in the direction of the scatterplot and the hit of positive correlation signals. Moreover, the absence of correlation does not necessarily mean more gene flow but an absence of isolation-by-distance which can also result from a strong drift and thus less gene flow, a drift which depends on both dispersion and population sizes. It is therefore essential to use as a complement, as we have done here and come up with a similar result, an approach that weights the geographical distances between populations according to their orientation with respect to a given angle axis passing through the barycenter, as implemented in PASSAGE.

Consideration for future work

Our results underline the importance of methodological development of wind dispersal models of *Culicoides* on land and not only over water. However, studies of landscape genetics remain indispensable and complementary in order to improve the accuracy of predictive models for *Culicoides* dispersal over land through integration of meteorological, landscape and activity-based parameters previously tested and validated in the landscape genetics work flow. In addition, population sampling of the same species at European level, *i.e.* a very large proportion of the known range of the species, would be necessary to observe a more marked structuring at European level and to estimate more precisely the gene flow. For example, *C. imicola* in Europe is structured in two large genetic clusters, “Europe central cluster” and “Europe western cluster” (Jacquet *et al*, 2016b). Moreover, the question of the phylogeographic history of *C. obsoletus* is still very little explored in view of its geographical distribution throughout Europe and North America. On the contrary, the study of the phylogeography of *C. imicola*, the main Afrotropical vector, has shown its distribution range out from the northern part of sub-Saharan Africa to the Mediterranean basin (Jacquet *et al*, 2015). This type of study would make it possible to estimate the effective size of the populations, a key factor in the dispersal of *C. obsoletus*. In future study, the use of High-throughput sequencing approaches using markers such as ddRadseq (Double-digest restriction-site associated DNA sequencing) can provide greater resolution in view of the large number of SNPs revealed (single nucleotide polymorphism) at a local scale and improve our understanding of the active and passive dispersal of *Culicoides*. It could also be relevant to look at on scale with finer genetic information. This could include more microsatellite markers or SNP to improve genetic resolution and observe the matching and assignment of each individual.

Conclusions

For the first time, this study provided the first complete landscape genetic analysis of *C. obsoletus*, a major vector species of animal viruses in Europe. This study showed that the large scale dispersion is able to homogenize the genetic structure of populations at the scale of a country. Our results demonstrate, for *C. obsoletus*, a very important inland dispersal and vectorization capacity, which has to be considered for further works on vector competence and epidemiological modeling of disease transmission. Wind direction could be a key factor in the dispersal of many vector insect species. Futures studies should increase their geographical extent to cover the entire species distribution area and better apprehend the limits of *Culicoides* gene flow. In addition to this biological knowledge, the study highlights several important areas for methodological improvement that may currently limit the inclusion of wind direction in landscape genetic analyses.

Abbreviations

BTV: bluetongue virus; SBV: Schmallenberg virus; *cox1*: cytochrome *c* oxidase subunit 1; N: number of individuals. SSR: simple sequence repeat. SNP: single nucleotide polymorphisms.

Acknowledgement

SD and MG are supported by the *Fonds National de la Recherche Scientifique* (FNRS, Belgium). AM and MJ have received a fellowships from the H2020-727393 PALE-Blu project funded the EU. We are most grateful to H el ene Vignes, Ronan Rivallan and Xavier Argout at the GPTR (Grand plateau technique regional de g enotypage) platform in Montpellier, France for their help in development of microsatellite markers and Aurore Manez for her help in genotyping. We would like to thank Sylvain Piry, Karine Berthier and Marie-Pierre Chapuis, for their constructive exchanges which helped to improve the relevance of our conclusions. The authors are grateful to all partners who assisted in *Culicoides* sampling and shipping. We thank William Wint (University of Oxford, PALE-Blu partner) for his continued support in extracting and preparing meteorological data.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files. The newly generated sequences were submitted in the GenBank database under the accession numbers MT828832-MT828844. All *cox1* sequences are available on request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was partially funded by the EU grant H2020-727393 PALE-Blu and by the VectorNet project (OC/EFSA/AHAW/2013/02-FWC1) funded by the European Centre for Disease Prevention and Control (ECDC) and the European Food Safety Authority (EFSA). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission, of the ECDC or of the EFSA.

Author contributions

AM, KH and CG designed the study. TB contributed to gather *Culicoides* samples and IR, MD contributed the identification of *Culicoides*. AM, LG, and M de W have performed molecular biology manipulations. AM, SD and MJ, developed techniques and workflow for landscape genetics analysis. AM analyzed the data. KH and CG contributed to the manuscript firstly written by AM. All authors read and commented the final manuscript version.

Figures and Tables

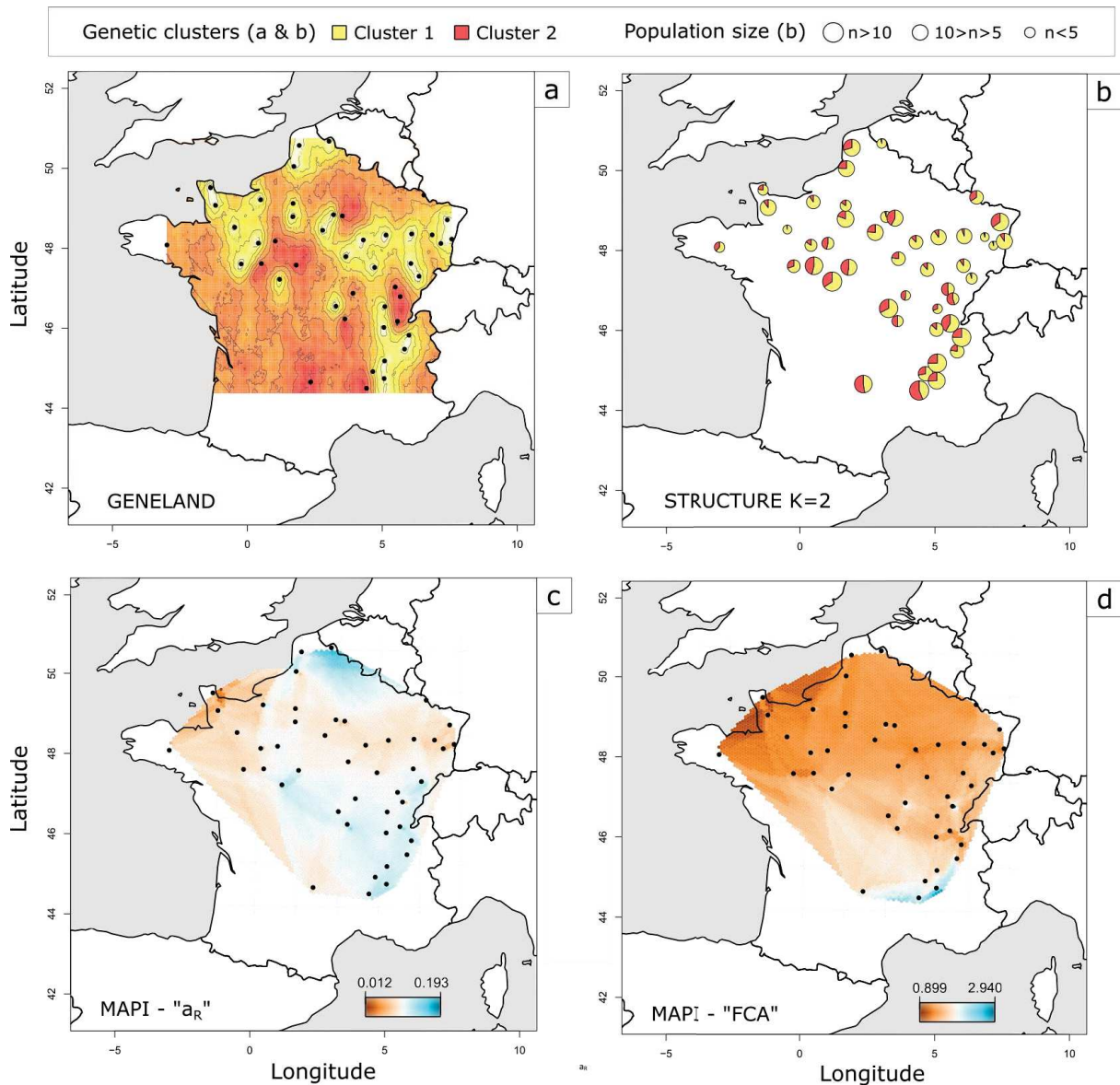


Figure 1. Genetic clustering and genetic differentiation of *C. obsoletus*. Results of the genetic clustering analyses performed with GENELAND (a) and STRUCTURE (b), as well as smoothing of pairwise measures performed with MAPI and based on (c) Rousset's (a_R) and (d) FCA inter-individual genetic distances (see the text for further details). One specific color has been assigned to each genetic cluster in panels (a) and (b). In panels (c) and (d), genetic dissimilarity is represented by a color scale ranging from red (lower genetic dissimilarity) to blue (higher genetic dissimilarity).

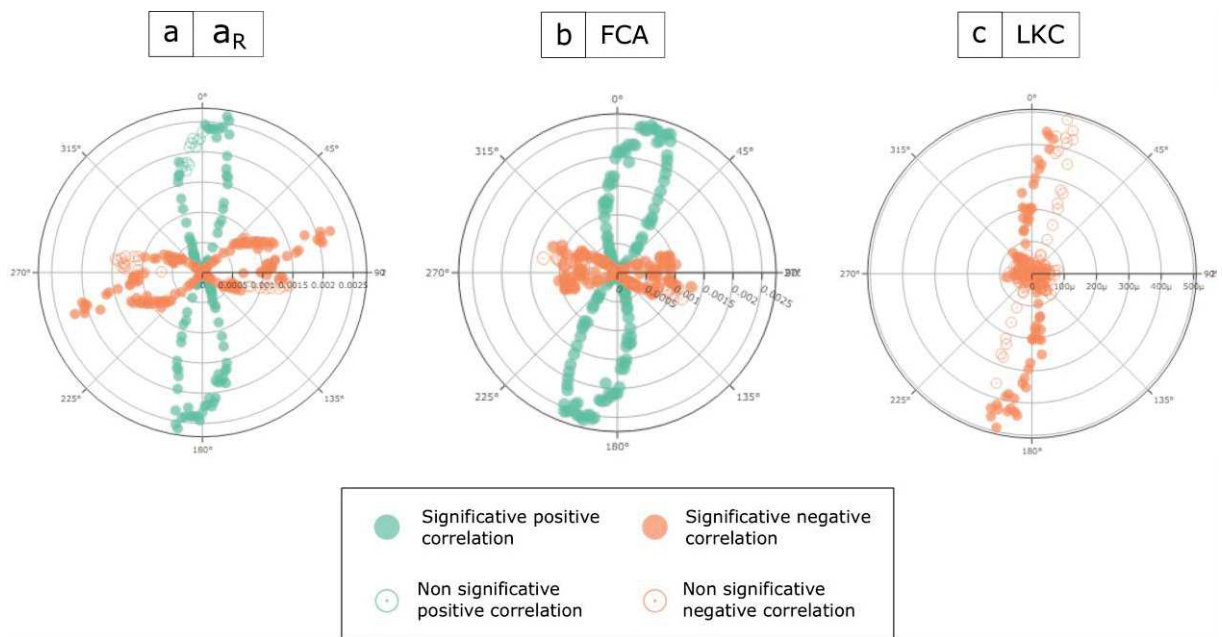


Figure 2. Results of anisotropic isolation-by-distance analyses. Polar plots show the correlation between geographical projected distances by angle and inter-individual genetic distances (a_R , LKC, and FCA).

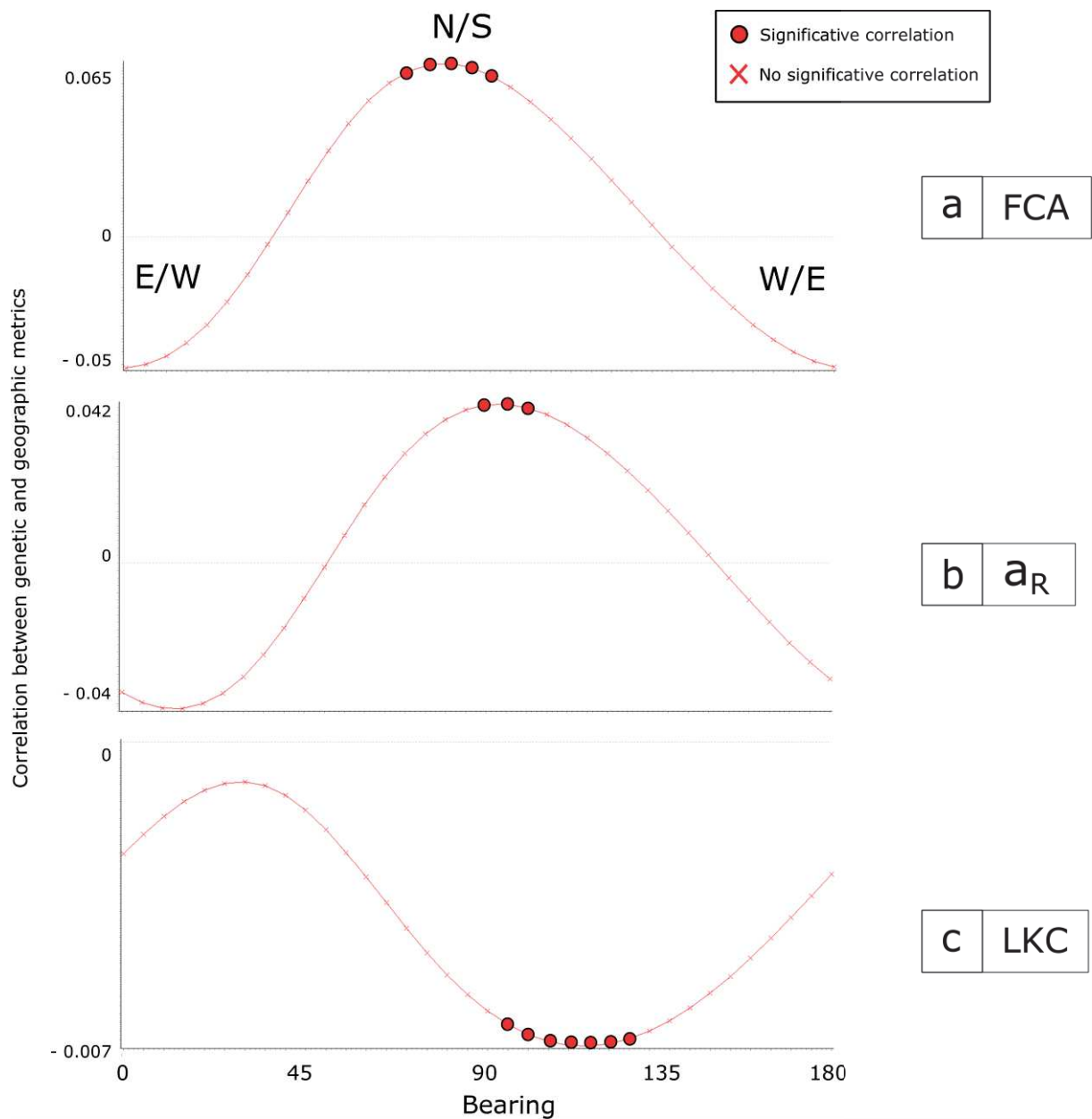


Figure 3. Bearing analysis: correlation between genetic (a_R , LKC, and FCA) and geographical distances as a function of the angle between sampling sites. Circles and crosses indicate significant and non-significant values, respectively. Abbreviations “E”, “W”, “S” and “N” refer to “East”, “West”, “South” and “North”, respectively.

Tableau 1. Results of multiple regressions on distance matrices (MRDM) and additional parameters derived from commonality analysis (CA). Pearson's correlation coefficient (r), significant regression coefficient (β), as well as unique (U) and common (C) contributions of environmental distances to the variance in the dependent variable. [C]/[R] indicate if the considered environmental raster was treated as a conductance or as a resistance factor, and k corresponds to the re-scaling parameter used to transform the initial raster file.

	k	r	β	U	C
FCA					
Cattle density [C]	1000	0.1733	0.1551	0.0221	0.0079
Grassland [R]	1000	0.1080	0.0637	0.0037	0.0079
ar	k	r	β	U	C
Elevation [R]	100	0.0254	-0.9945	0.00044	0.0002
Elevation [R]	1000	0.0291	0.9860	0.00048	0.0004
Grassland [R]	100	0.0448	0.2279	0.00005	0.0020
Grassland [R]	1000	0.0539	-0.1483	0.00002	0.0029
Urban areas [R]	100	0.0725	-0.1279	0.00014	0.0051
Urban areas [R]	1000	0.0879	0.2156	0.00049	0.0072

Supplementary informations

Tableau S1. Sampling sites and associated genetic diversity. N: sample size, Ho: observed heterozygosity, Hs: expected heterozygosity, F_{IS}: Fixation index

Population name (department number)	Longitude	Latitude	N	Ho	Hs	F _{IS}
MONTREAL-LA-CLUSE (01)	5.566	46.172	12	0.780	0.849	0.077
CHAUFFOUR-LES-BAILLY (10)	4.298	48.205	7	0.762	0.832	0.078
VAUDREVILLE (50)	-1.365	49.512	3	0.667	0.712	0.049
CANISY (50)	-1.178	49.070	10	0.736	0.773	0.036
GUDMONT-VILLIERS (52)	5.139	48.327	9	0.765	0.795	0.034
PONTIVY (56)	-2.981	48.080	3	0.606	0.864	0.304
WAMBRECHIES (59)	3.028	50.675	2	0.545	0.909	0.364
HUCQUELIERS (62)	1.921	50.570	11	0.711	0.813	0.123
SCHWENHEIM (67)	7.410	48.710	12	0.764	0.807	0.044
ORBAY (68)	7.171	48.119	1	0.818	-	-
TORCE-EN-VALLEE (72)	0.409	48.126	5	0.691	0.759	0.081
ERCOURT (80)	1.722	50.046	10	0.800	0.837	0.051
FRENELLE-LA-GRANDE (88)	6.088	48.356	9	0.727	0.796	0.092
LAPALISSE (03)	3.615	46.230	4	0.705	0.830	0.122
LAGORCE (07)	4.416	44.494	14	0.734	0.819	0.114
CASSANIOUZE (15)	2.346	44.655	12	0.727	0.794	0.071
UNVERRE (28)	1.034	48.178	5	0.782	0.830	0.041
CERE-LA-RONDE (37)	1.193	47.221	14	0.734	0.832	0.117
BUEIL-EN-TOURAIN (37)	0.519	47.615	12	0.716	0.790	0.080
PARCEY (39)	5.480	47.029	6	0.773	0.840	0.076
PLASNE (39)	5.665	46.788	5	0.727	0.820	0.104
LA MAROLLE-EN-SOLOGNE (41)	1.811	47.578	10	0.736	0.829	0.100
SAINT-DENIS-LES-REBAIS (77)	3.194	48.838	4	0.773	0.811	0.053
LIEUREY (27)	0.495	49.212	7	0.662	0.781	0.132
CHEVIRE-LE-ROUGE (49)	-0.229	47.605	6	0.848	0.821	-0.017
MORSAINS (51)	3.521	48.807	10	0.791	0.809	0.021
FILSTROFF (57)	6.541	49.328	7	0.736	0.817	0.090
LA CHAPELLE-D'ANDAINE (61)	-0.468	48.522	1	0.727	-	-
SCHWOBSHEIM (67)	7.574	48.227	10	0.736	0.796	0.064
LA CHAPELLE-THECLE (71)	5.096	46.540	2	0.773	0.773	-0.017
SAINT-THIBAUD-DE-COUZ (73)	5.827	45.475	7	0.753	0.807	0.057
HERICY (77)	2.791	48.448	10	0.809	0.819	0.012
GAMBAIS (78)	1.689	48.788	11	0.719	0.788	0.083
NOMPATELIZE (88)	6.849	48.333	1	0.818	-	-
VENOY (89)	3.650	47.792	7	0.766	0.801	0.038
CHAUSSY (95)	1.694	49.119	4	0.750	0.777	-0.001
LE PLANTAY (01)	5.059	46.018	7	0.712	0.783	0.100
NEUVY (03)	3.284	46.549	13	0.762	0.794	0.032
VERNOUX-EN-VIVARAIS (07)	4.652	44.915	8	0.670	0.799	0.176
CHANCEAUX (21)	4.717	47.519	7	0.740	0.802	0.073
ADAM-LES-PASSAVANT (25)	6.363	47.298	4	0.682	0.799	0.130
CREPOL (26)	5.086	45.180	13	0.685	0.752	0.078
COBONNE (26)	5.067	44.739	11	0.727	0.836	0.124
CHIDDES (58)	3.915	46.872	2	0.773	0.750	-0.050
MONT-LE-VERNOIS (70)	6.060	47.617	7	0.766	0.816	0.049
MARIGNY-SAINT-MARCEL (74)	5.990	45.823	13	0.720	0.811	0.112

Tableau S2. Reference sequences used for specific assignation.

Species	Accession number	Origin
<i>Culicoides abchazicus</i>	MK760081	GenBank
<i>Culicoides abchazicus</i>	MK760082	GenBank
<i>Culicoides alachua</i>	MK760088	GenBank
<i>Culicoides chiopterus</i>	KT278212	GenBank
<i>Culicoides chiopterus</i>	KM941505	GenBank
<i>Culicoides chiopterus</i>	KM940185	GenBank
<i>Culicoides dewulfi</i>	KJ162977	GenBank
<i>Culicoides montanus</i>	MK760206	GenBank
<i>Culicoides montanus</i>	MK760202	GenBank
<i>Culicoides montanus</i>	MK760207	GenBank
<i>Culicoides montanus</i>	MK760205	GenBank
<i>Culicoides obsoletus</i>	DQ162811	GenBank
<i>Culicoides obsoletus</i>	HM022839	GenBank
<i>Culicoides obsoletus</i>	HM022842	GenBank
<i>Culicoides obsoletus</i>	HM022856	GenBank
<i>Culicoides obsoletus</i>	HQ824374	GenBank
<i>Culicoides obsoletus</i>	KJ624104	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057928	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057935	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057940	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057929	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057930	GenBank
<i>Culicoides obsoletus</i> clade dark	KR057932	GenBank
<i>Culicoides obsoletus</i> clade O1	HQ824371	GenBank
<i>Culicoides obsoletus</i> clade O1	HQ824378	GenBank
<i>Culicoides obsoletus</i> clade O1	JQ620132	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824381	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824382	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824383	GenBank
<i>Culicoides obsoletus</i> clade O2	HQ824384	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ620144	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ978454	GenBank
<i>Culicoides obsoletus</i> clade O3	JQ978459	GenBank
<i>Culicoides sanguisuga</i>	MK760237	GenBank
<i>Culicoides sanguisuga</i>	MK760238	GenBank
<i>Culicoides scoticus</i>	HM022865	GenBank
<i>Culicoides scoticus</i>	HM022875	GenBank
<i>Culicoides scoticus</i>	KJ624126	GenBank
<i>Culicoides scoticus</i>	KT186881	GenBank
<i>Culicoides sinanoensis</i>	MK760244	GenBank
<i>Culicoides sinanoensis</i>	MK760242	GenBank
<i>Culicoides</i> sp.	GMGRC1056-13	BOLD Systems
<i>Culicoides</i> sp.	GMGRC1000-13	BOLD Systems
<i>Culicoides</i> sp.	GMGRD2587-13	BOLD Systems

Tableau S3. Origin and numbers of individuals used to build up the DNA library necessary for the development of the microsatellite markers.

Origin	Number of individuals	Collection date
France, Loire	1	June 2012
France, Pyrénées-Atlantique	1	June 2012
France, Morbihan	1	July 2012
France, Haute-Saone	1	June 2012
France, Corsica	5	December 2011
United Kingdom, Kent	1	2014

Tableau S4. Primers of the 13 microsatellite markers used to genotype *C. obsoletus* populations. “DYE”: fluorochrome name, “TFm”: half denaturation temperature in degree, “bp”: base pairs.

Multiplex	Locus	DYE	Repeat motif	Forward primer (5'-3')	Tm	Reverse primer (3'-5')	Tm	Size range (bp)	Accession numbers
Plex1	OBSms5	FAM	(TG)15	GCTCATTATGTAATTTGTGCC	56,4	CGAGACGGCATGTTTATCT	56,3	140-204	MT828835
	OBSms11	FAM	(CT)12	CATCCATCCACTTTGCTCT	56,1	GGACTGTCATTCAAACAAACA	56,1	262-293	MT828838
	OBSms2	NED	(TG)14	TGATAACAGATAACACGGAAGG	56,8	TGATGCAATAATGCCTCTGT	58,6	135-171	MT828832
	OBSms28	NED	(CA)8	GCCGCAATTTGTGTTTT	55,5	TTCCTTGACCTGCGAAC	55,5	239-264	MT828842
	OBSms3	VIC	(CA)13	ATTTACGCGACGATCATTT	55,8	AAAGCTCGAATACAGCACAA	56,2	140-184	MT828833
	OBSms26	VIC	(CA)5	CGCGATACATCAAGAGGA	55,7	GTAGGGCAGCAGCAGAA	55,9	200-230	MT828841
	OBSms4	PET	(GA)12	AATGTAGGCCAAACAGAAA	53,3	TTGCATAACAACATCATCAC	52,7	160-180	MT828834
Plex2	OBSms25	FAM	(TG)6	CCGCTTTACTCAATTCCC	55,7	ACACCCCTCTGCAAACA	55,8	200-230	MT828844
	OBSms29	FAM	(CA)5	GTTGACCGGGACACTTTT	55,8	CATGACAAGCACGAACG	55,4	252-268	MT828843
	OBSms6	NED	(TA)12	GACCTAATTCATAATCGTCCAA	55,8	TTAAAGCACGTTTCGATTTGT	56,1	175-201	MT828836
	OBSms13	VIC	(CA)12	CGTTCCATCGACTCATCTT	56,1	GCTTTTACTCCTTCGTGATGT	56,1	249-287	MT828839
	OBSms14	VIC	(AT)8	TCAAATCATTCCAAGAGCA	55,1	CCTCTCTGTGGTCTTTCGT	55,2	139-164	MT828840
	OBSms10	PET	(AC)17	GACATGGAATCCTTAATTGG	54,5	GATTGCTTCAGACAAACACA	54,9	224-260	MT828837

Tableau S5. Genetic diversity by locus. Ho: observed heterozygosity, Hs: expected heterozygosity, FIS: Fixation index, SAD: short alleles dominance.

Locus name	Alleles number	Ho	Hs	F _{IS}	SAD r ²	Null alleles frequency (95% confidence interval)	Association index
OBSms10	21	0.7146	0.8856	0.1931	0.0311	0.0947 [0.0639-0.1247]	0.0280
OBSms11	23	0.8335	0.8816	0.0546	0.0010	0.0320 [0.0097-0.0526]	0.0063
OBSms13	23	0.8721	0.9097	0.0413	0.0438	0.0224 [0.0028-0.0443]	0.0205
OBSms14	18	0.7536	0.7767	0.0298	0.0711	0.0248 [-0.0022-0.0525]	0.0225
OBSms2	23	0.7248	0.8383	0.1353	0.0405	0.0727 [0.0447-0.1008]	0.0372
OBSms25	14	0.1972	0.6380	0.6910	0.0318	0.5600 [0.4844-0.6345]	0.0128
OBSms26	20	0.5663	0.8817	0.3577	0.1140	0.2260 [0.1863-0.2715]	0.0309
OBSms28	12	0.7086	0.7517	0.0573	0.0834	0.0171 [-0.0124-0.0490]	0.0094
OBSms29	9	0.3611	0.4047	0.1079	0.1084	0.0594 [0.0201-0.1112]	0.0118
OBSms3	26	0.7915	0.8365	0.0538	0.0451	0.0269 [0.0014-0.0536]	0.0359
OBSms4	13	0.8011	0.8265	0.0307	0.1070	0.0186 [-0.0057-0.0444]	0.0010
OBSms5	40	0.7974	0.9185	0.1318	0.0735	0.0769 [0.0511-0.1047]	0.0131
OBSms6	32	0.7500	0.8478	0.1154	0.0096	0.0676 [0.0387-0.0972]	-0.0132

Tableau S6. Results of univariate analyses: determination coefficients R²'s estimated from univariate regressions between genetic and environmental distances. [C]/[R] indicate if the considered environmental raster was respectively treated as a conductance or resistance factor for the computation of environmental distances with circuit theory. Q: Difference between environmental MRDM R² and null raster MRDM R². The results for K=10 are not present here as they were non-significant.

	<i>k</i>	β	Q	<i>r</i>	p-value
FCA					
Cattle distribution [C]	100	0.0991	0.0092	0.0098	0.0010
Cattle distribution [C]	1000	0.1733	0.0294	0.0300	0.0010
Elevation [R]	100	0.0418	0.0011	0.0018	0.0050
Elevation [R]	1000	0.0434	0.0013	0.0019	0.0040
Land cover grassland [R]	100	0.1007	0.0095	0.0102	0.0020
Land cover grassland [R]	1000	0.1080	0.0110	0.0117	0.0010
Sheep distribution [C]	100	0.0585	0.0028	0.0034	0.0040
Sheep distribution [C]	1000	0.0834	0.0063	0.0070	0.0020
Null raster [R]		-0.0250	0.0000	0.0006	0.0050
ar					
Elevation [R]	100	0.0254	0.0002	0.0006	0.0080
Elevation [R]	1000	0.0291	0.0004	0.0009	0.0030
Land cover grassland [R]	100	0.0448	0.0015	0.0020	0.0010
Land cover grassland [R]	1000	0.0539	0.0024	0.0029	0.0010
Land cover urban areas [R]	100	0.0725	0.0048	0.0053	0.0010
Land cover urban areas [R]	1000	0.0879	0.0073	0.0077	0.0010
Null raster [R]		-0.0216	0.0000	0.0005	0.0010

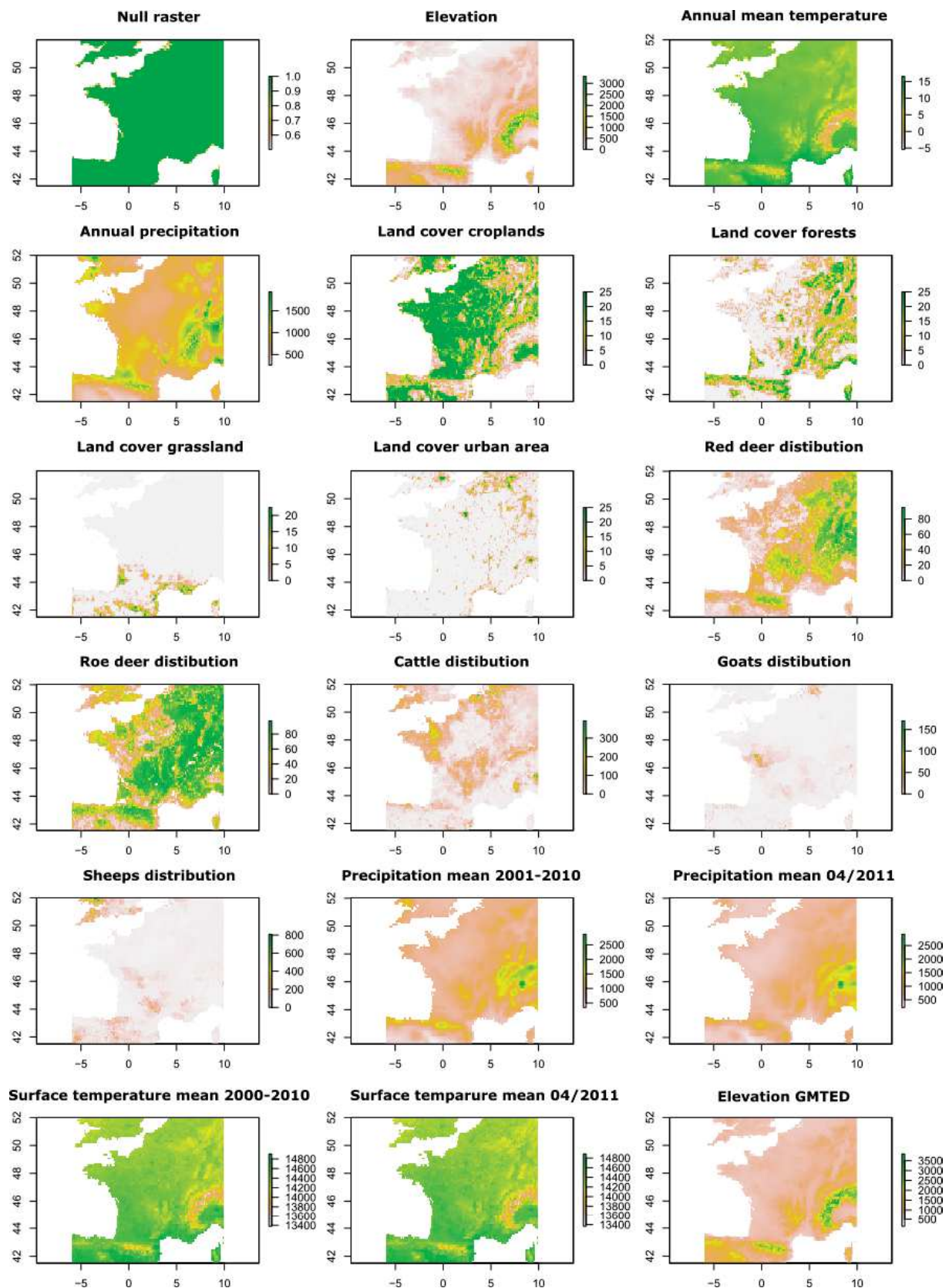


Figure S1. Environmental variables tested as potential factors that could impact inter-individual genetic differentiation of *C. obsoletus* in France (raster cell resolution: 0.04 arcmin).

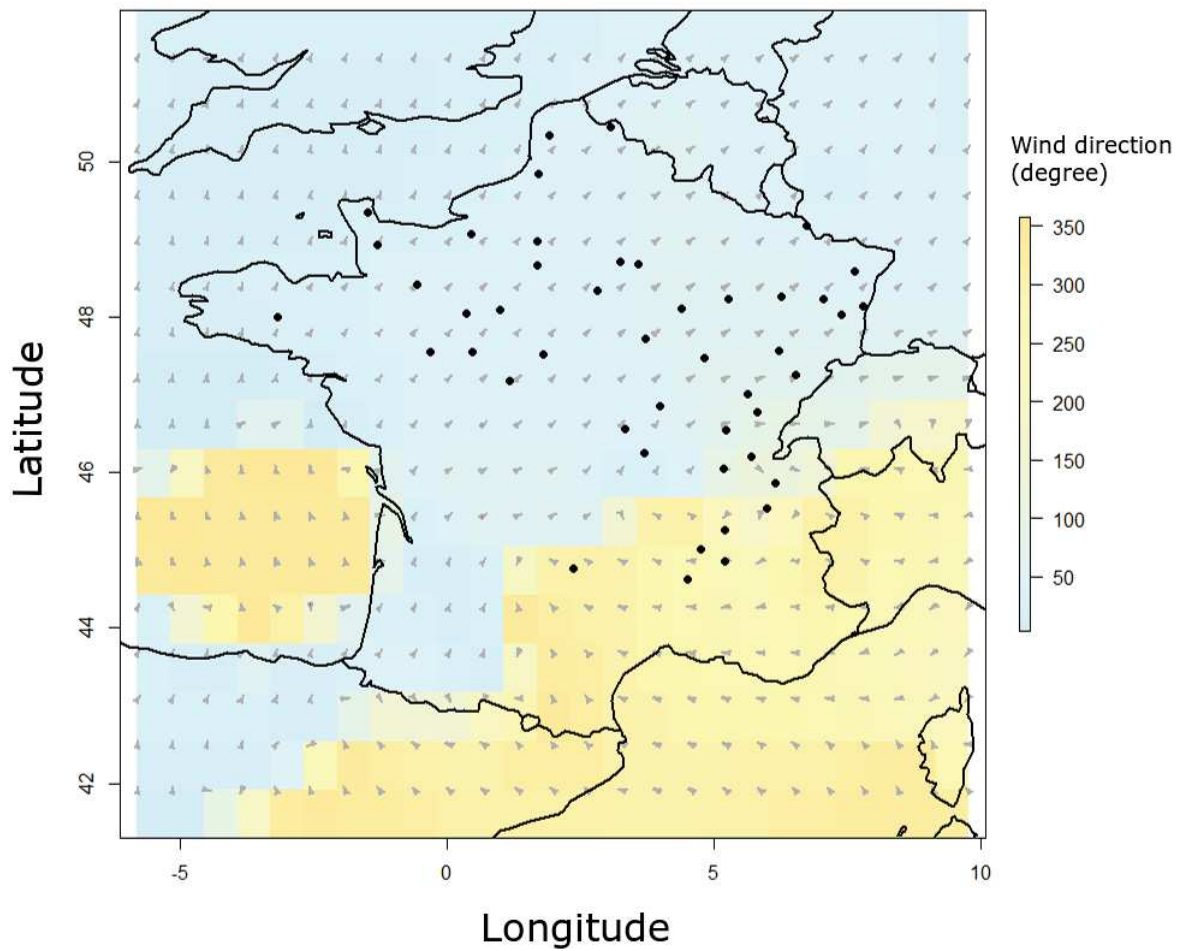


Figure S2. Map of wind direction averaged from 2000 to 2010. Sampling sites are represented by black points. The color scale represents the wind direction from 0 to 360 degrees from North. The arrowheads indicate the exact wind direction of each raster pixel.

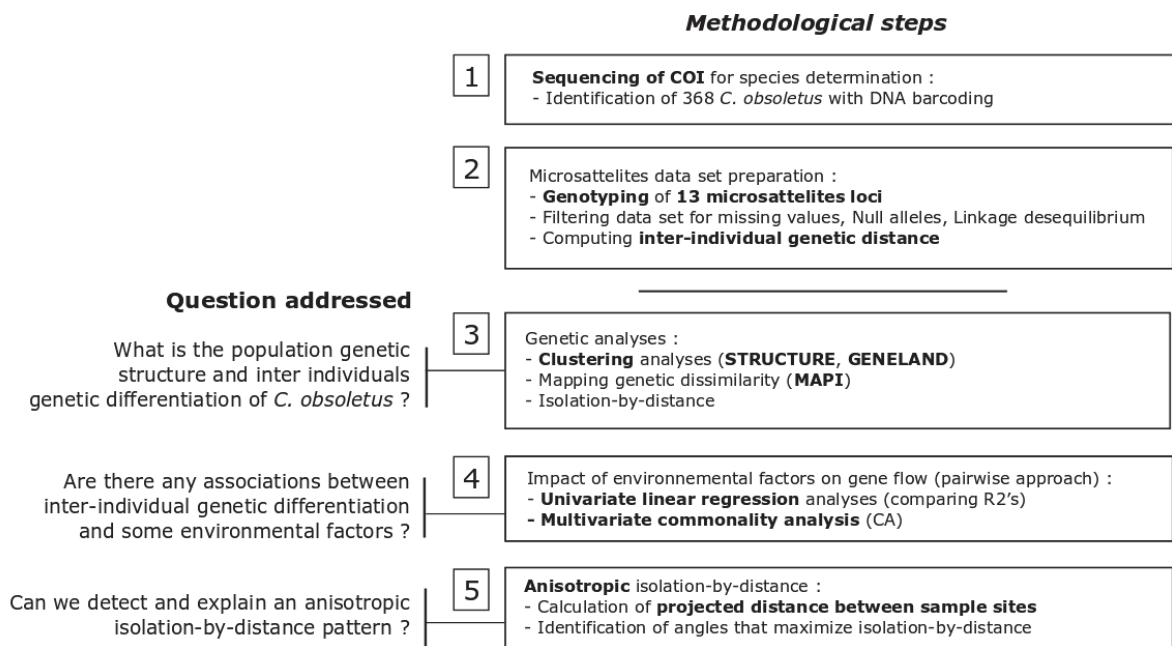


Figure S3. Analytical workflow

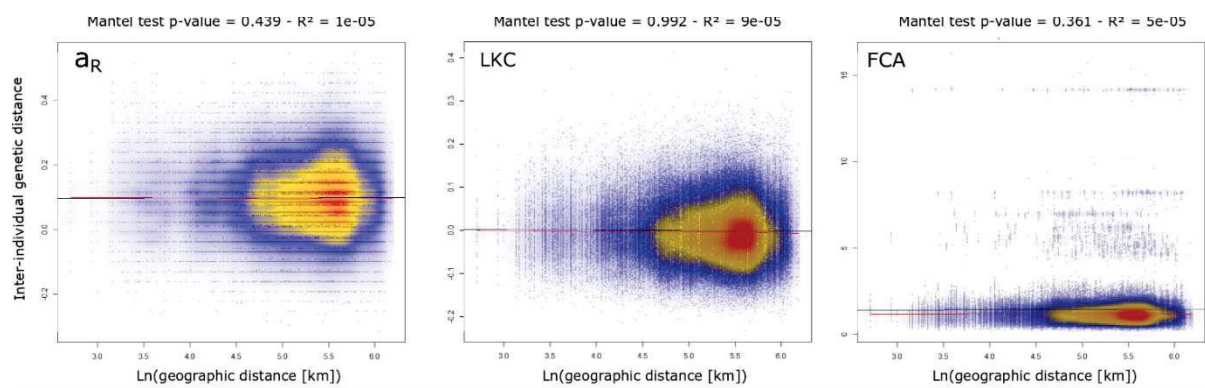


Figure S4. Isolation-by-distance analyses: density plots, Mantel tests and linear regressions performed with each inter-individual genetic distance considered in this study.

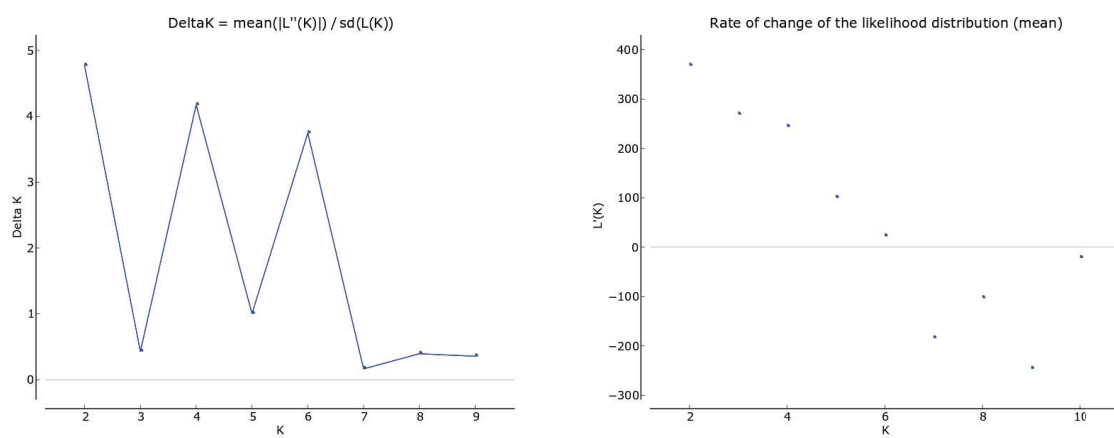


Figure S5. Identification of the optimal number of genetic clusters (K) inferred by STRUCTURE using the delta(K) and L'(K) methods.

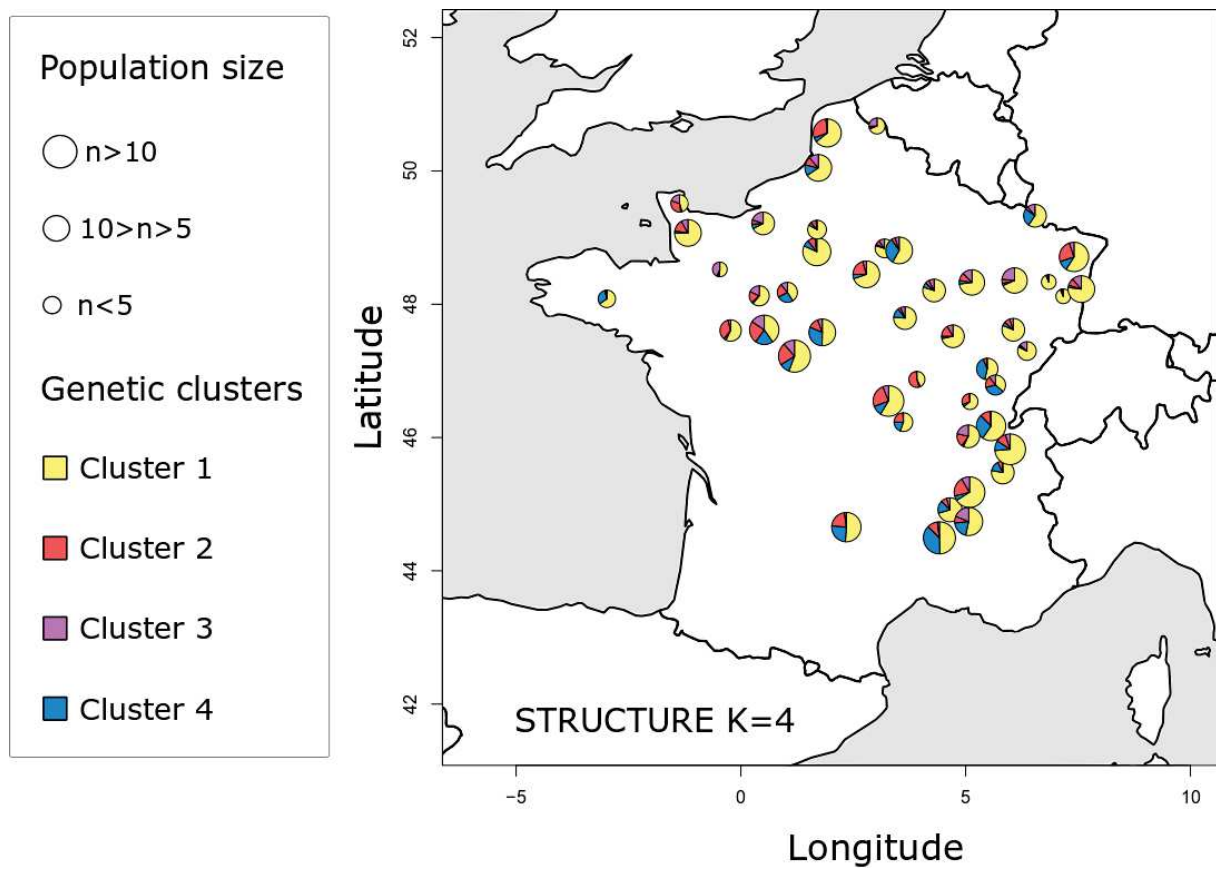


Figure S6. Population genetic structure results by clustering analyses performed STRUCTURE. One specific color has been assigned to each inferred genetic cluster.

References

- Adamack AT, Gruber B, Dray S (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* **5**(4): 384-387.
- Agren EC, Burgin L, Lewerin SS, Gloster J, Elvander M (2010). Possible means of introduction of bluetongue virus serotype 8 (BTV-8) to Sweden in August 2008: comparison of results from two models for atmospheric transport of the *Culicoides* vector. *Vet Rec* **167**(13): 484-488.
- Aguilar-Vega C, Fernandez-Carrion E, Sanchez-Vizcaino JM (2019). The possible route of introduction of bluetongue virus serotype 3 into Sicily by windborne transportation of infected *Culicoides* spp. *Transbound Emerg Dis* **66**(4): 1665-1673.
- Alba A, Casal J, Domingo M (2004). Possible introduction of bluetongue into the Balearic Islands, Spain, in 2000, via air streams. *Vet Rec* **155**(15): 460-461.
- Andrew Jones DT, Matthew Hort, and Ben Devenish (2007). The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III.
- Ayllon T, Nijhof AM, Weiher W, Bauer B, Allene X, Clausen PH (2014). Feeding behaviour of *Culicoides* spp. (Diptera: Ceratopogonidae) on cattle and sheep in northeast Germany. *Parasit Vectors* **7**: 34.
- Baylis M, Parkin H, Kreppel K, Carpenter S, Mellor PS, McIntyre KM (2010). Evaluation of housing as a means to protect cattle from *Culicoides* biting midges, the vectors of bluetongue virus. *Med Vet Entomol* **24**(1): 38-45.
- Belkhir K (2004). Genetix 4.05, logiciel sous Windows TM pour la genetique des populations. <http://www.genetixuniv-montp2fr/genetix/genetixhtm>.
- Bock DG, Caseys C, Cousens RD, Hahn MA, Heredia SM, Hubner S *et al* (2015). What we still don't know about invasion genetics. *Mol Ecol* **24**(9): 2277-2297.
- Brenner RJ, Wargo MJ, Stains GS, Mulla MS (1984). The Dispersal of *Culicoides mohave* (Diptera, Ceratopogonidae) in the Desert of Southern-California. *Mosquito News* **44**(3): 343-350.
- Burgin L, Ekstrom M, Dessai S (2017). Combining dispersion modelling with synoptic patterns to understand the wind-borne transport into the UK of the bluetongue disease vector. *Int J Biometeorol* **61**(7): 1233-1245.
- Burgin LE, Gloster J, Sanders C, Mellor PS, Gubbins S, Carpenter S (2013). Investigating incursions of bluetongue virus using a model of long-distance *Culicoides* biting midge dispersal. *Transbound Emerg Dis* **60**(3): 263-272.
- Carabajal CC, Harding DJ, Boy JP, Danielson JJ, Gesch DB, Suchdeo VP (2011). Evaluation of the Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010) Using ICESat Geodetic Control. *International Symposium on Lidar and Radar Mapping 2011: Technologies and Applications* **8286**.

- Carpenter S, Groschup MH, Garros C, Felipe-Bauer ML, Purse BV (2013). *Culicoides* biting midges, arboviruses and public health in Europe. *Antiviral Res* **100**(1): 102-113.
- Carpenter S, Szymaragd C, Barber J, Labuschagne K, Gubbins S, Mellor P (2008c). An assessment of *Culicoides* surveillance techniques in northern Europe: have we underestimated a potential bluetongue virus vector? *J Appl Ecol* **45**(4).
- Carpenter S, Wilson A, Mellor PS (2009). *Culicoides* and the emergence of bluetongue virus in northern Europe. *Trends Microbiol* **17**(4): 172-178.
- Cuellar AC, Kjaer LJ, Baum A, Stockmarr A, Skovgard H, Nielsen SA *et al* (2020). Modelling the monthly abundance of *Culicoides* biting midges in nine European countries using Random Forests machine learning. *Parasit Vectors* **13**(1): 194.
- Delecolle JC (1985). Nouvelle contribution à l'étude systématique et iconographique des espèces du genre *Culicoides*, (Diptéra) : (Cératopogonidae) du Nord-Est de la France. Strasbourg.
- Dellicour S, Prunier JG, Piry S, Eloy MC, Bertouille S, Licoppe A *et al* (2019). Landscape genetic analyses of *Cervus elaphus* and *Sus scrofa*: comparative study and analytical developments. *Heredity (Edinb)* **123**(2): 228-241.
- Diaz-Sanchez S, Hernandez-Jarguin A, Torina A, Fernandez de Mera IG, Estrada-Pena A, Villar M *et al* (2018). Biotic and abiotic factors shape the microbiota of wild-caught populations of the arbovirus vector *Culicoides imicola*. *Insect Mol Biol* **27**(6): 847-861.
- Dijkstra E, van der Ven IJ, Meiswinkel R, Holzel DR, Van Rijn PA, Meiswinkel R (2008). *Culicoides chiopterus* as a potential vector of bluetongue virus in Europe. *Vet Rec* **162**(13): 422.
- Dlugosch KM, Parker IM (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* **17**(1): 431-449.
- Duchesne P, Turgeon J (2012). FLOCK provides reliable solutions to the "number of populations" problem. *J Hered* **103**(5): 734-743.
- Ducheyne E, De Deken R, Becu S, Codina B, Nomikou K, Mangana-Vougiaki O *et al* (2007). Quantifying the wind dispersal of *Culicoides* species in Greece and Bulgaria. *Geospat Health* **1**(2): 177-189.
- Durr PA, Graham K, van Klinken RD (2017). Sellers' Revisited: A Big Data Reassessment of Historical Outbreaks of Bluetongue and African Horse Sickness due to the Long-Distance Wind Dispersion of *Culicoides* Midges. *Front Vet Sci* **4**: 98.
- Eagles D, Deveson T, Walker PJ, Zalucki MP, Durr P (2012). Evaluation of long-distance dispersal of *Culicoides* midges into northern Australia using a migration model. *Med Vet Entomol* **26**(3): 334-340.

- Eagles D, Melville L, Weir R, Davis S, Bellis G, Zalucki MP *et al* (2014). Long-distance aerial dispersal modelling of *Culicoides* biting midges: case studies of incursions into Australia. *BMC Vet Res* **10**: 135.
- Earl DA, vonHoldt BM (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**(2): 359-361.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**(5): 1792-1797.
- Elbers AR, Backx A, Mintiens K, Gerbier G, Staubach C, Hendrickx G *et al* (2008). Field observations during the Bluetongue serotype 8 epidemic in 2006. II. Morbidity and mortality rate, case fatality and clinical recovery in sheep and cattle in the Netherlands. *Prev Vet Med* **87**(1-2): 31-40.
- Elbers AR, Meiswinkel R (2014). *Culicoides* (Diptera: Ceratopogonidae) host preferences and biting rates in the Netherlands: comparing cattle, sheep and the black-light suction trap. *Vet Parasitol* **205**(1-2): 330-337.
- Fernández-López J, Schliep K (2018). rWind: download, edit and include wind data in ecological and evolutionary analysis. *Ecography* **42**(4): 804-810.
- Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *J Appl Ecol* **46**(2): 493-505.
- Garros C, Balenghien T, Carpenter S, Delecolle JC, Meiswinkel R, Pédarrieu A *et al* (2014). Towards the PCR-based identification of Palaearctic *Culicoides* biting midges (Diptera: Ceratopogonidae): results from an international ring trial targeting four species of the subgenus *Avaritia*. *Parasit Vectors* **7**: 223.
- Gilbert M, Nicolas G, Cinardi G, Van Boeckel TP, Vanwambeke S, Wint WGR *et al.* (2018). Harvard Dataverse.
- Ginestet C (2011). ggplot2: Elegant Graphics for Data Analysis. *J Roy Stat Soc Ser A (Stat Soc)* **174**(1): 245-246.
- Girod C, Vitalis R, Leblois R, Freville H (2011). Inferring population decline and expansion from microsatellite data: a simulation-based evaluation of the Msvar method. *Genetics* **188**(1): 165-179.
- Gloster J, Burgin L, Witham C, Athanassiadou M, Mellor PS (2008). Bluetongue in the United Kingdom and northern Europe in 2007 and key issues for 2008. *Vet Rec* **162**(10): 298-302.
- Goslee SC, Urban DL (2007). TheecodistPackage for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software* **22**(7): 1-19.

- Gubbins S, Carpenter S, Baylis M, Wood JL, Mellor PS (2008). Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. *J R Soc Interface* **5**(20): 363-371.
- Guillot G, Mortier F, Estoup A (2005). GENELAND: a computer package for landscape genetics. *Mol Ecol Notes* **5**(3): 712-715.
- Guis H, Caminade C, Calvete C, Morse AP, Tran A, Baylis M (2012). Modelling the effects of past and future climate on the risk of bluetongue emergence in Europe. *J R Soc Interface* **9**(67): 339-350.
- Hardy OJ (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Mol Ecol* **12**(6): 1577-1588.
- Hardy OJ, Vekemans X (2002). SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**(4): 618-620.
- Hendrickx G, Gilbert M, Staubach C, Elbers A, Mintiens K, Gerbier G *et al* (2008). A wind density model to quantify the airborne spread of *Culicoides* species during north-western Europe bluetongue epidemic, 2006. *Prev Vet Med* **87**(1-2): 162-181.
- Hijmans RJ (2019). geosphere: Spherical Trigonometry.
- Hoffmann B, Eschbaumer M, Beer M (2009). Real-time quantitative reverse transcription-PCR assays specifically detecting bluetongue virus serotypes 1, 6, and 8. *J Clin Microbiol* **47**(9): 2992-2994.
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* **9**(5): 1322-1332.
- Jacquet S, Garros C, Lombaert E, Walton C, Restrepo J, Allene X *et al* (2015). Colonization of the Mediterranean basin by the vector biting midge species *Culicoides imicola*: an old story. *Mol Ecol* **24**(22): 5707-5725.
- Jacquet S, Huber K, Guis H, Setier-Rio ML, Goffredo M, Allene X *et al* (2016a). Spatio-temporal genetic variation of the biting midge vector species *Culicoides imicola* (Ceratopogonidae) Kieffer in France. *Parasit Vectors* **9**: 141.
- Jacquet S, Huber K, Pages N, Talavera S, Burgin LE, Carpenter S *et al* (2016b). Range expansion of the Bluetongue vector, *Culicoides imicola*, in continental France likely due to rare wind-transport events. *Sci Rep* **6**: 27247.
- Jacquot M, Nomikou K, Palmarini M, Mertens P, Biek R (2017). Bluetongue virus spread in Europe is a consequence of climatic, landscape and vertebrate host factors as revealed by phylogeographic inference. *Proc Biol Sci* **284**(1864).
- Jay F, Sjodin P, Jakobsson M, Blum MG (2013). Anisotropic isolation by distance: the main orientations of human genetic differentiation. *Mol Biol Evol* **30**(3): 513-525.

- Kamvar ZN, Tabima JF, Grunwald NJ (2014). Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*.
- Kimmig SE, Beninde J, Brandt M, Schleimer A, Kramer-Schadt S, Hofer H *et al* (2020). Beyond the landscape: Resistance modelling infers physical and behavioural gene flow barriers to a mobile carnivore across a metropolitan area. *Mol Ecol* **29**(3): 466-484.
- Kluiters G, Swales H, Baylis M (2015b). Local dispersal of palaeartic *Culicoides* biting midges estimated by mark-release-recapture. *Parasit Vectors* **8**: 86.
- Landguth EL, Cushman SA, Murphy MA, Luikart G (2010). Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Mol Ecol Resour* **10**(5): 854-862.
- Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv Genet* **7**(2): 295-302.
- Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial Genetic-Structure of a Tropical Understory Shrub, *Psychotria Officinalis* (Rubiaceae). *Am J Bot* **82**(11): 1420-1425.
- Maclachlan NJ, Guthrie AJ (2010). Re-emergence of bluetongue, African horse sickness, and other orbivirus diseases. *Vet Res* **41**(6): 35.
- Manangwa O, De Meeus T, Grebaut P, Segard A, Byamungu M, Ravel S (2019). Detecting Wahlund effects together with amplification problems: Cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. *Mol Ecol Resour* **19**(3): 757-772.
- Manel S, Berthoud F, Bellemain E, Gaudeul M, Luikart G, Swenson JE *et al* (2007). A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Mol Ecol* **16**(10): 2031-2043.
- Manel S, Holderegger R (2013). Ten years of landscape genetics. *Trends Ecol Evol* **28**(10): 614-621.
- Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**(2): 209-220.
- McGrath G, More SJ, O'Neill R (2018). Hypothetical route of the introduction of Schmallenberg virus into Ireland using two complementary analyses. *Vet Rec* **182**(8): 226.
- McRae B, Shah VB (2008). Circuitscape: A Tool for Landscape Ecology.
- McRae BH (2006). Isolation by resistance. *Evolution* **60**(8): 1551-1561.
- Mehlhorn H, Walldorf V, Klimpel S, Jahn B, Jaeger F, Eschweiler J *et al* (2007). First occurrence of *Culicoides obsoletus*-transmitted Bluetongue virus epidemic in Central Europe. *Parasitol Res* **101**(1): 219-228.

- Mignotte A, Garros C, Gardes L, Balenghien T, Duhayon M, Rakotoarivony I *et al* (2020). The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palaearctic vector *Obsoletus/Scoticus* Complex (Diptera: Ceratopogonidae) at the European level. *Parasites & Vectors* **13**(1): 265.
- Murray MD (1987). Local Dispersal of the Biting-Midge *Culicoides-Brevitarsis* Kieffer (Diptera, Ceratopogonidae) in Southeastern Australia. *Aust J Zool* **35**(6): 559-573.
- Murray MD, Kirkland PD (1995). Bluetongue and Douglas virus activity in New South Wales in 1989 : further evidence for long-distance dispersal of the biting midge *Culicoides brevitarsis*. *Australian Veterinary Journal* **Vol. 72**.
- Neil A, David M, Jolyon m, Kate S, Willy W (2014). *A first attempt at modelling roe deer (Capreolus capreolus) distributions over Europe*.
- Newton RG, Spurrell DJ (1967). Examples of the Use of Elements for Clarifying Regression Analyses. *Applied Statistics* **16**(2): 165-172.
- Nimon K, Lewis M, Kane R, Haynes RM (2008). An R package to compute commonality coefficients in the multiple regression case: an introduction to the package and a practical example. *Behav Res Methods* **40**(2): 457-466.
- Nychka D, Furrer R, Paige J, Sain S (2017). *Fields: Tools for spatial data*.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB *et al* (2013). *Vegan: Community Ecology Package*. R Package Version. 2.0-10. *CRAN*.
- Onyango MG, Beebe NW, Gopurenko D, Bellis G, Nicholas A, Ogugo M *et al* (2015a). Assessment of population genetic structure in the arbovirus vector midge, *Culicoides brevitarsis* (Diptera: Ceratopogonidae), using multi-locus DNA microsatellites. *Vet Res* **46**: 108.
- Onyango MG, Michuki GN, Ogugo M, Venter GJ, Miranda MA, Elissa N *et al* (2015b). Delineation of the population genetic structure of *Culicoides imicola* in East and South Africa. *Parasit Vectors* **8**: 660.
- Pioz M, Guis H, Crespin L, Gay E, Calavas D, Durand B *et al* (2012). Why did bluetongue spread the way it did? Environmental factors influencing the velocity of bluetongue virus serotype 8 epizootic wave in France. *PLoS One* **7**(8): e43360.
- Piry S, Chapuis MP, Gauffre B, Papaix J, Cruaud A, Berthier K (2016). Mapping Averaged Pairwise Information (MAPI): a new exploratory tool to uncover spatial structure. *Methods Ecol Evol* **7**(12): 1463-1475.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2): 945-959.
- Prunier JG, Colyn M, Legendre X, Flamand MC (2017). Regression commonality analyses on hierarchical genetic distances. *Ecography* **40**(12): 1412-1425.

- Prunier JG, Kaufmann B, Fenet S, Picard D, Pompanon F, Joly P *et al* (2013a). Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: towards a better assessment of functional connectivity using an individual-based sampling scheme. *Mol Ecol* **22**(22): 5516-5530.
- Purse BV, Carpenter S, Venter GJ, Bellis G, Mullens BA (2015). Bionomics of temperate and tropical *Culicoides* midges: knowledge gaps and consequences for transmission of *Culicoides*-borne viruses. *Annu Rev Entomol* **60**: 373-392.
- Purse BV, Mellor PS, Rogers DJ, Samuel AR, Mertens PP, Baylis M (2005). Climate change and the recent emergence of bluetongue in Europe. *Nat Rev Microbiol* **3**(2): 171-181.
- Rice WR (1989). Analysing tables of statistic tests. *Evolution* **43**(1): 223-225.
- Rousset F (2000). Genetic differentiation between individuals. *J Evol Biol* **13**(1): 58-62.
- Sailleau C, Breard E, Viarouge C, Vitour D, Romey A, Garnier A *et al* (2017). Re-Emergence of Bluetongue Virus Serotype 8 in France, 2015. *Transbound Emerg Dis* **64**(3): 998-1000.
- Savini G, MacLachlan NJ, Sanchez-Vizcaino JM, Zientara S (2008). Vaccines against bluetongue in Europe. *Comp Immunol Microbiol Infect Dis* **31**(2-3): 101-120.
- Sedda L, Brown HE, Purse BV, Burgin L, Gloster J, Rogers DJ (2012). A new algorithm quantifies the roles of wind and midge flight activity in the bluetongue epizootic in northwest Europe. *Proc Biol Sci* **279**(1737): 2354-2362.
- Sedda L, Morley D, Brown HE (2015). Characteristics of Wind-Infective Farms of the 2006 Bluetongue Serotype 8 Epidemic in Northern Europe. *EcoHealth* **12**(3): 461-467.
- Shirk AJ, Landguth EL, Cushman SA (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Mol Ecol Resour* **17**(6): 1308-1317.
- Shults P, Ho A, Martin EM, McGregor BL, Vargo EL (2020). Genetic Diversity of *Culicoides stellifer* (Diptera: Ceratopogonidae) in the Southeastern United States Compared With Sequences From Ontario, Canada. *J Med Entomol* **57**(4): 1324-1327.
- Viennet E, Garros C, Gardes L, Rakotoarivony I, Allene X, Lancelot R *et al* (2013). Host preferences of Palaearctic *Culicoides* biting midges: implications for transmission of orbiviruses. *Med Vet Entomol* **27**(3): 255-266.
- Wan ZM (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens Environ* **140**: 36-45.
- Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol Ecol* **7**(11): 1569-1573.

- Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**(6): 1358-1370.
- White BP, Pilgrim EM, Boykin LM, Stein ED, Mazor RD (2014). Comparison of four species-delimitation methods applied to a DNA barcode data set of insect larvae for use in routine bioassessment. *Freshwater Science* **33**(1): 338-348.
- Winter DJ (2012). MMOD: an R library for the calculation of population differentiation statistics. *Mol Ecol Resour* **12**(6): 1158-1160.
- Wittmann EJ, Mello PS, Baylis M (2002). Effect of temperature on the transmission of orbiviruses by the biting midge, *Culicoides sonorensis*. *Med Vet Entomol* **16**(2): 147-156.

Discussion générale

Les objectifs principaux de ces travaux de thèse étaient (i) de déterminer par l'observation des patrons de distributions des fréquences alléliques, comment la dispersion et les évènements évolutifs plus anciens ont façonné l'arrangement spatial de la diversité génétique de *C. obsoletus*, (ii) et comment cette dispersion est favorisée ou limitée par les paysages et les caractéristiques des habitats dans lesquels cette espèce évolue. Une étape préalable d'identification taxonomique a été nécessaire afin de réaliser notre objectif principal et de ne pas biaiser nos conclusions. Etant donnée la diversité cryptique récemment publiée, une révision taxonomique du Complexe Obsoletus/Scoticus était indispensable afin de s'assurer de la bonne identification des individus inclus dans notre jeu de données (chapitre I, article 1). Cette première étape a permis (i) d'identifier le statut spécifique des spécimens collectés (ii), de cartographier la diversité du Complexe Obsoletus/Scoticus à l'échelle ouest paléarctique et (iii) de sélectionner les individus *C. obsoletus* pour la suite de nos travaux. En se basant sur les individus identifiés comme *C. obsoletus sensu stricto*, nous avons pu retracer l'histoire évolutive de *C. obsoletus* en Europe et identifier les caractéristiques démographiques, évolutives et temporelles qui ont façonné les patrons génétiques actuels (chapitre II, article 2). Enfin, nous avons travaillé à une échelle plus fine, au niveau national et en réalisant un échantillonnage individuel, pour identifier l'impact des facteurs environnementaux sur les flux de gènes de *C. obsoletus* (chapitre III, article 3).

Les résultats de nos travaux ont permis de (i) clarifier le schéma taxonomique du Complexe Obsoletus/Scoticus en région paléarctique, nos résultats soutenant l'existence de trois clades cryptiques au sein de *C. obsoletus*, de deux clades cryptiques au sein de *C. scoticus* et en questionnant le statut taxonomique de *C. montanus*, (ii) de mettre en évidence une forte dispersion de *C. obsoletus* au-dessus des terres, et (iii) de décrire l'histoire évolutive et démographique de *C. obsoletus* en Europe. L'ensemble de ces résultats met en lumière l'importance du paysage dans la connectivité paysagère des populations, et amène à repenser l'importance de la circulation du virus par le biais des insectes vecteurs.

I) La taxonomie moléculaire de *C. obsoletus* : l'arbre qui cache la forêt

Depuis l'introduction du virus de la FCO et l'émergence du virus de Schmollenberg dans le nord de l'Europe en l'absence de *C. imicola*, vecteur majeur en zone afrotropicale et méditerranéenne, la compétence vectorielle des espèces de *Culicoides* paléarctiques questionne. Du fait de leur abondance élevée dans les élevages de ruminants domestiques et de leurs préférences trophiques mammophiles (Garros and Balenghien, 2017), les espèces du Groupe Obsoletus ont rapidement été suspectées. Une fois que les preuves scientifiques furent établies avec l'isolement du virus dans des populations naturelles de *C. chiopterus* ou *C. obsoletus* et la démonstration au laboratoire de la répllication du virus chez *C. obsoletus* et *C. scoticus*, resta alors le manque de connaissances sur la diversité de ce groupe d'intérêt (Barber *et al*, 2018; Carpenter, 2016; Carpenter *et al*, 2006; Carpenter *et al*, 2008a; Veronesi *et al*, 2013). Cette connaissance est pourtant indispensable aux études de modélisation de la transmission des pathogènes et au développement de mesures de lutte anti vectorielle. En effet, des espèces proches ou cryptiques peuvent avoir des comportements, des biologies ou des écologies différentes et ainsi avoir des implications variables en termes de transmission.

Ces dernières années, les études de phylogénie moléculaire ou de morphométrie ont écarté *C. chiopterus* et *C. dewulfi* du Groupe Obsoletus et identifié l'existence de nouvelles espèces cryptiques, *C. obsoletus* clade O2 et *C. obsoletus* clade dark (Ander *et al*, 2013; Gomulski *et al*, 2005; Hajd Henni *et al*, 2014; Kluiters *et al*, 2016; Meiswinkel *et al*, 2004). Sans réel cadre et consensus à ce niveau taxonomique du groupe, les auteurs continuent parfois à utiliser des schémas taxonomiques anciens, à mélanger les appellations systématiques ou à désigner trop vite de la diversité observée comme « nouvelle espèce ». De plus, certaines études se basent sur l'étude du polymorphisme d'un faible nombre de marqueurs moléculaires, à des échelles géographiques restreintes ou avec de faibles nombres d'individus. Ainsi, la taxonomie du Complexe Obsoletus/Scoticus n'est toujours pas unifiée.

Notre étude est la première à être menée à l'échelle ouest européenne avec un jeu de données important (plus de 3 500 individus) couvrant 20 pays (sur 68 sites) permettant de fournir une idée plus précise de la diversité cryptique au sein du Complexe Obsoletus/Scoticus. Notre révision taxonomique du complexe réalisée avec une approche moléculaire nous a permis de délimiter et valider cinq Unités Moléculaires Taxonomiques Opérationnelles (MOTUs). En plus

de *C. obsoletus* et *C. scoticus*, les taxons *C. obsoletus* clade O2 et *C. obsoletus* clade dark sont soutenus moléculairement. Nous identifions des individus appartenant à une espèce non décrite ou décrite mais non représentée dans les bases de données génétiques. Notre étude interroge le statut taxonomique de *C. montanus*. De plus, *C. scoticus* présente une forte diversité intraspécifique avec l'existence de deux clades soutenus (clade 1 et clade 2).

L'utilisation de séquences provenant de bases de données moléculaires publiques comme GenBank ou BOLD montrent des limites. En effet, nous avons identifié la présence de séquences appartenant à la même espèce mais enregistrées sous des noms différents. Les séquences nommées « *C. obsoletus* O1 » et « *C. obsoletus* O3 » sont identifiées en réalité respectivement *C. obsoletus* et *C. obsoletus* clade dark, après comparaison avec d'autres séquences générées par notre échantillonnage européen (Ander *et al*, 2013). Des interrogations planent toujours autour du statut de *C. obsoletus* clade dark qui pourrait être en réalité *C. gornostaevae* Mirzaeva, 1984. Cette espèce est normalement de plus grande taille que *C. obsoletus* et a une distribution connue limitée à la zone boréale de la Sibérie (Meiswinkel *et al*, 2015). Malheureusement, l'absence de séquences de référence de *C. gornostaevae* dans les bases de données publiques, empêche les investigations supplémentaires avec les séquences disponibles de *C. obsoletus* clade dark. Bien que la faune des *Culicoides* soit bien décrite en Russie et plus largement en région est paléarctique, la barrière de la langue dans les monographies publiées en russe ou chinois, la difficulté d'échange du matériel biologique et l'absence de dépôt des séquences dans les bases de données internationales, empêchent la révision complète de la taxonomie du groupe à l'échelle de l'ensemble de la région paléarctique.

Lors des analyses de délimitation d'espèces, la validité taxonomique de *C. montanus* est a été questionnée du fait du faible niveau de distance génétique observé sur le marqueur mitochondrial Cox1 (distance de l'ordre du niveau intraspécifique) et du fait de la monophylie des deux clades. Les méthodes de délimitation moléculaire utilisées montrent des résultats différents. Dans notre étude, nous concluons au fait que l'approche multilocus Bayesian Poisson Tree Processes (bPTP) est plus robuste que l'approche bGMYC (Bayesian General Mixed Yule Coalescent). La méthode GMYC peut sur-évaluer la diversité spécifique en sur-délimitant (Carstens *et al*, 2013) et identifiant de nombreux clades. Ceci s'explique par le principe de l'analyse qui modélise le point de transition entre l'apparition de nouvelles espèces (cladogénèse) et de simples événements de coalescence allélique, en partant de l'hypothèse que les événements de spéciation sont plus rares que la coalescence allélique (Reid and Carstens, 2012). Sur la base de nos résultats et en considérant la combinaison de notre approche multi-

marqueur avec une taille d'échantillon importante comme robuste (Fujita *et al*, 2012), nous proposons l'hypothèse que *C. montanus* pourrait être non pas une espèce valide mais une sous population fortement divergente constituant un clade d'haplotypes soutenu au sein de *C. obsoletus*. Les outils de délimitations d'espèces se sont multipliés ces dernières années (Carstens *et al*, 2013), avec chacun leurs cadres et leurs limites. A l'avenir, il pourrait être intéressant, de combiner d'autres méthodes à celles testées, comme l'approche spedeSTEM (Ence and Carstens, 2011) pour tester cette hypothèse pour *C. montanus*. Les approches spedeSTEM et BPP (Yang, 2015), deux méthodes dérivées du modèle de coalescence multi-espèces, sont des approches complémentaires à la validation de la délimitation des espèces parce que chacune d'elles semble sujette à l'échec de l'autre (Carstens *et al*, 2013). Lorsque les deux programmes donnent des résultats compatibles, on peut raisonnablement déduire que la sortie de BPP n'a pas été induite en erreur en raison d'un arbre guide mal construit et que la sortie spedeSTEM n'a pas manqué de détecter les lignées indépendantes en raison d'arbres génétiques mal estimés. Inversement, l'incongruité dans les résultats indique probablement que l'un de ces problèmes est probablement évident (Carstens *et al*, 2013). Enfin, pour statuer avec plus de solidité sur le statut de *C. montanus*, et à défaut de pouvoir examiner le spécimen type, voire d'en extraire de l'ADN, l'inclusion d'individus échantillonnés dans la zone où il a été décrit pour la première fois (Kazakhstan) ou en Asie centrale où il est également présent, serait essentielle (Mathieu *et al*, 2011).

Conformément aux études précédentes, Möhlmann *et al*. (Möhlmann *et al*, 2019), nous observons une variation latitudinale de la diversité cryptique du complexe. Cependant, une variation saisonnière dans la structure de la communauté des *Culicoides* est possible (Ander *et al*, 2012) et notre étude souffre d'une hétérogénéité des dates de collecte. Cela peut également expliquer ces variations géographiques de la diversité spécifique au sein du complexe. De plus, le nombre de *Culicoides* capturés lors d'une nuit de piégeage dépend des conditions météorologiques (vent, précipitations) (Mellor *et al*, 2000), ce qui peut impacter la diversité collectée. Ainsi, dans un idéal, il aurait été souhaitable de conduire cette étude avec des échantillons capturés sur la même période (idéalement la période d'abondance des espèces pour potentiellement augmenter les chances de capturer des espèces peu abondantes, au printemps ou en été) et avec des conditions de capture identiques. Pour limiter les effets stochastiques des conditions météorologiques, il peut être préférable d'effectuer des sessions de piégeage de trois nuits de captures consécutives (Diarra *et al*, 2014; Diarra *et al*, 2015). Malgré cette limite, *C.*

obsoletus apparaît comme l'espèce dominante en Europe, confirmant son statut d'espèce généraliste tolérant une large gamme de conditions éco-climatiques. Il est intéressant de noter que la diversité spécifique est faible dans la région des Balkans avec *C. obsoletus* qui constitue l'espèce majoritairement présente. Si *C. obsoletus* est l'espèce dominante à l'échelle européenne, il existe des variations régionales. *Culicoides obsoletus* clade O2 est l'espèce la plus échantillonnée dans les pays nordiques alors que *C. montanus* prédomine au Maroc. *Culicoides scoticus* a été décrit comme plus abondant dans des milieux forestiers d'altitude moyenne (Ippoliti *et al*, 2016). La variation de l'abondance relative des différentes espèces cryptiques dans le Complexe Obsoletus/Scoticus pourrait illustrer des différences de niches écologiques (Goffredo *et al*, 2016). *Culicoides obsoletus* clade dark tient son nom d'une teinte foncée qui pourrait traduire une hausse de la valeur sélective des individus plus foncés dans certains environnements forestiers.

Les méthodes de délimitation permettent d'identifier des lignées évolutives distinctes, mais pour décrire des entités que l'on qualifie d'espèces au sens biologique du terme, d'autres étapes sont nécessaires. Un processus complet de caractérisation de la morphologie, de l'écologie et de la biologie de *C. montanus* est essentiel pour statuer définitivement sur la validité taxonomique de l'espèce. D'ailleurs, peu d'éléments de l'écologie de *C. montanus* sont connus. Ce processus de caractérisation est également à envisager pour les clades supportés phylogénétiquement et identifiés comme des espèces à part entière par l'analyse de délimitation d'espèces (*C. obsoletus* clade O2 et *C. obsoletus* clade dark). Cette étape sera longue et fastidieuse et demandera une forte expertise. Tous les individus devront être disséqués, montés entre lames et lamelles, typés moléculairement afin de pouvoir examiner *a posteriori* les éléments morphologiques et identifier des différences morphologiques ou morphométriques. Cette caractérisation devra être réalisée sur les deux sexes. En effet, dans les complexes d'espèces les femelles ne sont généralement pas distinguables morphologiquement, comme c'est le cas entre *C. obsoletus* et *C. scoticus* ou *C. montanus*, mais les mâles sont différenciables sur la base des appareils reproducteurs. Cependant, les techniques de piégeages utilisées ne permettent pas de capturer des adultes mâles en grand nombre.

L'existence d'une telle diversité questionne sur leurs compétences vectorielles respectives. La caractérisation de la compétence et des préférences trophiques de cette nouvelle diversité cryptique sera longue et compliquée sachant les difficultés méthodologiques à manipuler les *Culicoides* vivants au laboratoire et sur le terrain et à la difficulté de les élever. Sur la base des

évidences pour *C. obsoletus* de son rôle majeur dans la transmission des virus de la FCO et de Schmallenberg et de la faible divergence génétique entre les taxons du Complexe Obsoletus/Scoticus, nous invitons dès à présent, par principe de précaution, à considérer ces espèces comme de potentielles espèces vectrices des virus d'intérêt vétérinaire transmis par les *Culicoides* en Europe.

Pour conclure, nos résultats permettent de présenter une taxonomie moléculaire du Complexe Obsoletus/Scoticus clarifiée à l'échelle européenne avec des hypothèses fortes, mais cette histoire demeure encore incomplète. En effet, *C. obsoletus*, est également présent en région néarctique (points de captures au Canada et aux Etats-Unis). L'absence de séquences présentes dans les bases de données et l'intérêt moindre de cette espèce dans ces régions, peut-être du fait de son abondance plus limitée, font qu'à l'heure actuelle, il est impossible de statuer sur la diversité dans cette région. Une étude incluant la globalité de la zone de distribution de *C. obsoletus* permettrait de conforter ou non les hypothèses émises. Elle permettrait aussi par la même occasion de reconstruire l'histoire évolutive de *C. obsoletus* sur l'ensemble de son aire de distribution actuelle.

II) L'histoire évolutive de *C. obsoletus* en Europe : un morceau du puzzle

L'utilisation combinée d'un marqueur mitochondrial (Cox1) et de marqueurs microsatellites dans des analyses de phylogéographie et de génétique des populations nous a permis d'identifier (i) la structure génétique et phylogéographique à l'échelle continentale, (ii) l'histoire évolutive et démographique (iii) et la dynamique du flux génétique de *C. obsoletus* espèce autochtone en Europe. Contrairement à *C. imicola*, espèce afrotropicale dont le scénario d'invasion a été élucidé par Jacquet *et al.* (Jacquet *et al.*, 2015), nous présentons ici un scénario de recolonisation post-glaciaire de *C. obsoletus* en Europe.

Le scénario le plus probable que nous avons obtenu est une recolonisation d'est en ouest, depuis la région des Balkans et des refuges plus nordiques. Les populations ouest européennes seraient issues d'une admixture des populations est européennes. Ce scénario d'expansion géographique et démographique est construit sur les différents résultats de taxonomie moléculaire, de phylogéographie et de génétique des populations obtenus lors de nos travaux.

Tout d'abord, bien que *C. obsoletus* soit autochtone en Europe, il est retrouvé dans des proportions différentes en fonction de la zone géographique. L'étude de taxonomie moléculaire à l'échelle européenne du Complexe *Obsoletus/Scoticus* nous a permis d'observer une présence très importante de *C. obsoletus* dans la région des Balkans (Mignotte *et al.*, 2020). Cette étude confirme les études entomologiques antérieures dans la même région (Roumanie, Slovaquie Grèce, Albanie, Macédoine, Kosovo, Montenegro, Serbie, Bosnie-Herzégovine, Bulgarie et Croatie) (Goffredo *et al.*, 2004; Hristescu *et al.*, 2020; Pudar *et al.*, 2018; Sarvasova *et al.*, 2014). Au contraire, il est moins abondant dans le nord de l'Europe (Cuellar *et al.*, 2018; Mignotte *et al.*, 2020). Les refuges glaciaires sont habituellement considérés comme des hotspots de biodiversité, et l'on s'attendrait à observer une diminution de la diversité intraspécifique en dehors des refuges, à la suite des événements fondateurs successifs au cours de la colonisation post-glaciaire (Hewitt, 1999; Hewitt, 2004). Cependant, il a été montré, sur des espèces forestières que les populations génétiquement les plus diverses n'étaient pas situées dans le sud, au niveau des refuges glaciaires, mais à des latitudes intermédiaires, une conséquence probable du mélange (admixture) de lignées divergentes colonisant le continent à partir de refuges séparés (Petit *et al.*, 2003). Au sein du Complexe *Obsoletus/Scoticus* également, la plus large diversité spécifique est observée en Europe de l'ouest aux latitudes intermédiaires.

La structure phylogéographique de *C. obsoletus* à l'échelle européenne est basée sur le polymorphisme du marqueur mitochondrial Cox1, classiquement utilisé pour reconstruire la généalogie des espèces (de Meeus *et al*, 2007). Le réseau d'haplotypes montre une absence de structuration phylogéographique spatiale. Ainsi, des individus échantillonnés à des milliers de kilomètre d'écart (par exemple Finlande et Espagne) partagent le même haplotype, créant une absence d'isolement par la distance (hypothèse attendue à cette échelle d'étude). La configuration en étoile (star-like) du réseau d'haplotypes, avec des haplotypes majoritaires partagés par la quasi-totalité des populations étudiées et des haplotypes uniques étroitement liés à l'haplotype central communs est typiquement associée à un scénario d'événements de colonisation passée, suivi d'expansions démographiques récentes des populations européennes. Ces éléments sont aussi confirmés par d'autres analyses de génétique des populations (analyse de mismatch, test D de Fu et Tajima) indiquant des tailles de population en expansion.

Grâce à un travail réalisé avant ces travaux de thèse, nous avons pu bénéficier d'un jeu de 11 marqueurs microsatellites pour compléter nos études. Ce jeu de données microsatellites, utilisé pour la première fois, apporte un niveau d'information différent du marqueur mitochondrial. A l'échelle européenne, l'analyse bayésienne de regroupement sur les données microsatellites a identifié trois groupes génétiques majoritaires : un cluster balkanique (Grèce et Turquie), un cluster exclusivement constitué de la population des îles Canaries, et un cluster comprenant tout le reste de l'Europe. On peut considérer cette structuration comme extrêmement faible au vue de l'échelle géographique considérée. A cette échelle, des patrons de structuration faible sont retrouvés dans d'autres régions, par exemple détectés en Australie pour *C. brevitarsis* (Onyango *et al*, 2015a), dans le sud-est des Etats-Unis (Texas, Floride, Caroline du sud) pour *C. stellifer* (Shults *et al*, 2020) et en Afrique sub-saharienne pour *C. imicola* (Jacquet *et al*, 2016a).

Au Maroc, le faible effectif de *C. obsoletus* dans les populations sud marocaines, au profit de *C. montanus*, n'a pas permis d'inclure une population continentale proche de la population insulaire des îles Canaries dans les analyses de génétique des populations. Ceci aurait permis de tester l'origine de cette population, la plus méridionale de notre échantillonnage. Nous faisons l'hypothèse soit d'une origine nord-africaine via de la dispersion passive par les vents, soit d'une origine Européenne via des activités commerciales. La faible structuration génétique observée avec les marqueurs microsatellites à l'échelle continentale (Partie II) est retrouvée de façon exacerbée à l'échelle nationale (Partie III) où cette fois-ci, aucun cluster génétique spatial n'est observable. Cette faible structure s'accompagne d'une faible différenciation génétique

inter-populationnelle (Partie II) et inter-individuelle (Partie III). Après avoir vérifié l'efficacité, le polymorphisme et l'absence de biais de nos marqueurs (tests de déséquilibre de liaison, de détection d'allèle nul, de dominance d'allèle court, lecture en double aveugle des profils lors du génotypage), nous sommes confiants sur le fait que la diversité et la structure observée est bien le reflet de processus micro-évolutifs entre les populations. Ainsi, la combinaison de différents phénomènes évolutifs peut expliquer une si faible structure génétique.

Le premier phénomène est un flux de gènes intense, multidirectionnel et fréquent entre les populations, au point que la diversité génétique se serait homogénéisée à l'échelle continentale, contrebalançant ainsi totalement l'impact de la dérive génétique. Ce flux de gènes serait facilité par la dispersion active de proches en proches des *Culicoides* et la dispersion passive par le vent. La part relative des phénomènes de dispersion à courte et à longue distance sur les flux de gènes pourrait être étudiée avec des méthodes d'assignation bayésienne permettant l'estimation des taux de migration récents, implémentées dans les logiciels BAYESASS (Wilson and Rannala, 2003) et GENECLASS2 (Piry *et al*, 2004). GENECLASS2 est un logiciel qui utilise des génotypes multilocus, comme les microsatellites, pour assigner à des individus des populations d'origine. Cette assignation permet de détecter des individus migrants de premières générations et donc d'estimer la dispersion (Paetkau *et al*, 2004; Piry *et al*, 2004; Rannala and Mountain, 1997). Le logiciel BayesAss serait complémentaire car il permet d'estimer les taux de migration récents entre des populations mais aussi le nombre de générations depuis l'immigration. Cependant, ces méthodes, bien que prometteuses, possèdent des limites pour l'étude de la dispersion chez le modèle *Culicoides*. Par exemple, dans BayesAss, à chaque génération, la proportion de migrants dans une population ne peut pas dépasser un tiers du total de la population. Cette supposition d'un taux de migration faible va à l'encontre des connaissances actuelles sur la dispersion des *Culicoides*. Plus problématique encore, la précision de l'analyse est augmentée lorsque la différenciation de la population est élevée, ce qui n'est pas notre cas (Wilson and Rannala, 2003). Les résultats de telles analyses devront alors être considérés en tenant compte de ces limites. Il serait également judicieux d'utiliser un simulateur de flux de gènes dans l'espace comme implémenté dans le logiciel CDPOP (Cost Distance POPulations) (Landguth and Cushman, 2010), pour tenter d'évaluer plus précisément le nombre de migrants nécessaire à chaque génération pour obtenir le patron de structuration observée en France et en Europe.

Le second phénomène pour expliquer cette absence de structuration serait une très forte taille efficace de population suffisante pour contrebalancer l'action de la dérive génétique. La

domestication des ruminants et l'intensification récente des activités d'élevage peuvent être responsables de l'augmentation de la densité et du nombre des hôtes potentiels, permettant le déplacement de proches en proches fréquent et l'expansion démographique locale rapide, expliquant l'homogénéisation génétique. Les activités des populations humaines peuvent ainsi éliminer la barrière géographique et réduire l'isolement génétique dû à la distance. L'estimation des tailles efficaces apparaît alors comme une composante importante de futures études sur *C. obsoletus*. L'estimation précise des tailles efficaces de populations est délicate. Celles-ci pourraient être investiguées par des études directes de CMR, mais ces méthodes ne sont pas adaptées aux populations de grande taille (Cayuela *et al*, 2018). Il existe différentes familles de méthodes permettant d'inférer l'effectif efficace de populations à l'aide de marqueurs moléculaires. Les méthodes dites temporelles permettant d'estimer l'effectif efficace de variance nécessitent de disposer d'échantillonnages réalisés à différentes dates et de connaître le temps de génération de l'espèce étudiée. La deuxième grande famille de méthodes permet d'estimer l'effectif efficace de consanguinité par l'étude des variations spatiales des fréquences alléliques mais ces méthodes nécessitent de connaître les taux d'autofécondation et de recombinaison entre loci utilisés. Enfin, la dernière grande famille de méthodes utilise l'information des déséquilibres de liaison (Bartley *et al*, 1992) dans une seule population. Cependant celle-ci est biaisée quand la taille des échantillons est inférieure au N_e mais un logiciel (LDNe) propose une correction de ce biais (Waples and Do, 2008) mais donne rarement des résultats utilisables.

Une histoire évolutive contemporaine peut également engendrer cette absence presque totale de différenciation génétique. Ainsi, même si les flux de gènes actuels ne sont pas effectifs entre certaines localisations éloignées, un temps de divergence trop court ne permettrait pas encore de détecter l'impact de la dérive génétique sur la diversité génétique des populations étudiées. Ce type de patron est typiquement observé lors de phénomènes de recolonisation récente (Hewitt, 2004). Ce scénario est en accord avec l'expansion démographique récente mise en évidence avec le marqueur mitochondrial.

Les variations climatiques du Cénozoïque ont marqué de façon décisive l'évolution du vivant. Les alternances de refroidissement et de réchauffement ont cantonné les espèces vivantes à des habitats spécifiques. La recolonisation suite aux événements de glaciations du Pléistocène est bien documentée. De nombreuses études de génétique des populations ont décrit des expansions démographiques depuis les trois principaux refuges primaires européens : les péninsules ibérique, italienne (l'Apennins) (Italie) et balkanique (les Balkans et l'Anatolie (Turquie))

(Boissin *et al*, 2016; Ficetola *et al*, 2007; Provan and Bennett, 2008; Taberlet and Cheddadi, 2002; Taberlet *et al*, 1998). L'histoire démographique de la plupart des espèces européennes est marquée par une contraction (réduction de la taille des populations lors de la glaciation) puis une expansion démographique et géographique à la suite des épisodes glaciaires du Pléistocène (Bowler and Benton, 2005; Hewitt, 2004; Phillips *et al*, 2010; Widmer and Lexer, 2001). Cette expansion spatiale des populations s'accompagne souvent d'un gradient de réduction de la diversité génétique au cours du processus de colonisation (Comes and Kadereit, 1998), causé par des événements fondateurs en série (goulot d'étranglement) (Dlugosch and Parker, 2008; White *et al*, 2014). La diversité génétique est en moyenne plus élevée dans les zones de refuge par rapport aux zones de redéploiement post-glaciaire. Nous observons une richesse allélique plus importante de la région des Balkans suggérant que cette région puisse être considérée comme une zone d'origine des populations européennes (Hewitt, 1999; Petit *et al*, 2003; Prugnolle *et al*, 2005; Taberlet *et al*, 1998; White *et al*, 2013).

Différents scénarios de recolonisation ont été testés à partir de refuges glaciaires en Europe (en Espagne et dans les Balkans) par la méthode ABC. Le scénario le plus probable indique une origine est paléarctique des populations de *C. obsoletus* avec une recolonisation post-glaciaire depuis des refuges balkanique et nord européen, suivi d'une admixture entre les populations issues de ces deux refuges et d'une extension vers l'Europe de l'ouest. La divergence entre les populations est et ouest serait datée d'environ 10 000 ans. L'origine nord est paléarctique des populations ancestrales peut s'expliquer par l'existence de refuges dans les limites des calottes glaciaires où de nombreuses espèces végétales et la faune ont survécu au cours du dernier maximum glaciaire (Abbott *et al*, 2003; Lagercrantz and Ryman, 1990). Au vu de ses préférences trophiques larges et de sa grande plasticité écologique lui permettant de survivre à de hautes altitudes, il paraît probable que *C. obsoletus* ait pu persister dans des refuges glaciaires dispersés et septentrionaux (Provan and Bennett, 2008; Tzedakis *et al*, 2013).

En effet, certaines études ont suggéré que des refuges puissent être présents en dehors du sud de la péninsule des Balkans, comme en Europe occidentale, centrale et orientale (Schmitt and Varga, 2012; Stewart and Lister, 2001). Ainsi, au terme de l'épisode glaciaire, les populations situées dans la région balkanique et celles situées dans des refuges plus au nord ont pu entrer en contact et créer ce patron d'admixture (Schmitt and Varga, 2012; Taberlet *et al*, 1998). Il est à noter que le faible nombre de *C. obsoletus* identifiés en Italie lors de l'étape de barcoding moléculaire (Mignotte *et al*, 2020), ne nous a pas permis d'inclure les populations italiennes dans le jeu de données de géotypages microsatellites. Pour de futurs travaux, l'intégration de

populations de la péninsule italienne serait intéressante et permettrait de tester si cette région a servi de refuge durant les épisodes de glaciation afin de pouvoir appréhender l'intégralité de l'histoire évolutive de cette espèce dans la zone paléarctique occidentale.

Pour réaliser nos analyses ABC, nous avons utilisé le logiciel DIYABC pour inférer l'histoire évolutive des populations de *C. obsoletus*. Une des limites de ce logiciel (Cornuet *et al*, 2014) est qu'il ne permet pas d'inclure des phénomènes de dispersions récurrentes entre les populations dans les scénarios. Or dans le cas de *C. obsoletus*, qui est une espèce qui se disperse fortement, il aurait été informatif de pouvoir inclure de la migration dans nos scénarios. La dispersion de cette espèce est bien documentée, elle a certainement un rôle important dans l'absence de structure génétique et phylogéographique observée. Elle freine certainement aussi l'identification des événements de divergence ancienne car la variabilité génétique intra- et inter-populationnelle est brouillée par les flux de gènes contemporains (Edwards and Bensch, 2009). Des freins similaires ont été décrits par exemple lors de l'invasion accidentelle en Europe de *Leptoglossus occidentalis*, la punaise américaine du pin de l'ordre des hémiptères. Sa capacité de dispersion importante a perturbé la description des routes d'invasion de l'espèce (Lesieur *et al*, 2018). Une perspective intéressante est l'estimation des flux géniques par coalescence, implémentée dans le logiciel MIGRATE (Beerli and Felsenstein, 2001). Ce logiciel permet d'estimer des flux de gènes, entre populations, qu'ils soient historiques ou asymétriques, en tenant compte de la structure génétique globale. Malheureusement, ce type d'analyse peut être très chronophage, ce qui ne nous a pas permis son utilisation dans le cadre de cette thèse (Mardulyn *et al*, 2013).

Les discordances mito-nucléaires

L'utilisation combinée de marqueurs microsatellites (nucléaires) et mitochondriaux (cytoplasmiques) est complémentaire, car les deux types de marqueurs possèdent une héritabilité, des taux de mutation et des ploïdies différentes. En raison de ces caractéristiques, des discordances mitochondriale-nucléaire sont régulièrement signalées dans les études utilisant une approche multi-marqueurs (Kolleck *et al*, 2013; Rasic *et al*, 2015). C'est également ce que nous avons observé. La structure génétique obtenue avec nos 11 marqueurs microsatellites et celle obtenue avec le marqueur mitochondrial Cox1 sont différentes, bien qu'elles aient la caractéristique commune d'être faiblement structurées. Aucun patron de structuration spatiale de la diversité génétique n'est observable avec notre marqueur

mitochondrial. La structuration observée dans l'est du bassin méditerranéen et dans les îles Canaries avec les microsatellites n'est pas détectée avec le marqueur mitochondrial.

Une première explication repose sur la ploïdie et le mode d'héritabilité uniparentale maternelle de l'ADNmt. En théorie, ces caractéristiques réduisent de quatre fois la taille effective de ces marqueurs par rapport à des marqueurs nucléaires diploïdes avec une héritabilité biparentale (microsatellites). La conséquence directe de cette plus petite taille efficace est un taux d'extinction des allèles ancestraux plus élevé chez l'ADNmt (Zhang and Hewitt, 2003) que dans l'ADN nucléaire. La discordance mitonucléaire peut également survenir s'il existe des différences dans la manière dont la sélection agit sur le génome mitochondrial par rapport au génome nucléaire ou s'il y a un mouvement biaisé de l'un ou l'autre type de marqueur en raison d'asymétries démographiques, telles que des biais de dispersion liés au sexe (Rheindt and Edwards, 2011). Si la sélection des variantes d'ADNmt varie géographiquement, des schémas discordants entre l'ADN mitochondrial et nucléaire peuvent apparaître (Irwin, 2012). Nos résultats ne montrent pas de signal de sélection positive, mais seulement quelques sites sous sélection négative sur l'ADNmt. Cependant, ces sites sélectionnés négativement se trouvent dans toutes les zones d'échantillonnage, ce qui exclut la sélection comme facteur explicatif des schémas génétiques spatiaux observés (Murrell *et al*, 2012). La dispersion par les femelles peut également favoriser la dispersion de l'ADNmt en l'absence de mouvements concomitants de l'ADN nucléaire (Funk and Omland, 2003) Si la dispersion des *Culicoides* par les vents sur de longues distances est indépendante du sexe, ce n'est pas le cas de la dispersion active sur de courtes distances. En effet, chez les insectes se nourrissant de sang, les femelles parcourent généralement de plus longues distances que les mâles à la recherche de sites de ponte et de repas sanguins. La combinaison de ces différents mécanismes pourrait donc expliquer les légères différences de signal entre les deux types de marqueurs.

Espèces cryptiques comme indices d'évènements évolutifs

La question de l'isolement reproductif entre les espèces cryptiques du Complexe *Obsoletus/Scoticus* est importante. En effet, si celui-ci était encore incomplet, le fait de concentrer notre étude de génétique des populations et du paysage sur *C. obsoletus* au sens strict aurait pu contribuer à gommer une partie du signal de différenciation génétique observée. Malheureusement, pour pouvoir tester l'isolement reproductif entre espèces, il faudrait pouvoir disposer d'élevages en laboratoire des différentes espèces et clades de *Culicoides* afin de

réaliser des accouplements croisés. Nous l'avons décrit par ailleurs, la manipulation de *Culicoides* au laboratoire est très difficile et à ce jour, seules des colonies de deux espèces ont pu être établies au laboratoire (*C. sonorensis*, *C. nubeculosus*). Ainsi, il serait intéressant de lever notre restriction de départ et prendre en compte les espèces cryptiques les plus proches de *C. obsoletus*, dans les études phylogéographiques (*C. montanus*, *C. obsoletus clade O2*, *C. obsoletus clade dark*) pour espérer mieux reconstruire l'expansion géographique passée de *C. obsoletus*. Lors de la mise au point des marqueurs microsatellites pour *C. obsoletus*, des populations contenant un mélange d'individus des clades O2, dark et de *C. montanus*, ont été génotypées. Chez ces individus, on a pu déplorer beaucoup de loci non amplifiés et observer des allèles à certains loci en dehors des tailles attendues (Huber, communication personnelle). Ceci laisse à penser que ces clades présentent une divergence très forte de *C. obsoletus*. Cependant, l'inclusion dans une étude de l'ensemble de ces taxons ayant récemment divergé pourrait conduire à une meilleure compréhension des processus qui ont conduit à la différenciation génétique actuelle des populations et des espèces.

Une histoire évolutive encore incomplète

Si notre étude permet de proposer une histoire pour la région ouest paléarctique, l'histoire globale n'est pas encore complète. En effet, bien que son centre de distribution semble être en Europe, *C. obsoletus* est présent dans de nombreux pays de l'hémisphère nord (Anderson *et al*, 1991; Borkent, 2016; Gao *et al*, 2020; Mathieu, 2011; Sprygin *et al*, 2014). Il est présent sur l'ensemble de la région paléarctique du Portugal (Capela *et al*, 2003) jusqu'à la Chine (Gao *et al*, 2020). La présence de l'espèce en Amérique et en Asie questionne quant au scénario évolutif ayant permis cette large répartition géographique (Borkent, 2016). Nous émettons l'hypothèse, en raison de l'histoire évolutive récente européenne, que ce passage du continent asiatique au continent américain a pu se faire durant les derniers épisodes de glaciations du Pléistocène par le détroit de Béring, encore présent jusqu'à il y a 11 000 ans. *Culicoides obsoletus* est retrouvé des deux côtés de l'Océan Pacifique, côté oriental dans la région de Magadan à l'extrême est de la Russie (Sprygin *et al*, 2014), et côté occidental en Colombie Britannique sur la côte ouest du Canada (Anderson *et al*, 1991). Durant les épisodes successifs de glaciation du Pléistocène, le détroit de Béring, en raison de l'abaissement du niveau de la mer, permettait la connexion des deux principaux refuges de glaciation arctique, la Béringie occidentale (Yukon, Canada) et la Béringie orientale (Sibérie) (Bowden and Buddle, 2015; DeChaine, 2008; Klimaszewski *et al*, 2015). Par exemple, durant les glaciations du Pléistocène, des espèces de diptères du Yukon

comme *Meromyza columbi*, *Incertella incerta* et *Trioxoscelis fumipennis* ont persisté dans le refuge glaciaire de Béringie occidentale (Solecki *et al*, 2019). De plus, l'écosystème qui était retrouvé en Béringie est semblable à celui dans lequel *C. obsoletus* peut évoluer : un habitat à humidité moyenne, similaire aux forêts tempérées à feuilles caduques, un climat aride, de la steppe-toundra arbustive, avec la présence de plantes herbacées et de grands mammifères (bœufs musqués à cornes, chevaux, hommes). Il est raisonnable d'imaginer que *C. obsoletus*, grâce à son comportement trophique généraliste et à sa forte plasticité écologique, ait lui aussi pu persister dans ce refuge et même passer de la Béringie orientale à la Béringie occidentale, via ses fortes capacités de dispersion passive. Un tel scénario pourrait être mis à l'épreuve en incluant des individus provenant de l'ensemble de l'aire de répartition de *C. obsoletus*.

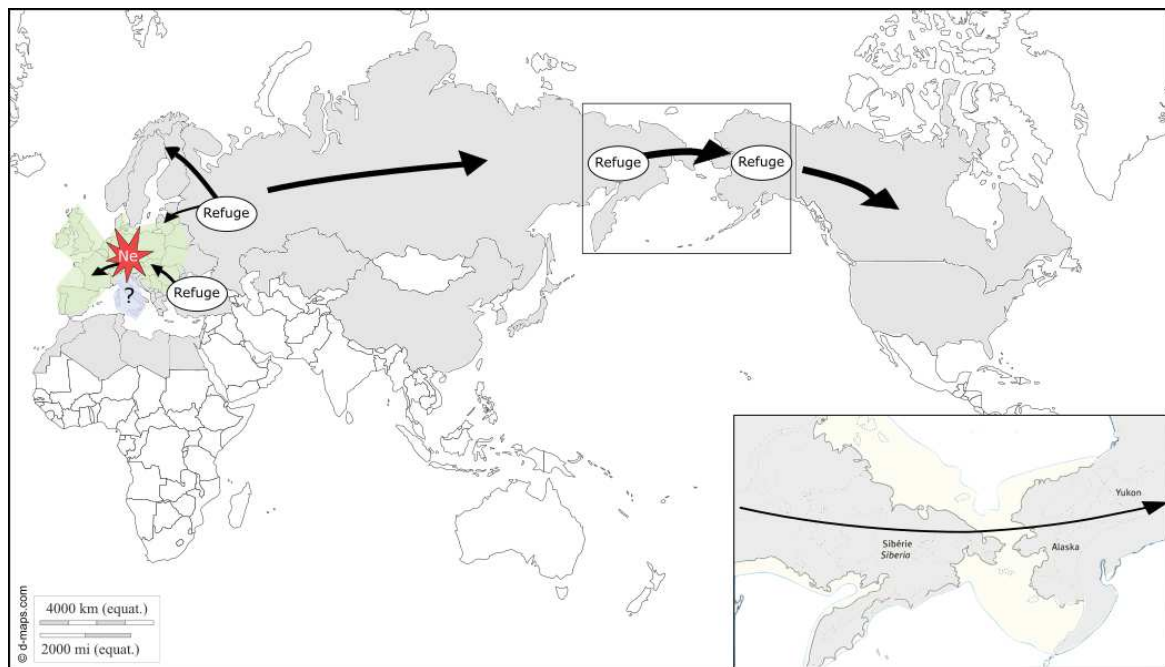


Figure 1. Hypothèse de scénario d'expansion mondiale de *C. obsoletus*. Les zones grises indiquent des pays dans lesquels *C. obsoletus* est présent. Les zones vertes représentent des zones d'admixture des populations. La zone bleue correspond à une zone en Europe qui n'a pas pu être incluse dans nos analyses mais qui devra être considérée dans le futur. Les cercles blancs indiquent l'emplacement de refuges glaciaires. Ne désigne une expansion démographique. Le beige désigne l'abaissement du niveau de la mer durant la glaciation.

Nos travaux mettent en évidence que *C. obsoletus* possède une forte capacité de dispersion lui permettant de brouiller les signaux de structuration génétique ancienne et d'homogénéiser la structure génétique contemporaine. Mais quels sont les facteurs environnementaux qui favorisent la dispersion des *Culicoides* ?

III) De la direction du vent à la densité d'hôtes : les facteurs environnementaux qui favorisent la dispersion des *Culicoides*

Pour la première fois, nous avons quantifié l'impact de l'hétérogénéité du paysage sur le flux de gènes d'une espèce de *Culicoides*. Une large sélection de facteurs environnementaux et météorologiques, classiquement décrits comme impactant la dispersion des *Culicoides*, a été incluse dans ces analyses. L'étude a été menée à l'échelle du pays (France), à l'aide de 11 marqueurs microsatellites et d'une approche individuelle visant à maximiser le nombre d'habitats représentés. Des analyses univariées et multivariées, étudiant la corrélation entre distance environnementale et génétique, ont mis en évidence l'absence ou le faible impact de la plupart des facteurs environnementaux et météorologiques testés sur les mesures inter-individuelles de différenciation génétique. Aucun isolement par distance n'a été détecté à l'échelle française, mais une analyse anisotrope d'isolement par distance a révélé un isolement significatif sur un axe nord-sud. Ces résultats conduisent à s'interroger sur l'impact de la dispersion des *Culicoides* à longue distance sur la transmission du virus, qui pourrait être sous-estimée.

Deux hypothèses peuvent être avancées pour expliquer l'absence de corrélation entre distance génétique et résistance paysagère.

La première est celle d'une dispersion à très faible distance, qui se passerait à une échelle bien plus petite que celle que nous avons considérée. Notre grain d'étude serait alors trop grand pour pouvoir détecter un signal (Phillipsen *et al*, 2015). Il pourrait alors être judicieux, lors de prochaines études sur la dispersion de *C. obsoletus*, d'utiliser une échelle géographique beaucoup plus fine, qui permettrait d'appréhender plus localement l'impact du paysage sur la dispersion des *Culicoides*. Pour cela, un plan d'échantillonnage centré autour de potentielles barrières à la dispersion (relief, cours d'eau, route, ville) avec des degrés de connectivité variés et des paysages hétérogènes devra être privilégié. Par exemple, un échantillonnage de *C. obsoletus* et *C. chiopterus*, d'ores et déjà disponible, effectué entre fermes distantes de quelques kilomètres, nous permettrait (i) de suivre plus finement la dispersion des *Culicoides* et (ii) de comparer la dispersion de deux espèces aux préférences trophiques très différents. A ce plan d'échantillonnage devront être couplées des méthodes d'analyse de parenté qui visent à

identifier les lignées génétiques de la génération précédente. Ces méthodes requièrent un échantillonnage très important car les liens de parenté entre individus doivent être reconstitués entièrement. La deuxième hypothèse est celle d'une dispersion très peu impactée par les facteurs environnementaux qui ont été testés. À la vue des larges préférences trophiques de *C. obsoletus* et de la dispersion stratifiée des *Culicoides*, il semble que la seconde hypothèse est à privilégier.

L'utilisation couplée d'approches de régression MRDM (*multiple regression of distance matrices*) et de commonalité nous a tout de même permis de mettre en évidence un impact significatif de la densité d'hôtes sur la différenciation génétique de *C. obsoletus*. Cependant, la part explicative de ce facteur est faible et un tel résultat n'est pas surprenant à la vue de la faible différenciation génétique observée. La densité d'hôtes comme facteur de conductance pour les *Culicoides* est cohérente avec la biologie de l'espèce. *Culicoides obsoletus* est opportuniste, et les bovins font partie de ses hôtes potentiels (Bartsch *et al*, 2009; Garros *et al*, 2011; Lassen *et al*, 2011). Ce résultat indiquerait une dispersion active importante à courte distance liée à la recherche d'hôtes.

L'utilisation d'une approche anisotropique nous a permis d'identifier un isolement anisotropique significatif en France, pouvant indiquer une dispersion par le vent. A l'heure actuelle, la direction des vents ne peut pas être incluse directement dans un *workflow* comme celui que nous avons utilisé. Les éléments quantitatifs du paysage (précipitations, taux d'humidité, dénivelé) et les éléments qualitatifs (couverture paysagère) peuvent être représentés dans un raster par une valeur numérique. Or cela est impossible pour la direction du vent où chaque pixel représente un vecteur indiquant un angle. Une approche prometteuse est celle proposée par le package « Rwind » (Fernández-López and Schliep, 2018). Ce package permet via les fonctions « flow.dispersion » et « costDistance » de générer une matrice de conductance par le vent, tenant compte à la fois de la vitesse et de la direction du vent. Malheureusement, la trop faible différenciation génétique que nous avons rencontrée ne nous a pas permis d'exploiter cet outil. Celui-ci pourrait être résolutif dans le cadre d'une étude à plus faible échelle.

L'approche individuelle présente de nombreux avantages en génétique du paysage dont le plus important est la multiplication du nombre de points d'échantillonnage dans le paysage et est très adaptée aux espèces présentant un continuum de la répartition. Cependant, dans le cas de l'étude à l'échelle française, elle ne nous a pas permis de trancher quant à l'explication de l'isolement anisotropique significative observé sur l'axe sud-nord, entre un effet de bord et un impact du vent.

Isolement par l'environnement et « gravity model »

Les flux de gènes entre populations sont dépendants de la résistance paysagère durant la phase mobile de la dispersion. Nous l'avons évalué ici par nos travaux avec une approche MRDM entre matrices de distances environnementales et génétiques. Cependant, les caractéristiques biotiques et abiotiques locales des populations influencent également les flux de gènes lors de la phase d'émigration et d'immigration (accès aux ressources, disponibilité des partenaires de reproduction ou capacité d'accueil limitante) (Murphy *et al*, 2010). Ainsi, le paysage n'impacte pas seulement la phase de transfert entre deux sites, mais également les phases d'émigration et d'immigrations dépendantes des conditions environnementales locales. Ces limitations de flux de gènes dues aux caractéristiques locales sont décrites comme « l'isolement par l'environnement ». Les flux de gènes sont alors la combinaison de l'isolement par la résistance (entre les populations) mais également l'isolement par l'environnement (caractéristiques locales des populations) (Pfluger and Balkenhol, 2014). Des approches dites « gravity model » (Murphy *et al*, 2010), intègrent des caractéristiques paysagères entre deux sites mais également celles du site au niveau local, permettant ainsi d'estimer l'isolement par la résistance et l'isolement par l'environnement. Ce type d'approche a été utilisé pour l'étude de la génétique du paysage du virus de la FCO (Jacquot *et al*, 2017), du virus *Puumala* ou de la bactérie *Borrelia burgdorferi* (Barrios *et al*, 2012). Cependant, l'intégration de l'isolement par l'environnement requiert des moyens supplémentaires importants. Les facteurs environnementaux faisant varier la qualité d'un site et impactant l'activité des *Culicoides* sont nombreux : température, humidité, couverture paysagère, densité d'hôtes, direction du vent, vitesse du vent, turbulences climatiques, précipitations ou encore la présence de sites de repos (Acevedo *et al*, 2010; Cuellar *et al*, 2020; Diarra *et al*, 2014; Kettle *et al*, 1998; Mellor *et al*, 2000; Tsutsui *et al*, 2011). La quantité d'informations à inclure dans le modèle devient alors trop grande, d'autant plus que la biologie des *Culicoides* est encore inconnue sur de nombreux aspects. Ainsi, quantifier la qualité des habitats est une tâche compliquée. De plus, il serait nécessaire d'estimer la densité des populations d'hôtes domestiques et sauvages sur chaque site d'échantillonnage, ce qui constituerait une information précieuse sur la qualité de l'habitat local.

Optimisation du coût du paysage

L'absence de corrélation entre les distances génétiques et la résistance paysagère n'est pas forcément synonyme d'absence d'impact du paysage sur la dispersion de *C. obsoletus*. Lors de nos analyses, nous avons attribué des coûts à la résistance paysagère en désignant des seuils de résistance paysagère (K=10, K=100, K=1000) (Dellicour *et al*, 2019). Plus cette valeur augmente, plus le facteur environnemental représente une résistance ou une conductance importante. Il était prévu que ces valeurs de coût soient optimisées pour chaque variable environnementale grâce au package R « ResistanceGA ». Cependant, à la vue de la très faible structuration génétique globale, et de la très faible part de différenciation génétique expliquée par les facteurs environnementaux, cette approche n'a pas pu être mise en œuvre.

La génétique du paysage appliquée à d'autres espèces de *Culicoides*

Nos travaux illustrent une limite de la génétique du paysage. Comment estimer l'impact du paysage sur les flux de gènes d'une espèce lorsque ces flux de gènes sont si importants qu'aucune structure génétique n'est détectable ?

L'utilisation d'une espèce plus rare, moins généraliste, et avec des préférences trophiques plus spécifiques pourrait être un autre choix pour estimer la connectivité paysagère des *Culicoides* (Engler *et al*, 2014; Louy *et al*, 2007). *Culicoides chiopterus* est une espèce de la région paléarctique décrite comme vectrice et beaucoup moins abondante que *C. obsoletus*. Sa biologie et son écologie sont liées aux bovins pour le repas de sang et les gîtes de ponte. Nous faisons l'hypothèse qu'en raison de ses caractéristiques biologiques, l'espèce devrait avoir une structure génétique plus marquée et subir une résistance paysagère plus importante en relation avec la densité de bovins. De plus, *C. chiopterus* semble également posséder une capacité de dispersion importante comme en témoigne la capture d'individus gorgés sur des bovins à plus de deux kilomètres des premiers bovins (Garros *et al*, 2011). Une approche comparative de la distribution spatiale des fréquences alléliques entre *C. obsoletus* et *C. chiopterus* serait informative.

Nos travaux constituent une première étape de caractérisation de la dispersion stratifiée des *Culicoides* par approche indirecte au-dessus des masses continentales. Ils soulignent l'importance des capacités de dispersion active et passive au-dessus des terres des *Culicoides*. Quels impacts ce niveau de dispersion peut-il avoir sur les mesures de lutte anti-vectorielle et comment modifie-t-il notre compréhension de la circulation des maladies vectorielles ?

IV) Des simulations de circulation et une lutte anti vectorielle adaptée à de tels flux de gènes ?

La dispersion des *Culicoides* au-dessus des masses d'eaux et des terres constitue une voie importante de dispersion des pathogènes transmis. Nos travaux ont montré de faibles niveaux de différenciation génétique entre les populations de *C. obsoletus* à toutes les échelles spatiales étudiées. Ces connaissances sont essentielles pour dans la mise en place de mesures de lutte anti-vectorielle adaptées. Par exemple, des études de génétique des populations ont permis de développer des moyens de lutte anti-vectorielle efficaces contre les mouches responsables de la transmission de trypanosomes (Bouyer *et al*, 2007; Solano *et al*, 2010; Solano *et al*, 2009).

La dispersion importante des *Culicoides* est régulièrement mise en évidence lors d'études de modélisation épidémiologique du virus de la FCO, comme le facteur principal de circulation du virus. (Ducheyne *et al*, 2007; Ducheyne *et al*, 2011) conclut même que lors de l'épizootie de 2006 en France, la dispersion du vecteur via des processus de vols passifs et actifs, aurait été suffisante à elle seule pour expliquer la propagation du virus. La forte connectivité génétique mise en évidence par nos travaux invite à réestimer certains paramètres utilisés dans les études de modélisation de propagation des virus transmis. Le rôle de la dispersion à longue distance des *Culicoides* est souvent sous-estimé ou mal évalué. Des modèles statistiques retraçant la propagation du virus de la FCO dans le nord-ouest de l'Europe ont pourtant pointé du doigt la responsabilité du transport atmosphérique passif des *Culicoides* (Hendrickx *et al*, 2008; Sedda *et al*, 2012). L'une des dernières études de modélisation de transmission du virus de la FCO entre cantons français (Courtejoie *et al*, 2019) n'inclut pas ces différents modes de dispersion. Seuls les événements de vol sur de courtes distances, basés sur les estimations d'études de CMR (à moins de 5 km par semaine), sont pris en compte. Cependant, l'étude conclut à une transmission du virus intra-canton assurée en majeure partie par la dispersion des vecteurs.

Sumner et ses collaborateurs (2017) ont effectué une étude de modélisation similaire en Grande-Bretagne en utilisant une unité épidémiologique différente, celle de l'exploitation agricole. Les auteurs concluaient également que la majeure partie (plus de 90%) de la transmission du virus était due à la dispersion des *Culicoides* entre exploitations (Sumner *et al*, 2017). Courtejoie et al (2019) fait l'hypothèse qu'une partie des transmissions attribuées à la dispersion des *Culicoides* serait en fait liée aux mouvements d'animaux appartenant à la même exploitation entre différents sites de pâture, préconisant la restriction des pratiques de pâture lors d'épizooties (Courtejoie *et al*, 2019). Bien que certains modèles épidémiologiques très complets aient été

développés en incluant l'ensemble des composantes de la dispersion des *Culicoides*, il semble que ces derniers soient, du fait de la multiplication du nombre de paramètres à évaluer, très sensibles au manque de précision de ces derniers (Graesboll *et al*, 2012; Graesboll *et al*, 2014; Graesboll *et al*, 2016). Ainsi, notre étude permet d'apporter davantage d'informations dans l'estimation de la dispersion des *Culicoides*, qui permettront d'alimenter les prochaines études épidémiologiques.

Ceci permettra d'optimiser des outils comme le SWOTS (spatio temporal wind-outbreak trajectory simulation). Le SWOTS est un algorithme développé par Sedda et ses collaborateurs (2012) permettant de quantifier la dispersion de *Culicoides* par le vent et la circulation des virus transmis au-dessus des terres. Celui-ci ne se limite pas comme la plupart des modèles développés récemment aux mouvements au-dessus de la mer (Sedda *et al*, 2012). Seulement, l'estimation des paramètres de dispersion est très longue (environ 7 jours de calculs). Ainsi, les informations que nous apportons par nos travaux sur les flux de gènes des *Culicoides* permettront d'améliorer les performances de tels modèles de dispersion. L'étude de Sedda a également permis d'identifier des « super-diffuseurs », c'est-à-dire des exploitations qui ont infecté de nombreuses autres fermes durant l'épizootie de FCO de 2006. Il serait alors particulièrement intéressant d'effectuer une nouvelle étude de génétique du paysage avec un échantillonnage basé sur ces « super-diffuseurs », pour voir si ce sont les caractéristiques paysagères locales qui ont été à l'origine de ce statut ou les caractéristiques intrinsèques propres aux pratiques agricoles de ces exploitations.

Les moyens de lutte anti-vectorielle

Les moyens actuels de lutte antivectorielle contre les *Culicoides* visent à diminuer l'abondance des populations, par l'utilisation d'insecticides (larvicide et adulticide), de répulsifs ou la destruction des habitats larvaires (Bandeira *et al*, 2017; Carpenter *et al*, 2008b; Harrup *et al*, 2016; Meloni *et al*, 2018; Mullens *et al*, 2015). Dans le cas d'espèces autochtones et très abondantes comme *C. obsoletus*, il n'est pas envisageable d'engager une lutte anti-vectorielle chimique ou environnementale, seules des actions locales, à effets très limités, au niveau de chaque exploitation peuvent être envisagées en cas d'épizootie. De plus, ces mesures, communes à de nombreuses espèces d'insectes vecteurs ou ravageurs manquent de spécificité et d'efficacité, et ont montré leurs limites s'agissant du respect des espèces non cibles et de l'environnement (Carpenter *et al*, 2008b; Harrup *et al*, 2016; Mehlhorn *et al*, 2009; Mullens *et*

al, 2015; Venail, 2014; Venail *et al*, 2012). Elles doivent rester la dernière option de lutte. De plus, comme nos travaux l'ont montré, la dispersion est telle que la recolonisation des zones traitées par des populations adjacentes ou même distantes par le vent serait hautement probable. La prévention des acteurs de la santé animale en amont, la vaccination des animaux et la transmission d'informations sur le rôle important joué par les vecteurs lors des épizooties restent à privilégier (Ganter, 2014).

Perspectives d'études

Ces dernières années, un intérêt croissant a été porté à l'étude des communautés endosymbiotiques héréditaires des vecteurs d'intérêt et à leurs impacts pour les espèces vectrices (Feldhaar and Gross, 2009; Floate *et al*, 2007). Ces endosymbiontes apportent une protection contre les différents stress rencontrés (prédation, parasitisme, stress abiotique) ou un apport en nutriments pouvant compléter certains régimes alimentaires (Dunbar *et al*, 2007; Duron *et al*, 2008; Duron *et al*, 2018; Rio *et al*, 2016).

Cette augmentation de la valeur sélective des hôtes fait des endosymbiontes des acteurs importants de l'histoire évolutive des arthropodes (Ferrari and Vavre, 2011; McLean *et al*, 2016; Sudakaran *et al*, 2017). La transmission s'effectue de manière verticale, de la mère à sa descendance. Afin de maximiser leur propagation, certaines bactéries héréditaires effectuent un parasitisme reproductif, c'est-à-dire un biais de sexe ratio de l'hôte en faveur des femelles (Hurst and Frost, 2015; Stouthamer *et al*, 1999; Werren *et al*, 2008). Cette capacité détournée permet d'utiliser les endosymbiontes en lutte anti-vectorielle (Iturbe-Ormaetxe *et al*, 2011), comme pour *Aedes aegypti* (baisse de la survie des individus infecté) (McMeniman *et al*, 2009) ou *Anopheles stephensi* (limite l'infection par *Plasmodium falciparum*) (Bian *et al*, 2013). Parmi les plus étudiés, l'endosymbiote *Wolbachia*, présent chez près de 66% de toutes les espèces d'insectes (Hilgenboecker *et al*, 2008), peut, dans le cas où le mâle est infecté, empêcher la création de nouvelles générations viables. Ces caractéristiques ont permis son utilisation dans la lutte vectorielle.

Déjà bien étudié chez d'autres espèces de diptères comme les moustiques et la mouche tsé-tsé (Caragata *et al*, 2016; Medlock *et al*, 2013), l'étude des endosymbiontes des *Culicoides* est encore naissante. Des infections par les bactéries endosymbiotiques *Wolbachia* et *Cardinium* ont déjà été mise en évidence chez les *Culicoides* (Covey *et al*, 2020; Ghosh *et al*, 2019; Lewis *et al*, 2014; Morag *et al*, 2012; Pages *et al*, 2017). Récemment, des études ont mis en évidence

la transmission maternelle (présence de la bactérie dans le tissu ovarien) de symbiotes *Rickettsia*, dont le génome présente des caractéristiques potentiellement associées à l'adaptation à l'hôte (Pilgrim *et al*, 2017). De plus la localisation dans les corps adipeux de ces symbiotes peut laisser présager de potentielles interactions avec les virus (Pilgrim *et al*, 2020). À la vue de leurs effets importants sur la biologie et l'écologie des vecteurs, établir l'impact de ces endosymbiontes sur la compétence vectorielle des *Culicoides* est donc une perspective intéressante (Siozios *et al*, 2019).

Caractériser la structure des communautés bactériennes et investiguer celle-ci en comparaison de celle de l'hôte vecteur permettrait aussi d'estimer la connectivité des populations. Si des patrons similaires étaient observés, cela pourrait renforcer nos conclusions concernant la dispersion importante des *Culicoides*. Au contraire, des communautés bactériennes distinctes et spécifiques pourraient confirmer nos résultats de délimitation d'espèces au sein du Complexe *Obsoletus/Scoticus* ou souligner l'importance de l'adaptation à des conditions biotiques et abiotiques locales.

Ainsi, ces organismes associés aux *Culicoides* pourraient être utilisés comme marqueurs de la connectivité des populations. Retracer les changements dans la composition du microbiome serait alors une approche complémentaire aux marqueurs génétiques et permettrait de retracer la dispersion et l'histoire évolutive du vecteur. Cela permettrait également de comprendre le rôle des bactéries symbiotiques sur l'infection et la propagation des virus.

Conclusion générale

Connaître la biologie et l'écologie d'une espèce vectrice est capital pour anticiper et limiter la dispersion de pathogènes. Nos travaux apportent des connaissances nouvelles sur la diversité cryptique de *C. obsoletus*, vecteur majeur des virus de la FCO et de Schmallenberg, en Europe grâce à des approches de taxonomie moléculaire. Nos travaux ont également permis de caractériser la structure génétique, les capacités de dispersion et l'influence de la connectivité paysagère sur la dispersion de *C. obsoletus* en Europe. Ces résultats nous permettent de proposer une histoire évolutive de l'espèce. L'ensemble des données générées au cours de cette étude est inédit et l'échantillonnage réalisé nous a permis de couvrir largement la zone ouest paléarctique par un transect allant de la Scandinavie aux îles Canaries (du nord au sud) et des îles Canaries à la Turquie (d'ouest en est).

Nos travaux permettent de clarifier le schéma de distribution et les relations phylogénétiques des espèces appartenant au Complexe *Obsoletus/Scoticus*. Cinq MOTUs ont été identifiés et validés, *C. obsoletus*, *C. scoticus*, *C. obsoletus* clade O2, *C. obsoletus* clade dark et une MOTU correspondant à une espèce non identifiée. Ces trois dernières espèces n'ont pas été formellement décrites jusqu'ici. Des études de leurs morphologies et de leurs écologies sont nécessaires pour publier des descriptions complètes respectant le Code International de la Nomenclature Zoologique. Le statut taxonomique de *C. montanus*, auparavant considéré comme une espèce valide, est remis en question et nous proposons que celui-ci soit considéré comme un clade phylogénétique bien soutenu au sein de *C. obsoletus* et non plus comme une espèce valide. Serions-nous à un début de spéciation ?

La taxonomie clarifiée du Complexe *Obsoletus/Scoticus* amène à prolonger les questions habituellement posées pour les espèces valides aux nouvelles entités validées. Ont-elles toutes les mêmes capacités vectorielles ? Les mêmes préférences trophiques ? La même co-évolution avec le virus ? Les mêmes communautés endosymbiotiques ?

Alors que jusqu'à présent, la dispersion au-dessus des terres était considérée comme limitée, à cause de la friction paysagère, l'ensemble de ce travail de thèse permet de confirmer que les *Culicoides* possèdent de fortes capacités de dispersion au-dessus de terres. En effet, des flux de gènes intenses ont permis d'homogénéiser la structure génétique des populations à l'échelle d'un pays et d'un continent et de brouiller le signal d'isolement par la distance. Cette dispersion est si importante que deux populations éloignées de milliers de kilomètres peuvent présenter la

même composition génétique. Ceci est possible grâce à la dispersion stratifiée des *Culicoides*. La prochaine étape sera de caractériser la part relative et la fréquence exacte des deux phénomènes, grâce à un échantillonnage à une échelle locale, entre fermes séparées de quelques kilomètres de distance situées dans un paysage contrasté et en réalisant une assignation spécifique de parenté.

Nous avons évalué l'impact d'un large éventail de facteurs environnementaux sur le flux de gènes de *C. obsoletus*. Ces résultats ouvrent la voie à des études complémentaires de génétique du paysage appliquée aux *Culicoides*. La différenciation génétique trop faible de notre jeu de données ne nous a pas permis d'investiguer en profondeur l'impact du vent. Sans doute, il doit être considéré et intégré dans les prochaines études de génétique du paysage car il participe à l'homogénéisation des populations. Actuellement, la prise en compte de la direction du vent dans les analyses de génétique du paysage est limitée, et nous soulignons le besoin d'améliorations méthodologiques sur ce sujet. Nos résultats démontrent pour *C. obsoletus* une capacité de dispersion à l'intérieur des terres très importante, qui doit être prise en compte pour les travaux futurs sur la compétence des vecteurs et la modélisation épidémiologique de la transmission des maladies.

La faible structuration des *Culicoides* est également le témoin d'une histoire évolutive récente. La recolonisation post-glaciaire est un excellent témoin de la dispersion de *C. obsoletus*, puisqu'elle lui a permis dès la fin de la période de glaciation de recoloniser rapidement des niches écologiques vacantes, comme en témoignent l'expansion démographique et la divergence récente des populations ouest européennes. Il semble que *C. obsoletus* a pu persister dans des refuges nord Européens en raison de ses larges préférences trophiques et écologiques. Il serait intéressant dans le futur de s'interroger sur l'histoire évolutive d'autres espèces de *Culicoides* présentant des niches écologiques plus restreintes. De plus, vu la très large la répartition paléarctique et néarctique de *C. obsoletus*, ceci ne représente qu'un morceau du puzzle. Ainsi qu'en est-il de l'histoire évolutive mondiale de *C. obsoletus* ?

Références

- Abbott KC (2011). A dispersal-induced paradox: synchrony and stability in stochastic metapopulations. *Ecol Lett* **14**(11): 1158-1169.
- Abbott RJ, James JK, Milne RI, Gillies AC (2003). Plant introductions, hybridization and gene flow. *Philos Trans R Soc Lond B Biol Sci* **358**(1434): 1123-1132.
- Abellan P, Svenning JC (2014). Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biol J Linn Soc* **113**: 13–28.
- Acevedo P, Ruiz-Fons F, Estrada R, Marquez AL, Miranda MA, Gortazar C *et al* (2010). A broad assessment of factors determining *Culicoides imicola* abundance: modelling the present and forecasting its future in climate change scenarios. *PLoS One* **5**(12): e14236.
- Adamack AT, Gruber B, Dray S (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* **5**(4): 384-387.
- Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulinck H *et al* (2003). The application of ‘least-cost’ modelling as a functional landscape model. *Landscape Urban Plann* **64**(4): 233-247.
- Afonso A, Abrahantes JC, Conraths F, Veldhuis A, Elbers A, Roberts H *et al* (2014). The Schmallenberg virus epidemic in Europe-2011-2013. *Prev Vet Med* **116**(4): 391-403.
- Agapow P-M, Burt A (2001). Indices of multilocus linkage disequilibrium. *Mol Ecol Notes* **1**(1-2): 101-102.
- Agren EC, Burgin L, Lewerin SS, Gloster J, Elvander M (2010). Possible means of introduction of bluetongue virus serotype 8 (BTV-8) to Sweden in August 2008: comparison of results from two models for atmospheric transport of the *Culicoides* vector. *Vet Rec* **167**(13): 484-488.
- Aguilar-Vega C, Fernandez-Carrion E, Sanchez-Vizcaino JM (2019). The possible route of introduction of bluetongue virus serotype 3 into Sicily by windborne transportation of infected *Culicoides* spp. *Transbound Emerg Dis* **66**(4): 1665-1673.
- Alba A, Casal J, Domingo M (2004). Possible introduction of bluetongue into the Balearic Islands, Spain, in 2000, via air streams. *Vet Rec* **155**(15): 460-461.
- Allio R, Donega S, Galtier N, Nabholz B (2017). Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as a Molecular Marker. *Mol Biol Evol* **34**(11): 2762-2772.
- Alvarado - Serrano DF, Hickerson MJ, Freckleton R (2015). Spatially explicit summary statistics for historical population genetic inference. *Methods Ecol Evol* **7**(4): 418-427.

- Amit Roy SR (2014). Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology* **02**(02).
- Ander M, Meiswinkel R, Chirico J (2012). Seasonal dynamics of biting midges (Diptera: Ceratopogonidae: *Culicoides*), the potential vectors of bluetongue virus, in Sweden. *Vet Parasitol* **184**(1): 59-67.
- Ander M, Troell K, Chirico J (2013). Barcoding of biting midges in the genus *Culicoides*: a tool for species determination. *Med Vet Entomol* **27**(3): 323-331.
- Anderson DP, Sturtevant BR (2011). Pattern analysis of eastern spruce budworm *Choristoneura fumiferana* dispersal. *Ecography* **34**(3): 488-497.
- Anderson GS, Belton P, Kleider N (1988). The hypersensitivity of horses to *Culicoides* bites in british columbia. *Can Vet J* **29**(9): 718-723.
- Anderson GS, Belton P, Kleider N (1991). *Culicoides obsoletus* (Diptera: Ceratopogonidae) as a causal agent of *Culicoides* hypersensitivity (sweet itch) in British Columbia. *J Med Entomol* **28**(5): 685-693.
- Andersson E, Bodin Ö (2009). Practical tool for landscape planning? An empirical investigation of network based models of habitat fragmentation. *Ecography* **32**(1): 123-132.
- Andrew Jones DT, Matthew Hort, and Ben Devenish (2007). The U.K. Met Office's Next-Generation Atmospheric Dispersion Model, NAME III.
- Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat Rev Genet* **17**(2): 81-92.
- Angelone S, Kienast F, Holderegger R (2011). Where movement happens: scale-dependent landscape effects on genetic differentiation in the European tree frog. *Ecography* **34**(5): 714-722.
- Arbogast BS (2001). Phylogeography: The History and Formation of Species. *Am Zool* **41**(1): 134-135.
- Archie EA, Luikart G, Ezenwa VO (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends Ecol Evol* **24**(1): 21-30.
- Auger-Rozenberg MA, Roques A (2012). Seed wasp invasions promoted by unregulated seed trade affect vegetal and animal biodiversity. *Integr Zool* **7**(3): 228-246.
- Augot D, Sauvage F, Jouet D, Simphal E, Veuille M, Couloux A *et al* (2010). Discrimination of *Culicoides obsoletus* and *Culicoides scoticus*, potential bluetongue vectors, by morphometrical and mitochondrial cytochrome oxidase subunit I analysis. *Infect Genet Evol* **10**(5): 629-637.
- Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J *et al* (2011). Economic impacts of non-native forest insects in the continental United States. *PLoS One* **6**(9): e24587.

- Auriault M (1979). Contribution à l'étude biologique et écologique de *Culicoides grahamii* (Austen, 1909) (Diptera, Ceratopogonidae) : 4. Dynamique des populations. *Cahiers ORSTOM Série Entomologie Médicale et Parasitologie* **17 (1)**: 31-36.
- Austerlitz F, Dutech C, Smouse PE, Davis F, Sork VL (2007). Estimating anisotropic pollen dispersal: a case study in *Quercus lobata*. *Heredity (Edinb)* **99(2)**: 193-204.
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE *et al* (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics* **18(1)**: 489-522.
- Ayllon T, Nijhof AM, Weiher W, Bauer B, Allene X, Clausen PH (2014). Feeding behaviour of *Culicoides spp.* (Diptera: Ceratopogonidae) on cattle and sheep in northeast Germany. *Parasit Vectors* **7**: 34.
- Backx A, Heutink CG, van Rooij EM, van Rijn PA (2007). Clinical signs of bluetongue virus serotype 8 infection in sheep and goats. *Vet Rec* **161(17)**: 591-592.
- Backx A, Heutink R, van Rooij E, van Rijn P (2009). Transplacental and oral transmission of wild-type bluetongue virus serotype 8 in cattle after experimental infection. *Vet Microbiol* **138(3-4)**: 235-243.
- Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C (2013). Individual dispersal, landscape connectivity and ecological networks. *Biol Rev Camb Philos Soc* **88(2)**: 310-326.
- Baguette M, Van Dyck H (2007). Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology* **22(8)**: 1117-1129.
- Bakhoun MT, Fall AG, Fall M, Bassene CK, Baldet T, Seck MT *et al* (2016). Insight on the larval habitat of Afrotropical *Culicoides* Latreille (Diptera: Ceratopogonidae) in the Niayes area of Senegal, West Africa. *Parasit Vectors* **9**: 462.
- Bakhoun MT, Labuschagne K, Huber K, Fall M, Mathieu B, Venter G *et al* (2018). Phylogenetic relationships and molecular delimitation of *Culicoides* Latreille (Diptera: Ceratopogonidae) species in the Afrotropical region: interest for the subgenus Avaritia. *Syst Entomol* **43(2)**: 355-371.
- Baldet T, Delecolle JC, Cetre-Sossah C, Mathieu B, Meiswinkel R, Gerbier G (2008). Indoor activity of *Culicoides* associated with livestock in the bluetongue virus (BTV) affected region of northern France during autumn 2006. *Prev Vet Med* **87(1-2)**: 84-97.
- Balkenhol N, Fortin MJ (2015). Basics of Study Design: Sampling Landscape Heterogeneity and Genetic Variation for Landscape Genetic Studies. *Landscape Genetics: Concepts, Methods, Applications*: 58-76.
- Balkenhol N, Holbrook JD, Onorato D, Zager P, White C, Waits LP (2014). A multi-method approach for analyzing hierarchical genetic structures: a case study with cougars *Puma concolor*. *Ecography* **37(6)**: 552-563.

- Balkenhol N, Waits LP, Dezzani RJ (2009). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* **32**(5): 818-830.
- Bandeira M, Brito GA, da Penha A, Santos CLC, Rebelo JMM (2017). The influence of environmental management and animal shelters in vector control of *Culicoides* (Diptera, Ceratopogonidae) in northeastern Brazil. *J Vector Ecol* **42**(1): 113-119.
- Barber J, Harrup LE, Silk R, Veronesi E, Gubbins S, Bachanek-Bankowska K *et al* (2018). Blood-feeding, susceptibility to infection with Schmallenberg virus and phylogenetics of *Culicoides* (Diptera: Ceratopogonidae) from the United Kingdom. *Parasit Vectors* **11**(1): 116.
- Barcelo C, Miranda MA (2018). Bionomics of livestock-associated *Culicoides* (biting midge) bluetongue virus vectors under laboratory conditions. *Med Vet Entomol* **32**(2): 216-225.
- Barrios JM, Verstraeten WW, Maes P, Aerts JM, Farifteh J, Coppin P (2012). Using the gravity model to estimate the spatial spread of vector-borne diseases. *Int J Environ Res Public Health* **9**(12): 4346-4364.
- Bartley D, Bagley M, Gall G, Bentley B (1992). Use of Linkage Disequilibrium Data to Estimate Effective Size of Hatchery and Natural Fish Populations. *Conserv Biol* **6**(3): 365-375.
- Bartsch S, Bauer B, Wiemann A, Clausen PH, Steuber S (2009). Feeding patterns of biting midges of the *Culicoides obsoletus* and *Culicoides pulicaris* groups on selected farms in Brandenburg, Germany. *Parasitol Res* **105**(2): 373-380.
- Baylis M (2017). Potential impact of climate change on emerging vector-borne and other infections in the UK. *Environ Health* **16**(Suppl 1): 112.
- Baylis M, Parkin H, Kreppel K, Carpenter S, Mellor PS, McIntyre KM (2010). Evaluation of housing as a means to protect cattle from *Culicoides* biting midges, the vectors of bluetongue virus. *Med Vet Entomol* **24**(1): 38-45.
- Beaumont MA (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annu Rev Ecol Evol Syst* **41**(1): 379-406.
- Beaumont MA, Rannala B (2004). The Bayesian revolution in genetics. *Nat Rev Genet* **5**(4): 251-261.
- Beaumont MA, Zhang W, Balding DJ (2002). Approximate Bayesian computation in population genetics. *Genetics* **162**(4): 2025-2035.
- Beer M, Conraths FJ, van der Poel WH (2013). 'Schmallenberg virus'--a novel orthobunyavirus emerging in Europe. *Epidemiol Infect* **141**(1): 1-8.

- Beerli P, Felsenstein J (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc Natl Acad Sci U S A* **98**(8): 4563-4568.
- Belkhir K (2004). Genetix 4.05, logiciel sous Windows TM pour la genetique des populations. <http://www.genetixuniv-montp2fr/genetix/genetixhtm>.
- Bellis GA, Gopurenko D, Cookson B, Postle AC, Halling L, Harris N *et al* (2015). Identification of incursions of *Culicoides* Latreille species (Diptera: Ceratopogonidae) in Australasia using morphological techniques and DNA barcoding. *Austral Entomology* **54**(3): 332-338.
- Benton TG, Bowler DE (2012). Linking dispersal to spatial dynamics *Dispersal Ecology and Evolution*, pp 251-265.
- Bertin A, Alvarez E, Gouin N, Gianoli E, Montecinos S, Lek S *et al* (2015). Effects of wind-driven spatial structure and environmental heterogeneity on high-altitude wetland macroinvertebrate assemblages with contrasting dispersal modes. *Freshwater Biology* **60**(2): 297-310.
- Bertorelle G, Benazzo A, Mona S (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Mol Ecol* **19**(13): 2609-2625.
- Bhagwat SA, Willis KJ (2008). Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *J Biogeogr* **35**(3): 464-482.
- Bian G, Joshi D, Dong Y, Lu P, Zhou G, Pan X *et al* (2013). *Wolbachia* invades *Anopheles stephensi* populations and induces refractoriness to *Plasmodium* infection. *Science* **340**(6133): 748-751.
- Biek R (2007). Evolutionary dynamics and spatial genetic structure of epizootic hemorrhagic disease virus in the eastern United States. *Infect Genet Evol* **7**(5): 651-655.
- Biek R, Real LA (2010). The landscape genetics of infectious disease emergence and spread. *Mol Ecol* **19**(17): 3515-3531.
- Bishop AL, Barchia IM, Spohr LJ (2000). Models for the dispersal in Australia of the arbovirus vector, *Culicoides brevitarsis* Kieffer (Diptera: Ceratopogonidae). *Prev Vet Med* **47**(4): 243-254.
- Bjornstad ON, Ims RA, Lambin X (1999). Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends Ecol Evol* **14**(11): 427-432.
- Blackwell A (2001). Recent advances on the ecology and behaviour of *Culicoides spp.* in Scotland and the prospects for control. *Vet Bull* **11**: 1-8.
- Blackwell A, Mordue AJ, Mordue W (1994). Identification of bloodmeals of the Scottish biting midge, *Culicoides impunctatus*, by indirect enzyme-linked immunosorbent assay (ELISA). *Med Vet Entomol* **8**(1): 20-24.

- Bock DG, Caseys C, Cousens RD, Hahn MA, Heredia SM, Hubner S *et al* (2015). What we still don't know about invasion genetics. *Mol Ecol* **24**(9): 2277-2297.
- Bohonak AJ, Jenkins DG (2003). Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecol Lett* **6**(8): 783-796.
- Boissin E, Micu D, Janczyszyn-Le Goff M, Neglia V, Bat L, Todorova V *et al* (2016). Contemporary genetic structure and postglacial demographic history of the black scorpionfish, *Scorpaena porcus*, in the Mediterranean and the Black Seas. *Mol Ecol* **25**(10): 2195-2209.
- Borkent A (2014). The pupae of the biting midges of the world (Diptera: Ceratopogonidae), with a generic key and analysis of the phylogenetic relationships between genera. *Zootaxa* **3879**(1): 1-327.
- Borkent A (2016). World Species of Biting Midges (Diptera: Ceratopogonidae).
- Bortolus A (2008). Error Cascades in the Biological Sciences: The Unwanted Consequences of Using Bad Taxonomy in Ecology. *AMBIO: A Journal of the Human Environment* **37**(2): 114-118.
- Bourquia M, Garros C, Rakotoarivony I, Gardes L, Huber K, Boukhari I *et al* (2019). Update of the species checklist of *Culicoides* Latreille, 1809 biting midges (Diptera: Ceratopogonidae) of Morocco. *Parasit Vectors* **12**(1): 459.
- Bouyer J, Dicko AH, Cecchi G, Ravel S, Guerrini L, Solano P *et al* (2015). Mapping landscape friction to locate isolated tsetse populations that are candidates for elimination. *Proc Natl Acad Sci U S A* **112**(47): 14575-14580.
- Bouyer J, Ravel S, Dujardin JP, de Meeus T, Vial L, Thevenon S *et al* (2007). Population structuring of *Glossina palpalis gambiensis* (Diptera: Glossinidae) according to landscape fragmentation in the Mouhoun river, Burkina Faso. *J Med Entomol* **44**(5): 788-795.
- Bowden JJ, Buddle CM (2015). Determinants of ground-dwelling spider assemblages at a regional scale in the Yukon Territory, Canada. *Ecoscience* **17**(3): 287-297.
- Bowler DE, Benton TG (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biol Rev Camb Philos Soc* **80**(2): 205-225.
- Brar S, Tsui CK, Dhillon B, Bergeron MJ, Joly DL, Zambino PJ *et al* (2015). Colonization history, host distribution, anthropogenic influence and landscape features shape populations of white pine blister rust, an invasive alien tree pathogen. *PLoS One* **10**(5): e0127916.
- Braverman Y, Chechik F (1996). Air streams and the introduction of animal diseases borne on *Culicoides* (Diptera, Ceratopogonidae) into Israel. *Rev Sci Tech* **15**(3): 1037-1052.

- Brenner RJ, Wargo MJ, Stains GS, Mulla MS (1984). The Dispersal of *Culicoides mohave* (Diptera, Ceratopogonidae) in the Desert of Southern-California. *Mosquito News* **44**(3): 343-350.
- Brito PH, Edwards SV (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135**(3): 439-455.
- Brookfield JF (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* **5**(3): 453-455.
- Broquet T, Petit EJ (2009). Molecular Estimation of Dispersal for Ecology and Population Genetics. *Annu Rev Ecol Evol Syst* **40**(1): 193-216.
- Bryant JE, Holmes EC, Barrett AD (2007). Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. *PLoS Pathog* **3**(5): e75.
- Bumbarov V, Golender N, Jenckel M, Wernike K, Beer M, Khinich E *et al* (2020). Characterization of bluetongue virus serotype 28. *Transbound Emerg Dis* **67**(1): 171-182.
- Burgess SC, Baskett ML, Grosberg RK, Morgan SG, Strathmann RR (2015). When is dispersal for dispersal? Unifying marine and terrestrial perspectives. *Biological Reviews* **91**(3).
- Burgin L, Ekstrom M, Dessai S (2017). Combining dispersion modelling with synoptic patterns to understand the wind-borne transport into the UK of the bluetongue disease vector. *Int J Biometeorol* **61**(7): 1233-1245.
- Burgin LE, Gloster J, Sanders C, Mellor PS, Gubbins S, Carpenter S (2013). Investigating incursions of bluetongue virus using a model of long-distance *Culicoides* biting midge dispersal. *Transbound Emerg Dis* **60**(3): 263-272.
- Bush GL, Butlin RK (2004). Sympatric Speciation in Insects. *Adaptive Speciation*: 229-248.
- Calvo JH, Calvete C, Martinez-Royo A, Estrada R, Miranda MA, Borrás D *et al* (2009). Variations in the mitochondrial cytochrome c oxidase subunit I gene indicate northward expanding populations of *Culicoides imicola* in Spain. *Bull Entomol Res* **99**(6): 583-591.
- Campbell LP, Alexander AM (2017). Landscape Genetics of *Aedes mcintoshi* (Diptera: Culicidae), an Important Vector of Rift Valley Fever Virus in Northeastern Kenya. *J Med Entomol* **54**(5): 1258-1265.
- Capela R, Purse BV, Pena I, Wittman EJ, Margarita Y, Capela M *et al* (2003). Spatial distribution of *Culicoides* species in Portugal in relation to the transmission of African horse sickness and bluetongue viruses. *Med Vet Entomol* **17**(2): 165-177.
- Carabajal CC, Harding DJ, Boy JP, Danielson JJ, Gesch DB, Suchdeo VP (2011). Evaluation of the Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010) Using ICESat Geodetic Control. *International Symposium on Lidar and Radar Mapping 2011: Technologies and Applications* **8286**.

- Caragata EP, Dutra HLC, Moreira LA (2016). Exploiting Intimate Relationships: Controlling Mosquito-Transmitted Disease with *Wolbachia*. *Trends Parasitol* **32**(3): 207-218.
- Carpenter S (2016). Progress and knowledge gaps in *Culicoides* genetics, genomics and population modelling: 2003 to 2014. *Vet Ital* **52**(3-4): 201-211.
- Carpenter S, Groschup MH, Garros C, Felipe-Bauer ML, Purse BV (2013). *Culicoides* biting midges, arboviruses and public health in Europe. *Antiviral Res* **100**(1): 102-113.
- Carpenter S, Lunt HL, Arav D, Venter GJ, Mellor PS (2006). Oral susceptibility to bluetongue virus of *Culicoides* (Diptera: Ceratopogonidae) from the United Kingdom. *J Med Entomol* **43**(1): 73-78.
- Carpenter S, McArthur C, Selby R, Ward R, Nolan DV, Luntz AJ *et al* (2008a). Experimental infection studies of UK *Culicoides* species midges with bluetongue virus serotypes 8 and 9. *Vet Rec* **163**(20): 589-592.
- Carpenter S, Mellor PS, Torr SJ (2008b). Control techniques for *Culicoides* biting midges and their application in the U.K. and northwestern Palearctic. *Med Vet Entomol* **22**(3): 175-187.
- Carpenter S, Szymaragd C, Barber J, Labuschagne K, Gubbins S, Mellor P (2008c). An assessment of *Culicoides* surveillance techniques in northern Europe: have we underestimated a potential bluetongue virus vector? *J Appl Ecol* **45**(4).
- Carpenter S, Wilson A, Mellor PS (2009). *Culicoides* and the emergence of bluetongue virus in northern Europe. *Trends Microbiol* **17**(4): 172-178.
- Carrel MA, Emch M, Nguyen T, Todd Jobe R, Wan XF (2012). Population-environment drivers of H5N1 avian influenza molecular change in Vietnam. *Health Place* **18**(5): 1122-1131.
- Carstens BC, Pelletier TA, Reid NM, Satler JD (2013). How to fail at species delimitation. *Mol Ecol* **22**(17): 4369-4383.
- Cayuela H, Rougemont Q, Prunier JG, Moore JS, Clobert J, Besnard A *et al* (2018). Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Mol Ecol* **27**(20): 3976-4010.
- Cetre-Sossah C, Baldet T, Delecolle JC, Mathieu B, Perrin A, Grillet C *et al* (2004). Molecular detection of *Culicoides spp.* and *Culicoides imicola*, the principal vector of bluetongue (BT) and African horse sickness (AHS) in Africa and Europe. *Vet Res* **35**(3): 325-337.
- Chaine AS, Legendre S, Clobert J (2013). The co-evolution of multiply-informed dispersal: information transfer across landscapes from neighbors and immigrants. *PeerJ* **1**: e44.
- Chakraborty R, Zhong Y, Jin L, Budowle B (1994). Nondetectability of restriction fragments and independence of DNA fragment sizes within and between loci in RFLP typing of DNA. *Am J Hum Genet* **55**(2): 391-401.

- Chang X, Zhong D, Lo E, Fang Q, Bonizzoni M, Wang X *et al* (2016). Landscape genetic structure and evolutionary genetics of insecticide resistance gene mutations in *Anopheles sinensis*. *Parasit Vectors* **9**(1): 228.
- Chapman JW, Reynolds DR, Wilson K (2015). Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecol Lett* **18**(3): 287-302.
- Chauhan HC, Biswas SK, Chand K, Rehman W, Das B, Dadawala AI *et al* (2014). Isolation of bluetongue virus serotype 1 from aborted goat fetuses. *Rev Sci Tech* **33**(3): 803-812.
- Chen C, Durand E, Forbes F, François O (2007). Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* **7**(5): 747-756.
- Cho HC, Chong CS (1974). Notes On Biting Midges Of The Genus *Culicoides* From South Korea: With Special Reference To Unrecorded Species And Distribution. *Kisaengchunghak Chapchi* **12**(1): 45-75.
- Ciosi M, Miller NF, Kim KS, Giordano R, Estoup A, Guillemaud T (2008). Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. (1365-294X (Electronic)).
- Clarke RT, Rothery P, Raybould AF (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *J Agric Biol Environ Stat* **7**(3): 361-372.
- Clobert J, Ims RA, Rousset F (2004). Causes, Mechanisms and Consequences of Dispersal *Ecology, Genetics and Evolution of Metapopulations*, pp 307-335.
- Clobert J, Le Galliard JF, Cote J, Meylan S, Massot M (2009). Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecol Lett* **12**(3): 197-209.
- Comes HP, Kadereit JW (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends Plant Sci* **3**(11): 432-438.
- Corander J, Sirén J, Arjas E (2007). Bayesian spatial modeling of genetic population structure. *Computational Statistics* **23**(1): 111-129.
- Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004). BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20**(15): 2363-2369.
- Corander J, Waldmann PF, Sillanpää MJ (2003). Bayesian analysis of genetic differentiation between populations. *Genetics* **163**(1): 367-374.
- Corman VM, Ithete NL, Richards LR, Schoeman MC, Preiser W, Drosten C *et al* (2014). Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J Virol* **88**(19): 11297-11303.

- Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R *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**(8): 1187-1189.
- Cote H, Garant D, Robert K, Mainguy J, Pelletier F (2012). Genetic structure and rabies spread potential in raccoons: the role of landscape barriers and sex-biased dispersal. *Evol Appl* **5**(4): 393-404.
- Cote J, Bocedi G, Debeffe L, Chudzinska ME, Weigang HC, Dytham C *et al* (2017). Behavioural synchronization of large-scale animal movements - disperse alone, but migrate together? *Biol Rev Camb Philos Soc* **92**(3): 1275-1296.
- Courtejoie N, Cauchemez S, Zanella G, Durand B (2019). A network-based approach to modelling bluetongue spread in France. *Prev Vet Med* **170**: 104744.
- Covey H, Hall RH, Krafur A, Matthews ML, Shults PT, Brelsfoard CL (2020). Cryptic *Wolbachia* (Rickettsiales: Rickettsiaceae) Detection and Prevalence in *Culicoides* (Diptera: Ceratopogonidae) Midge Populations in the United States. *J Med Entomol* **57**(4): 1262-1269.
- Csillery K, Blum MG, Gaggiotti OE, Francois O (2010). Approximate Bayesian Computation (ABC) in practice. *Trends Ecol Evol* **25**(7): 410-418.
- Cuellar AC, Kjaer LJ, Baum A, Stockmarr A, Skovgard H, Nielsen SA *et al* (2020). Modelling the monthly abundance of *Culicoides* biting midges in nine European countries using Random Forests machine learning. *Parasit Vectors* **13**(1): 194.
- Cuellar AC, Kjaer LJ, Kirkeby C, Skovgard H, Nielsen SA, Stockmarr A *et al* (2018). Spatial and temporal variation in the abundance of *Culicoides* biting midges (Diptera: Ceratopogonidae) in nine European countries. *Parasit Vectors* **11**(1): 112.
- Cullingham CI, Kyle CJ, Pond BA, Rees EE, White BN (2009). Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Mol Ecol* **18**(1): 43-53.
- Dallas JF, Cruickshank RH, Linton YM, Nolan DV, Patakakis M, Braverman Y *et al* (2003). Phylogenetic status and matrilineal structure of the biting midge, *Culicoides imicola*, in Portugal, Rhodes and Israel. *Med Vet Entomol* **17**(4): 379-387.
- Darpe K, Kf L, Venter G, L H, D K, Potgieter C *et al* (2009). *Investigation of saliva proteins from different Culicoides species and their interaction with the BTV particle.*
- de Meeus T, McCoy KD, Prugnolle F, Chevillon C, Durand P, Hurtrez-Bousses S *et al* (2007). Population genetics and molecular epidemiology or how to "debusquer la bete". *Infect Genet Evol* **7**(2): 308-332.
- De Queiroz K (2007). Species concepts and species delimitation. *Syst Biol* **56**(6): 879-886.

- De Regge N, Deblauwe I, De Deken R, Vantieghem P, Madder M, Geysen D *et al* (2012). Detection of Schmallenberg virus in different *Culicoides spp.* by real-time RT-PCR. *Transbound Emerg Dis* **59**(6): 471-475.
- Debeila TJ (2010). Characterisation of Selected *Culicoides* Population in South Africa Using Genetic Markers. University of Pretoria, Faculty of Veterinary Science.
- Deblauwe I, de Witte JC, de Deken G, de Deken R, Madder M, van Erk S *et al* (2012). A new tool for the molecular identification of *Culicoides* species of the Obsoletus group: the glass slide microarray approach. *Med Vet Entomol* **26**(1): 83-91.
- DeChaine EG (2008). A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants. *Plant Ecol Divers* **1**(2): 197-207.
- Delecolle JC (1985). Nouvelle contribution à l'étude systématique et iconographique des espèces du genre *Culicoides*, (Diptera) : (Ceratopogonidae) du Nord-Est de la France. Strasbourg.
- Dellicour S, Flot JF (2015). Delimiting Species-Poor Data Sets using Single Molecular Markers: A Study of Barcode Gaps, Haplowebs and GMYC. *Syst Biol* **64**(6): 900-908.
- Dellicour S, Flot JF (2018). The hitchhiker's guide to single-locus species delimitation. *Mol Ecol Resour* **18**(6): 1234-1246.
- Dellicour S, Mardulyn P (2014). spads 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. *Mol Ecol Resour* **14**(3): 647-651.
- Dellicour S, Prunier JG, Piry S, Eloy MC, Bertouille S, Licoppe A *et al* (2019). Landscape genetic analyses of *Cervus elaphus* and *Sus scrofa*: comparative study and analytical developments. *Heredity (Edinb)* **123**(2): 228-241.
- Dellicour S, Rose R, Pybus OG (2016). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics* **17**: 82.
- Despres L (2019). One, two or more species? Mitonuclear discordance and species delimitation. *Mol Ecol* **28**(17): 3845-3847.
- Deyoung RW, Zamorano A, Mesenbrink BT, Campbell TA, Leland BR, Moore GM *et al* (2009). Landscape-Genetic Analysis of Population Structure in the Texas Gray Fox Oral Rabies Vaccination Zone. *J Wildl Manage* **73**(8): 1292-1299.
- Diarra M, Fall M, Fall AG, Diop A, Seck MT, Garros C *et al* (2014). Seasonal dynamics of *Culicoides* (Diptera: Ceratopogonidae) biting midges, potential vectors of African horse sickness and bluetongue viruses in the Niayes area of Senegal. *Parasit Vectors* **7**: 147.
- Diarra M, Fall M, Lancelot R, Diop A, Fall AG, Dicko A *et al* (2015). Modelling the Abundances of Two Major *Culicoides* (Diptera: Ceratopogonidae) Species in the Niayes Area of Senegal. *PLoS One* **10**(6): e0131021.

- Diaz-Sanchez S, Hernandez-Jarguin A, Torina A, Fernandez de Mera IG, Estrada-Pena A, Villar M *et al* (2018). Biotic and abiotic factors shape the microbiota of wild-caught populations of the arbovirus vector *Culicoides imicola*. *Insect Mol Biol* **27**(6): 847-861.
- Dijkstra E, van der Ven IJ, Meiswinkel R, Holzel DR, Van Rijn PA, Meiswinkel R (2008). *Culicoides chiopterus* as a potential vector of bluetongue virus in Europe. *Vet Rec* **162**(13): 422.
- Diniz-Filho JA, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, de Campos Telles MP *et al* (2013). Mantel test in population genetics. *Genet Mol Biol* **36**(4): 475-485.
- Dlugosch KM, Parker IM (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* **17**(1): 431-449.
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* **14**(1): 209-214.
- Drake V, Reynolds D (2012). *Radar Entomology: Observing Insect Flight and Migration*.
- Dube CE, Planes S, Zhou Y, Berteaux-Lecellier V, Boissin E (2017). Genetic diversity and differentiation in reef-building *Millepora* species, as revealed by cross-species amplification of fifteen novel microsatellite loci. *PeerJ* **5**: e2936.
- Dubovi EJ, Hawkins M, Griffin RA, Jr., Johnson DJ, Ostlund EN (2013). Isolation of Bluetongue virus from canine abortions. *J Vet Diagn Invest* **25**(4): 490-492.
- Duchesne P, Turgeon J (2012). FLOCK provides reliable solutions to the "number of populations" problem. *J Hered* **103**(5): 734-743.
- Ducheyne E, De Deken R, Becu S, Codina B, Nomikou K, Mangana-Vougiaki O *et al* (2007). Quantifying the wind dispersal of *Culicoides* species in Greece and Bulgaria. *Geospat Health* **1**(2): 177-189.
- Ducheyne E, Lange M, Van der Stede Y, Meroc E, Durand B, Hendrickx G (2011). A stochastic predictive model for the natural spread of bluetongue. *Prev Vet Med* **99**(1): 48-59.
- Dunbar HE, Wilson AC, Ferguson NR, Moran NA (2007). Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol* **5**(5): e96.
- Duron O, Bouchon D, Boutin S, Bellamy L, Zhou L, Engelstadter J *et al* (2008). The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol* **6**(1): 27.
- Duron O, Morel O, Noel V, Buysse M, Binetruy F, Lancelot R *et al* (2018). Tick-Bacteria Mutualism Depends on B Vitamin Synthesis Pathways. *Curr Biol* **28**(12): 1896-1902 e1895.

- Durr PA, Graham K, van Klinken RD (2017). Sellers' Revisited: A Big Data Reassessment of Historical Outbreaks of Bluetongue and African Horse Sickness due to the Long-Distance Wind Dispersion of *Culicoides* Midges. *Front Vet Sci* **4**: 98.
- Dutech C, Sork VL, Irwin AJ, Smouse PE, Davis FW (2005). Gene flow and fine-scale genetic structure in a wind-pollinated tree species, *Quercus lobata* (Fagaceae). *Am J Bot* **92**(2): 252-261.
- Duvallet G, Fontenille D, Robert V (2017). *Entomologie médicale et vétérinaire*.
- Eagles D, Deveson T, Walker PJ, Zalucki MP, Durr P (2012). Evaluation of long-distance dispersal of *Culicoides* midges into northern Australia using a migration model. *Med Vet Entomol* **26**(3): 334-340.
- Eagles D, Melville L, Weir R, Davis S, Bellis G, Zalucki MP *et al* (2014). Long-distance aerial dispersal modelling of *Culicoides* biting midges: case studies of incursions into Australia. *BMC Vet Res* **10**: 135.
- Earl DA, vonHoldt BM (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**(2): 359-361.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**(5): 1792-1797.
- Edwards S, Bensch S (2009). Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Mol Ecol* **18**(14): 2930-2933; discussion 2934-2936.
- Elbers AR, Backx A, Mintiens K, Gerbier G, Staubach C, Hendrickx G *et al* (2008). Field observations during the Bluetongue serotype 8 epidemic in 2006. II. Morbidity and mortality rate, case fatality and clinical recovery in sheep and cattle in the Netherlands. *Prev Vet Med* **87**(1-2): 31-40.
- Elbers AR, Meiswinkel R (2014). *Culicoides* (Diptera: Ceratopogonidae) host preferences and biting rates in the Netherlands: comparing cattle, sheep and the black-light suction trap. *Vet Parasitol* **205**(1-2): 330-337.
- Ellegren H (2014). Genome sequencing and population genomics in non-model organisms. *Trends Ecol Evol* **29**(1): 51-63.
- Ellis JS, Knight ME, Darvill B, Goulson D (2006). Extremely low effective population sizes, genetic structuring and reduced genetic diversity in a threatened bumblebee species, *Bombus sylvarum* (Hymenoptera: Apidae). *Mol Ecol* **15**(14): 4375-4386.
- Emaresi G, Pellet J, Dubey S, Hirzel AH, Fumagalli L (2009). Landscape genetics of the Alpine newt (*Mesotriton alpestris*) inferred from a strip-based approach. *Conserv Genet* **12**(1): 41-50.

- Ence DD, Carstens BC (2011). SpedeSTEM: a rapid and accurate method for species delimitation. *Mol Ecol Resour* **11**(3): 473-480.
- Engler JO, Balkenhol N, Filz KJ, Habel JC, Rodder D (2014). Comparative landscape genetics of three closely related sympatric Hesperid butterflies with diverging ecological traits. *PLoS One* **9**(9): e106526.
- Epopa PS, Millogo AA, Collins CM, North A, Tripet F, Benedict MQ *et al* (2017). The use of sequential mark-release-recapture experiments to estimate population size, survival and dispersal of male mosquitoes of the *Anopheles gambiae* complex in Bana, a west African humid savannah village. *Parasit Vectors* **10**(1): 376.
- Esselstyn JA, Evans BJ, Sedlock JL, Anwarali Khan FA, Heaney LR (2012). Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc Biol Sci* **279**(1743): 3678-3686.
- Estoup A, Guillemaud T (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Mol Ecol* **19**(19): 4113-4130.
- Estoup A, Solignac M, Cornuet JM, Goudet J, Scholl A (1996). Genetic differentiation of continental and island populations of *Bombus terrestris* (Hymenoptera: Apidae) in Europe. *Mol Ecol* **5**(1): 19-31.
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**(8): 2611-2620.
- Excoffier L, Foll M, Petit RJ (2009). Genetic Consequences of Range Expansions. *Annu Rev Ecol Evol Syst* **40**(1): 481-501.
- Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **491**(1985): 479-491.
- Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**(4): 1567-1587.
- Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ *et al* (2014). HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **346**(6205): 56-61.
- Feldhaar H, Gross R (2009). Insects as hosts for mutualistic bacteria. *Int J Med Microbiol* **299**(1): 1-8.
- Fernández - López J, Schliep K (2018). rWind: download, edit and include wind data in ecological and evolutionary analysis. *Ecography* **42**(4): 804-810.
- Ferrari J, Vavre F (2011). Bacterial symbionts in insects or the story of communities affecting communities. *Philos Trans R Soc Lond B Biol Sci* **366**(1569): 1389-1400.

- Ficetola GF, Garner TW, De Bernardi F (2007). Genetic diversity, but not hatching success, is jointly affected by postglacial colonization and isolation in the threatened frog, *Rana latastei*. *Mol Ecol* **16**(9): 1787-1797.
- Floate KD, Kyei-Poku GK, Coghlin PC (2007). Overview and relevance of *Wolbachia* bacteria in biocontrol research. *Biocontrol Sci Technol* **16**(8): 767-788.
- Flot JF (2015). Species Delimitation's Coming of Age. *Syst Biol* **64**(6): 897-899.
- Foll M, Gaggiotti O (2006). Identifying the environmental factors that determine the genetic structure of populations. *Genetics* **174**(2): 875-891.
- Fontaneto D, Flot JF, Tang C (2015). Guidelines for DNA taxonomy, with a focus on the meiofauna. *Mar Biodivers*: 1-19.
- Fountain-Jones NM, Packer C, Troyer JL, VanderWaal K, Robinson S, Jacquot M *et al* (2017). Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. *J Anim Ecol* **86**(6): 1469-1482.
- Fountain-Jones NM, Pearse WD, Escobar LE, Alba-Casals A, Carver S, Davies TJ *et al* (2018). Towards an eco-phylogenetic framework for infectious disease ecology. *Biol Rev Camb Philos Soc* **93**(2): 950-970.
- Francois O, Ancelet S, Guillot G (2006). Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* **174**(2): 805-816.
- Francois O, Durand E (2010). Spatially explicit Bayesian clustering models in population genetics. *Mol Ecol Resour* **10**(5): 773-784.
- François O, Waits LP (2015). Clustering and Assignment Methods in Landscape Genetics *Landscape Genetics*, pp 114-128.
- Frank BM, Piccolo JJ, Baret PV (2011). A review of ecological models for brown trout: towards a new demogenetic model. *Ecology of Freshwater Fish* **20**(2): 167-198.
- Frankham R (2009). Where are we in conservation genetics and where do we need to go? *Conserv Genet* **11**(2): 661-663.
- Frantz AC, Bertouille S, Eloy MC, Licoppe A, Chaumont F, Flamand MC (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (*Cervus elaphus*), but not wild boars (*Sus scrofa*). *Mol Ecol* **21**(14): 3445-3457.
- Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *J Appl Ecol* **46**(2): 493-505.
- Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**(2): 915-925.

- Fuentes-Pardo AP, Ruzzante DE (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Mol Ecol* **26**(20): 5369-5406.
- Fujisawa T, Barraclough TG (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst Biol* **62**(5): 707-724.
- Fujita MK, Leache AD, Burbrink FT, McGuire JA, Moritz C (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol Evol* **27**(9): 480-488.
- Funk DJ, Omland KE (2003). Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annu Rev Ecol Evol Syst* **34**(1): 397-423.
- Gaggiotti OE, Smouse PE (1996). Stochastic Migration and Maintenance of Genetic Variation in Sink Populations. *The American Naturalist* **147**(6): 919-945.
- Ganter M (2014). Bluetongue disease—Global overview and future risks. *Small Rumin Res* **118**(1-3): 79-85.
- Gao X, Wang L, Liu T, Xiao J, Wang H (2020). Effect of agro-ecological landscape on the distribution of *Culicoides obsoletus* in northeast China. *Pest Manag Sci*.
- Garcia-Diaz P (2016). Invasion Ecology, 2nd edition by J. L.Lockwood, M. F.Hoopes and M. P.Marchetti. Wiley-Blackwell, West Sussex, 2013. xi + 444 pp. Price AUD\$94.95 (paperback, also available in hardcover and e-book). ISBN 978-1-4443-3364-0. *Austral Ecol* **41**(4): e2-e2.
- Garros C, Balenghien T (2017). Chapitre 14. Les *Culicoides* (Diptera : Ceratopogonidae). In: Éditions I (ed) *Entomologie médicale et vétérinaire*, pp 345-365.
- Garros C, Balenghien T, Carpenter S, Delecolle JC, Meiswinkel R, Pédarrieu A *et al* (2014). Towards the PCR-based identification of Palaearctic *Culicoides* biting midges (Diptera: Ceratopogonidae): results from an international ring trial targeting four species of the subgenus *Avaritia*. *Parasit Vectors* **7**: 223.
- Garros C, Gardes L, Allene X, Rakotoarivony I, Viennet E, Rossi S *et al* (2011). Adaptation of a species-specific multiplex PCR assay for the identification of blood meal source in *Culicoides* (Ceratopogonidae: Diptera): applications on Palaearctic biting midge species, vectors of Orbiviruses. *Infect Genet Evol* **11**(5): 1103-1110.
- Garza JC, Williamson EG (2001). Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* **10**(2): 305-318.
- Gerber AS, Loggins R, Kumar S, Dowling TE (2001). Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? *Annu Rev Genet* **35**: 539-566.

- Gerry AC, Sarto i Monteys V, Moreno Vidal JO, Francino O, Mullens BA (2009). Biting rates of *Culicoides* midges (Diptera: Ceratopogonidae) on sheep in northeastern Spain in relation to midge capture using UV light and carbon dioxide-baited traps. *J Med Entomol* **46**(3): 615-624.
- Gethmann J, Probst C, Conraths FJ (2020). Economic Impact of a Bluetongue Serotype 8 Epidemic in Germany. *Front Vet Sci* **7**: 65.
- Ghosh A, Jaspersen D, Cohnstaedt LW, Brelsfoard CL (2019). Transfection of *Culicoides sonorensis* biting midge cell lines with *Wolbachia pipientis*. *Parasit Vectors* **12**(1): 483.
- Gilbert M, Nicolas G, Cinardi G, Van Boeckel TP, Vanwambeke S, Wint WGR *et al.* (2018). Harvard Dataverse.
- Gillespie RG, Baldwin BG, Waters JM, Fraser CI, Nikula R, Roderick GK (2012). Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol Evol* **27**(1): 47-56.
- Ginestet C (2011). ggplot2: Elegant Graphics for Data Analysis. *J Roy Stat Soc Ser A (Stat Soc)* **174**(1): 245-246.
- Gippet JM, Liebhold AM, Fenn-Moltu G, Bertelsmeier C (2019). Human-mediated dispersal in insects. *Curr Opin Insect Sci* **35**: 96-102.
- Girod C, Vitalis R, Leblois R, Freville H (2011). Inferring population decline and expansion from microsatellite data: a simulation-based evaluation of the Msvar method. *Genetics* **188**(1): 165-179.
- Gloster J, Burgin L, Witham C, Athanassiadou M, Mellor PS (2008). Bluetongue in the United Kingdom and northern Europe in 2007 and key issues for 2008. *Vet Rec* **162**(10): 298-302.
- Goetghebuer M (1952). Le genre *Culicoides* (Diptères: Ceratopogonides) et ses représentants en Belgique. *Biologische Jaarboek* **19**: 185-191.
- Goffredo M, Delecolle JC, Semproni G, Lika A (2004). *Culicoides* (Diptera: Ceratopogonidae) in Albania: results of the 2002 entomological survey for bluetongue. *Vet Ital* **40**(3): 209-211.
- Goffredo M, Meiswinkel R, Federici V, Di Nicola F, Mancini G, Ippoliti C *et al* (2016). The '*Culicoides obsoletus* group' in Italy: relative abundance, geographic range, and role as vector for Bluetongue virus. *Vet Ital* **52**(3-4): 235-241.
- Gómez A, Lunt DH (2007). Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Springer (ed) *Phylogeography of Southern European Refugia*
- Gomulski LM, Meiswinkel R, Delécolle J-C, Goffredo M, Gasperi G (2005). Phylogenetic relationships of the subgenus *Avaritia*^[1]_[SEP]Fox, 1955 including *Culicoides obsoletus* (Diptera, Ceratopogonidae) in Italy based on internal transcribed spacer 2 ribosomal DNA sequences. *Syst Entomol* **30**(4): 619-631.

- Gomulski LM, Meiswinkel R, Delecolle JC, Goffredo M, Gasperi G (2006). Phylogeny of the subgenus *Culicoides* and related species in Italy, inferred from internal transcribed spacer 2 ribosomal DNA sequences. *Med Vet Entomol* **20**(2): 229-238.
- Gonzalez M, Lopez S, Mullens BA, Baldet T, Goldarazena A (2013). A survey of *Culicoides* developmental sites on a farm in northern Spain, with a brief review of immature habitats of European species. *Vet Parasitol* **191**(1-2): 81-93.
- Goslee SC, Urban DL (2007). TheecodistPackage for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software* **22**(7): 1-19.
- Gotelli NJ (1991). Demographic Models for *Leptogorgia Virgulata*, A Shallow-Water Gorgonian. *Ecology* **72**(2): 457-467.
- Goudet J, Raymond M, de Meeus T, Rousset F (1996). Testing differentiation in diploid populations. *Genetics* **144**(4): 1933-1940.
- Goulson D, Kaden JC, Lepais O, Lye GC, Darvill B (2011). Population structure, dispersal and colonization history of the garden bumblebee *Bombus hortorum* in the Western Isles of Scotland. *Conserv Genet* **12**(4): 867-879.
- Graesboll K, Bodker R, Enoe C, Christiansen LE (2012). Simulating spread of Bluetongue Virus by flying vectors between hosts on pasture. *Sci Rep* **2**(1): 863.
- Graesboll K, Enoe C, Bodker R, Christiansen LE (2014). Optimal vaccination strategies against vector-borne diseases. *Spat Spatiotemporal Epidemiol* **11**(1877-5853 (Electronic)): 153-162.
- Graesboll K, Sumner T, Enoe C, Christiansen LE, Gubbins S (2016). A Comparison of Dynamics in Two Models for the Spread of a Vector-Borne Disease. *Transbound Emerg Dis* **63**(2): 215-223.
- Gratz NG (2004). Critical review of the vector status of *Aedes albopictus*. *Med Vet Entomol* **18**(3): 215-227.
- Graves TA, Beier P, Royle JA (2013). Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Mol Ecol* **22**(15): 3888-3903.
- Gubbins S, Carpenter S, Baylis M, Wood JL, Mellor PS (2008). Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. *J R Soc Interface* **5**(20): 363-371.
- Gubbins S, Turner J, Baylis M, van der Stede Y, van Schaik G, Abrahantes JC *et al* (2014). Inferences about the transmission of Schmallenberg virus within and between farms. *Prev Vet Med* **116**(4): 380-390.
- Guillot G, Leblois R, Coulon A, Frantz AC (2009). Statistical methods in spatial genetics. *Mol Ecol* **18**(23): 4734-4756.

- Guillot G, Mortier F, Estoup A (2005). GENELAND: a computer package for landscape genetics. *Mol Ecol Notes* **5**(3): 712-715.
- Guis H, Caminade C, Calvete C, Morse AP, Tran A, Baylis M (2012). Modelling the effects of past and future climate on the risk of bluetongue emergence in Europe. *J R Soc Interface* **9**(67): 339-350.
- Hagler JR, Jackson CG (2001). Methods for marking insects: Current techniques and future prospects. *Annu Rev Entomol* **46**(6): 511-543.
- Hagler JR, Naranjo SE, Machtley SA, Blackmer F (2014). Development of a standardized protein immunomarking protocol for insect mark-capture dispersal research. *J Appl Entomol* **138**(10): 772-782.
- Hajd Henni L, Sauvage F, Ninio C, Depaquit J, Augot D (2014). Wing geometry as a tool for discrimination of *Obsoletus* group (Diptera: Ceratopogonidae: *Culicoides*) in France. *Infect Genet Evol* **21**: 110-117.
- Hajibabaei M, Singer GA, Clare EL, Hebert PD (2007). Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biol* **5**(1): 24.
- Hakima B, Hwang HS, Lee KY (2020). Molecular identification of *Culicoides* (Diptera: Ceratopogonidae) species in Algeria. *Acta Trop* **202**: 105261.
- Handley LJJ, Estoup A, Evans DM, Thomas CE, Lombaert E, Facon B *et al* (2011). Ecological genetics of invasive alien species. *BioControl* **56**(4): 409-428.
- Hanski I (1998). Metapopulation dynamics. *Nature* **396**(6706): 41-49.
- Hanski I, Gaggiotti O (2004). *Ecology, genetics, and evolution of metapopulations*, Vol 67.
- Harbach RE (1994). Review of the Internal Classification of the Genus *Anopheles* (Diptera, Culicidae) - the Foundation for Comparative Systematics and Phylogenetic Research. *Bull Entomol Res* **84**(3): 331-342.
- Hardy AC, Cheng L (1986). Studies in the Distribution of Insects by Aerial Currents .3. Insect Drift over the Sea. *Ecol Entomol* **11**(3): 283-290.
- Hardy OJ (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Mol Ecol* **12**(6): 1577-1588.
- Hardy OJ, Vekemans X (2002). SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**(4): 618-620.
- Harrup LE, Bellis GA, Balenghien T, Garros C (2015). *Culicoides Latreille* (Diptera: Ceratopogonidae) taxonomy: current challenges and future directions. *Infect Genet Evol* **30**: 249-266.

- Harrup LE, Laban S, Purse BV, Reddy YK, Reddy YN, Byregowda SM *et al* (2016). DNA barcoding and surveillance sampling strategies for *Culicoides* biting midges (Diptera: Ceratopogonidae) in southern India. *Parasit Vectors* **9**: 461.
- Hastings RP, Sonuga-Barke EJ, Remington B (1993). An analysis of labels for people with learning disabilities. *Br J Clin Psychol* **32**(4): 463-465.
- Hebert PD, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**(1512): 313-321.
- Hendrickx G, Gilbert M, Staubach C, Elbers A, Mintiens K, Gerbier G *et al* (2008). A wind density model to quantify the airborne spread of *Culicoides* species during north-western Europe bluetongue epidemic, 2006. *Prev Vet Med* **87**(1-2): 162-181.
- Hendry G, Godwin G (1988). Biting midges in Scottish forestry: A costly irritant or a trivial nuisance? *Scottish Forestry* **42**: 113-119.
- Hewitt GM (1999). Post-glacial re-colonization of European biota. *Biol J Linn Soc* **68**(1-2): 87-112.
- Hewitt GM (2001). Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Mol Ecol* **10**(3): 537-549.
- Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* **359**(1442): 183-195; discussion 195.
- Hijmans RJ (2019). geosphere: Spherical Trigonometry.
- Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008). How many species are infected with *Wolbachia*? A statistical analysis of current data. *FEMS Microbiol Lett* **281**(2): 215-220.
- Hlaing T, Tun-Lin W, Somboon P, Socheat D, Setha T, Min S *et al* (2010). Spatial genetic structure of *Aedes aegypti* mosquitoes in mainland Southeast Asia. *Evol Appl* **3**(4): 319-339.
- Hoffmann B, Eschbaumer M, Beer M (2009). Real-time quantitative reverse transcription-PCR assays specifically detecting bluetongue virus serotypes 1, 6, and 8. *J Clin Microbiol* **47**(9): 2992-2994.
- Holderegger R, Kamm U, Gugerli F (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* **21**(6): 797-807.
- Holderegger R, Wagner HH (2008). Landscape genetics. *Bioscience* **58**(3): 199-207.
- Hristescu D, Barbuceanu F, Dascalu L, Nitescu C, Goffredo M, Santilli A *et al* (2020). Species composition and relative abundance of the genus *Culicoides* (Diptera: Ceratopogonidae) in Romania. *Parasit Vectors* **13**(1): 393.

- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* **9**(5): 1322-1332.
- Huestis DL, Dao A, Diallo M, Sanogo ZL, Samake D, Yaro AS *et al* (2019). Windborne long-distance migration of malaria mosquitoes in the Sahel. *Nature* **574**(7778): 404-408.
- Hurst GD, Frost CL (2015). Reproductive parasitism: maternally inherited symbionts in a biparental world. *Cold Spring Harb Perspect Biol* **7**(5): a017699.
- Ippoliti C, Goffredo M, Salini R, Bocci M, Quaglia M, Federici V *et al* (2016). Eco-climatic indicators for three *Culicoides* species of the *Obsoletus* complex in Italy. *Vet Ital* **52**(3-4): 213-222.
- Irwin DE (2012). Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *Am Nat* **180**(1): 35-49.
- Irwin DE, Bensch S, Irwin JH, Price TD (2005). Speciation by distance in a ring species. *Science* **307**(5708): 414-416.
- Itoua A, Vattier-Bernard G, Trouillet J (1987). Ceratopogonidae (Diptera) of the Congolese Mayombe. II. Contribution to the study of a *Culicoides* community. *Ann Parasitol Hum Comp* **62**(5): 474-483.
- Iturbe-Ormaetxe I, Walker T, SL ON (2011). *Wolbachia* and the biological control of mosquito-borne disease. *EMBO Rep* **12**(6): 508-518.
- Jacquet S, Garros C, Lombaert E, Walton C, Restrepo J, Allene X *et al* (2015). Colonization of the Mediterranean basin by the vector biting midge species *Culicoides imicola*: an old story. *Mol Ecol* **24**(22): 5707-5725.
- Jacquet S, Huber K, Guis H, Setier-Rio ML, Goffredo M, Allene X *et al* (2016a). Spatio-temporal genetic variation of the biting midge vector species *Culicoides imicola* (Ceratopogonidae) Kieffer in France. *Parasit Vectors* **9**: 141.
- Jacquet S, Huber K, Pages N, Talavera S, Burgin LE, Carpenter S *et al* (2016b). Range expansion of the Bluetongue vector, *Culicoides imicola*, in continental France likely due to rare wind-transport events. *Sci Rep* **6**: 27247.
- Jacquot M, Nomikou K, Palmarini M, Mertens P, Biek R (2017). Bluetongue virus spread in Europe is a consequence of climatic, landscape and vertebrate host factors as revealed by phylogeographic inference. *Proc Biol Sci* **284**(1864).
- James JE, Lanfear R, Eyre-Walker A (2016). Molecular Evolutionary Consequences of Island Colonization. *Genome Biol Evol* **8**(6): 1876-1888.
- Jay F, Sjodin P, Jakobsson M, Blum MG (2013). Anisotropic isolation by distance: the main orientations of human genetic differentiation. *Mol Biol Evol* **30**(3): 513-525.

- Jenkins A, Young M (2010). Investigation of *Culicoides* spp. preference for light colour and source using light emitting diodes and fluorescent light. *South African Journal of Animal Sciences* **40**.
- Johansen CA, Farrow RA, Morrisen A, Foley P, Bellis G, Van Den Hurk AF *et al* (2003). Collection of wind-borne haematophagous insects in the Torres Strait, Australia. *Med Vet Entomol* **17**(1): 102-109.
- Johnson CG, Taylor LR, Southwood TRE (1962). High Altitude Migration of *Oscinella frit* L. (Diptera: Chloropidae). *The Journal of Animal Ecology* **31**(2): 373-383.
- Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* **11**(1): 94.
- Jombart T, Devillard S, Dufour AB, Pontier D (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity (Edinb)* **101**(1): 92-103.
- Juliano SA, Lounibos LP (2005). Ecology of invasive mosquitoes: effects on resident species and on human health. *Ecol Lett* **8**(5): 558-574.
- Kamvar ZN, Tabima JF, Grunwald NJ (2014). Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*.
- Kaufmann C, Steinmann IC, Hegglin D, Schaffner F, Mathis A (2012). Spatio-temporal occurrence of *Culicoides* biting midges in the climatic regions of Switzerland, along with large scale species identification by MALDI-TOF mass spectrometry. *Parasit Vectors* **5**: 246.
- Keller D, Holderegger R, van Strien MJ, Bolliger J (2014). How to make landscape genetics beneficial for conservation management? *Conserv Genet* **16**(3): 503-512.
- Kenis M, Branco M (2010). Impact of alien terrestrial arthropods in Europe. Chapter 5. *BioRisk* **4**: 51-71.
- Kenis M, Hurley BP, Hajek AE, Cock MJW (2017). Classical biological control of insect pests of trees: facts and figures. *Biol Invasions* **19**(11): 3401-3417.
- Kerdelhue C, Zane L, Simonato M, Salvato P, Rousselet J, Roques A *et al* (2009). Quaternary history and contemporary patterns in a currently expanding species. *BMC Evol Biol* **9**: 220.
- Kettle DS, Edwards PB, Barnes A (1998). Factors affecting numbers of *Culicoides* in truck traps in coastal Queensland. *Med Vet Entomol* **12**(4): 367-377.
- Kimmig SE, Beninde J, Brandt M, Schleimer A, Kramer-Schadt S, Hofer H *et al* (2020). Beyond the landscape: Resistance modelling infers physical and behavioural gene flow barriers to a mobile carnivore across a metropolitan area. *Mol Ecol* **29**(3): 466-484.

- Kimura M (1953). "Stepping stone" model of population. *Ann Rep Nat Inst Genet Japan* **3**: 62-63.
- Kirk H, Dorn S, Mazzi D (2013). Worldwide population genetic structure of the oriental fruit moth (*Grapholita molesta*), a globally invasive pest. *BMC Ecol* **13**: 12.
- Kirkeby C, Bodker R, Stockmarr A, Lind P, Heegaard PM (2013). Quantifying dispersal of european *Culicoides* (Diptera: Ceratopogonidae) vectors between farms using a novel mark-release-recapture technique. *PLoS One* **8**(4): e61269.
- Kleider N, Lees MJ (1984). *Culicoides* hypersensitivity in the horse: 15 cases in southwestern british columbia. *Can Vet J* **25**(1): 26-32.
- Klimaszewski J, Webster RP, Bourdon C, Pelletier G, Godin B, Langor DW (2015). Review of Canadian species of the genus *Mocyta* Mulsant & Rey (Coleoptera, Staphylinidae, Aleocharinae), with the description of a new species and a new synonymy. *Zookeys* **487**(487): 111-139.
- Kluiters G, Hunter K, Baylis M (2015a). Validation of fluorescent dust marking of *Culicoides* biting midges and the design of a self-marking technique. *Parasit Vectors* **8**: 54.
- Kluiters G, Pages N, Carpenter S, Gardes L, Guis H, Baylis M *et al* (2016). Morphometric discrimination of two sympatric sibling species in the Palaearctic region, *Culicoides obsoletus* Meigen and *C. scoticus* Downes & Kettle (Diptera: Ceratopogonidae), vectors of bluetongue and Schmallenberg viruses. *Parasit Vectors* **9**: 262.
- Kluiters G, Swales H, Baylis M (2015b). Local dispersal of palaeartic *Culicoides* biting midges estimated by mark-release-recapture. *Parasit Vectors* **8**: 86.
- Kolleck J, Yang M, Zinner D, Roos C (2013). Genetic diversity in endangered Guizhou snub-nosed monkeys (*Rhinopithecus brelichi*): contrasting results from microsatellite and mitochondrial DNA data. *PLoS One* **8**(8): e73647.
- Kosakovsky Pond SL, Frost SD (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* **22**(5): 1208-1222.
- Kozakiewicz CP, Burrige CP, Funk WC, VandeWoude S, Craft ME, Crooks KR *et al* (2018). Pathogens in space: Advancing understanding of pathogen dynamics and disease ecology through landscape genetics. *Evol Appl* **11**(10): 1763-1778.
- Krysan JL (1986). Introduction: Biology, Distribution, and Identification of Pest Diabrotica. In: Krysan JL and Miller TA (eds) *Methods for the Study of Pest Diabrotica*. Springer New York: New York, NY, pp 1-23.
- Lagercrantz U, Ryman N (1990). Genetic Structure of Norway Spruce (*Picea Abies*): Concordance of Morphological and Allozymic Variation. *Evolution* **44**(1): 38-53.
- Landguth EL, Cushman SA (2010). cdpop: A spatially explicit cost distance population genetics program. *Mol Ecol Resour* **10**(1): 156-161.

- Landguth EL, Cushman SA, Murphy MA, Luikart G (2010). Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Mol Ecol Resour* **10**(5): 854-862.
- Lassen SB, Nielsen SA, Skovgard H, Kristensen M (2011). Molecular identification of bloodmeals from biting midges (Diptera: Ceratopogonidae: *Culicoides* Latreille) in Denmark. *Parasitol Res* **108**(4): 823-829.
- Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv Genet* **7**(2): 295-302.
- Leache AD, Fujita MK (2010). Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc Biol Sci* **277**(1697): 3071-3077.
- Leblois R, Rousset F, Estoup A (2004). Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* **166**(2): 1081-1092.
- Lebreton JD, Hines JE, Pradel R, Nichols JD, Spendelov JA (2003). Estimation by capture-recapture of recruitment and dispersal over several sites. *Oikos* **101**(2): 253-264.
- Lecocq T, Dellicour S, Michez D, Lhomme P, Vanderplanck M, Valterova I *et al* (2013). Scent of a break-up: phylogeography and reproductive trait divergences in the red-tailed bumblebee (*Bombus lapidarius*). *BMC Evol Biol* **13**: 263.
- Lecocq T, Rasmont P, Harpke A, Schweiger O (2016). Improving International Trade Regulation by Considering Intraspecific Variation for Invasion Risk Assessment of Commercially Traded Species: The *Bombus terrestris* Case. *Conservation Letters* **9**(4): 281-289.
- Lefebvre T, Vargo EL, Zimmermann M, Dupont S, Kutnik M, Bagnères AG (2016). Subterranean termite phylogeography reveals multiple postglacial colonization events in southwestern Europe. *Ecol Evol* **6**(16): 5987-6004.
- Lefevre T, Raymond M, Thomas F (2016). *Biologie évolutive*. De Boeck supérieur.
- Lehmann T, Hawley WA, Kamau L, Fontenille D, Simard F, Collins FH (1996). Genetic differentiation of *Anopheles gambiae* populations from East and west Africa: comparison of microsatellite and allozyme loci. *Heredity (Edinb)* **77** (Pt 2)(2): 192-200.
- Lemey P, Rambaut A, Welch JJ, Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Mol Biol Evol* **27**(8): 1877-1885.
- Lesieur V, Lombaert E, Guillemaud T, Courtial B, Strong W, Roques A *et al* (2018). The rapid spread of *Leptoglossus occidentalis* in Europe: a bridgehead invasion. *J Pest Sci* **92**(1): 189-200.

- Levins R (1969). Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *Bulletin of the Entomological Society of America* **15**(3): 237-240.
- Lewis SE, Rice A, Hurst GD, Baylis M (2014). First detection of endosymbiotic bacteria in biting midges *Culicoides pulicaris* and *Culicoides punctatus*, important Palaearctic vectors of bluetongue virus. *Med Vet Entomol* **28**(4): 453-456.
- Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**(11): 1451-1452.
- Lichstein JW (2006). Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol* **188**(2): 117-131.
- Lillie TH, Marquardt WC, Jones RH (1981). The Flight Range of *Culicoides variipennis* (Diptera, Ceratopogonidae). *Can Entomol* **113**(5): 419-426.
- Linden A, Gregoire F, Nahayo A, Hanrez D, Mousset B, Massart AL *et al* (2010). Bluetongue virus in wild deer, Belgium, 2005-2008. *Emerg Infect Dis* **16**(5): 833-836.
- Lindh E, Argentini C, Remoli ME, Fortuna C, Faggioni G, Benedetti E *et al* (2019). The Italian 2017 Outbreak Chikungunya Virus Belongs to an Emerging *Aedes albopictus*-Adapted Virus Cluster Introduced From the Indian Subcontinent. *Open Forum Infect Dis* **6**(1): ofy321.
- Linley JR, Davies JB (1971). Sandflies and Tourism in Florida and the Bahamas and Caribbean Area1. *J Econ Entomol* **64**(1): 264-278.
- Linton YM, Mordue AJ, Cruickshank RH, Meiswinkel R, Mellor PS, Dallas JF (2002). Phylogenetic analysis of the mitochondrial cytochrome oxidase subunit I gene of five species of the *Culicoides imicola* species complex. *Medical and Veterinary Entomology* **16**(2): 139-146.
- Loiselle BA, Sork VL, Nason J, Graham C (1995). Spatial Genetic-Structure of a Tropical Understory Shrub, *Psychotria Officinalis* (Rubiaceae). *Am J Bot* **82**(11): 1420-1425.
- Lopes JS, Beaumont MA (2010). ABC: a useful Bayesian tool for the analysis of population data. *Infect Genet Evol* **10**(6): 826-833.
- Lopez-Olvera JR, Falconi C, Fernandez-Pacheco P, Fernandez-Pinero J, Sanchez MA, Palma A *et al* (2010). Experimental infection of European red deer (*Cervus elaphus*) with bluetongue virus serotypes 1 and 8. *Vet Microbiol* **145**(1-2): 148-152.
- Lorca-Oro C, Lopez-Olvera JR, Ruiz-Fons F, Acevedo P, Garcia-Bocanegra I, Oleaga A *et al* (2014). Long-term dynamics of bluetongue virus in wild ruminants: relationship with outbreaks in livestock in Spain, 2006-2011. *PLoS One* **9**(6): e100027.
- Louy D, Habel JC, Schmitt T, Assmann T, Meyer M, Muller P (2007). Strongly diverging population genetic patterns of three skipper species: the role of habitat fragmentation and dispersal ability. *Conserv Genet* **8**(3): 671-681.

- Lowe WH, Allendorf FW (2010). What can genetics tell us about population connectivity? *Mol Ecol* **19**(15): 3038-3051.
- Lutz Tischendorf, Fahrig L (2000). How should we measure landscape connectivity? *Landscape Ecology* **15**: 633-641.
- Luximon N, Petit EJ, Broquet T (2014). Performance of individual vs. group sampling for inferring dispersal under isolation-by-distance. *Mol Ecol Resour* **14**(4): 745-752.
- Ma Y, Xu J, Yang Z, Wang X, Lin Z, Zhao W *et al* (2013). A video clip of the biting midge *Culicoides anophelis* ingesting blood from an engorged *Anopheles* mosquito in Hainan, China. *Parasit Vectors* **6**(1): 326.
- MacArthur RH, and E. O. Wilson. (1967). The Theory of Island Biogeography. *Princeton, NJ: Princeton University Press.*
- Maclachlan NJ (2011). Bluetongue: history, global epidemiology, and pathogenesis. *Prev Vet Med* **102**(2): 107-111.
- Maclachlan NJ, Guthrie AJ (2010). Re-emergence of bluetongue, African horse sickness, and other orbivirus diseases. *Vet Res* **41**(6): 35.
- Maclachlan NJ, Zientara S, Wilson WC, Richt JA, Savini G (2019). Bluetongue and epizootic hemorrhagic disease viruses: recent developments with these globally re-emerging arboviral infections of ruminants. *Curr Opin Virol* **34**: 56-62.
- Magee D, Beard R, Suchard MA, Lemey P, Scotch M (2015). Combining phylogeography and spatial epidemiology to uncover predictors of H5N1 influenza A virus diffusion. *Arch Virol* **160**(1): 215-224.
- Manangwa O, De Meeus T, Grebaut P, Segard A, Byamungu M, Ravel S (2019). Detecting Wahlund effects together with amplification problems: Cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. *Mol Ecol Resour* **19**(3): 757-772.
- Manel S, Berthoud F, Bellemain E, Gaudeul M, Luikart G, Swenson JE *et al* (2007). A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Mol Ecol* **16**(10): 2031-2043.
- Manel S, Gaggiotti OE, Waples RS (2005). Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol Evol* **20**(3): 136-142.
- Manel S, Holderegger R (2013). Ten years of landscape genetics. *Trends Ecol Evol* **28**(10): 614-621.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* **18**(4): 189-197.

- Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**(2): 209-220.
- Mardulyn P, Goffredo M, Conte A, Hendrickx G, Meiswinkel R, Balenghien T *et al* (2013). Climate change and the spread of vector-borne diseases: using approximate Bayesian computation to compare invasion scenarios for the bluetongue virus vector *Culicoides imicola* in Italy. *Mol Ecol* **22**(9): 2456-2466.
- Martinez-de la Puente J, Martinez J, Ferraguti M, Morales-de la Nuez A, Castro N, Figuerola J (2012). Genetic characterization and molecular identification of the bloodmeal sources of the potential bluetongue vector *Culicoides obsoletus* in the Canary Islands, Spain. *Parasit Vectors* **5**: 147.
- Mathieu B (2011). Les espèces de *Culicoides* du sous-genre *Avaritia* (Diptera : Ceratopogonidae) dans le monde : révision systématique et taxonomique des espèces d'intérêt dans la transmission d'Orbivirus. 253.
- Mathieu B, Delecolle JC, Garros C, Balenghien T, Setier-Rio ML, Candolfi E *et al* (2011). Simultaneous quantification of the relative abundance of species complex members: application to *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae), potential vectors of bluetongue virus. *Vet Parasitol* **182**(2-4): 297-306.
- Mathieu B, Perrin A, Baldet T, Delecolle JC, Albina E, Cetre-Sossah C (2007). Molecular identification of Western European species of *obsoletus* complex (Diptera: Ceratopogonidae) by an internal transcribed spacer-1 rDNA multiplex polymerase chain reaction assay. *J Med Entomol* **44**(6): 1019-1025.
- Mayden RL (1999). Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. *J Nematol* **31**(2): 95-116.
- Mayer F, Piel FB, Cassel-Lundhagen A, Kirichenko N, Grumiau L, Okland B *et al* (2015). Comparative multilocus phylogeography of two Palaeartic spruce bark beetles: influence of contrasting ecological strategies on genetic variation. *Mol Ecol* **24**(6): 1292-1310.
- Mayr E (1982). Speciation and Macroevolution. *Evolution* **36**(6): 1119-1132.
- McGrath G, More SJ, O'Neill R (2018). Hypothetical route of the introduction of Schmallenberg virus into Ireland using two complementary analyses. *Vet Rec* **182**(8): 226.
- McLean AH, Parker BJ, Hrcek J, Henry LM, Godfray HC (2016). Insect symbionts in food webs. *Philos Trans R Soc Lond B Biol Sci* **371**(1702): 20150325.
- McMeniman CJ, Lane RV, Cass BN, Fong AW, Sidhu M, Wang YF *et al* (2009). Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. *Science* **323**(5910): 141-144.
- McMichael AJ (2013). Globalization, climate change, and human health. *N Engl J Med* **368**(14): 1335-1343.

- McRae B, Shah VB (2008). Circuitscape: A Tool for Landscape Ecology.
- McRae BH (2006). Isolation by resistance. *Evolution* **60**(8): 1551-1561.
- McRae BH, Beier P (2007). Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci U S A* **104**(50): 19885-19890.
- Medlock J, Atkins KE, Thomas DN, Aksoy S, Galvani AP (2013). Evaluating paratransgenesis as a potential control strategy for African trypanosomiasis. *PLoS Negl Trop Dis* **7**(8): e2374.
- Meentemeyer RK, Haas SE, Vaclavik T (2012). Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. *Annu Rev Phytopathol* **50**(1): 379-402.
- Mehlhorn H, Walldorf V, Klimpel S, Jahn B, Jaeger F, Eschweiler J *et al* (2007). First occurrence of *Culicoides obsoletus*-transmitted Bluetongue virus epidemic in Central Europe. *Parasitol Res* **101**(1): 219-228.
- Mehlhorn H, Walldorf V, Klimpel S, Schaub G, Kiel E, Focke R *et al* (2009). Bluetongue disease in Germany (2007-2008): monitoring of entomological aspects. *Parasitol Res* **105**(2): 313-319.
- Meirmans PG (2015). Seven common mistakes in population genetics and how to avoid them over confidence need diff analyses diff scales needed may not be able to do patterns errors. *Mol Ecol* **24**: 3223-3231.
- Meiswinkel R, De Bree F, Bossers-De Vries R, Elbers AR (2015). An unrecognized species of the *Culicoides obsoletus* complex feeding on livestock in The Netherlands. *Vet Parasitol* **207**(3-4): 324-328.
- Meiswinkel R, Gomulski LM, Delecolle JC, Goffredo M, Gasperi G (2004). The taxonomy of *Culicoides* vector complexes - unfinished business. *Vet Ital* **40**(3): 151-159.
- Meiswinkel R, Scolamacchia F, Dik M, Mudde J, Dijkstra E, Van Der Ven IJ *et al* (2014). The Mondrian matrix: *Culicoides* biting midge abundance and seasonal incidence during the 2006-2008 epidemic of bluetongue in the Netherlands. *Med Vet Entomol* **28**(1): 10-20.
- Meiswinkel R, van Rijn P, Leijts P, Goffredo M (2007). Potential new *Culicoides* vector of bluetongue virus in northern Europe. *Vet Rec* **161**(16): 564-565.
- Mellor PS, Boorman J, Baylis M (2000). *Culicoides* biting midges: their role as arbovirus vectors. *Annu Rev Entomol* **45**: 307-340.
- Mellor PS, Boorman JPT, Wilkinson PJ, Martinez-Gomez F (1983). Potential vectors of bluetongue and African horse sickness viruses in Spain. *Vet Rec* **112**(10): 229-230.

- Mellor PS, Carpenter S, Harrup L, Baylis M, Mertens PP (2008). Bluetongue in Europe and the Mediterranean Basin: history of occurrence prior to 2006. *Prev Vet Med* **87**(1-2): 4-20.
- Meloni G, Cossu M, Foxi C, Vento L, Circosta S, Burrari E *et al* (2018). Combined larvicidal and adulticidal treatments to control *Culicoides* biting midges (Diptera: Ceratopogonidae): Results of a pilot study. *Vet Parasitol* **257**: 28-33.
- Menzies FD, McCullough SJ, McKeown IM, Forster JL, Jess S, Batten C *et al* (2008). Evidence for transplacental and contact transmission of bluetongue virus in cattle. *Vet Rec* **163**(7): 203-209.
- Miao J, Wu Y-Q, Gong Z-J, He Y-Z, Duan Y, Jiang Y-L (2012). Long-Distance Wind-Borne Dispersal of *Sitodiplosis mosellana* Géhin (Diptera:Cecidomyiidae) in Northern China. *J Insect Behav* **26**(1): 120-129.
- Mignotte A, Garros C, Dellicour S, Gilbert M, Gardès L, Balenghien T *et al* (Submitted in 2020). *Important Dispersal Capacity of One Main Species Vector of Bluetongue and Schmallenberg Viruses Culicoides Obsoletus (Diptera: Ceratopogonidae) Revealed by Landscape Genetic Analyses.*
- Mignotte A, Garros C, Gardes L, Balenghien T, Duhayon M, Rakotoarivony I *et al* (2020). The tree that hides the forest: cryptic diversity and phylogenetic relationships in the Palearctic vector Obsoletus/Scoticus Complex (Diptera: Ceratopogonidae) at the European level. *Parasites & Vectors* **13**(1): 265.
- Miller MP (2005). Alleles In Space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. *J Hered* **96**(6): 722-724.
- Miller N, Estoup A, Toepfer S, Bourguet D, Lapchin L, Derridj S *et al* (2005). Multiple transatlantic introductions of the western corn rootworm. *Science* **310**(5750): 992-992.
- Miranda MA, Borrás D, Rincon C, Alemany A (2003). Presence in the Balearic Islands (Spain) of the midges *Culicoides imicola* and *Culicoides obsoletus* group. *Medical and Veterinary Entomology* **17**(1): 52-54.
- Mohlmann TWR, Bekendam AM, van Kemenade I, Wennergren U, Favia G, Takken W *et al* (2019). Latitudinal diversity of biting midge species within the *Obsoletus* group across three habitats in Europe. *Med Vet Entomol* **33**(3): 420-426.
- Monaco F, Benedetto L, Di Marcello V, Lelli R, Goffredo M (2010). Development and preliminary evaluation of a real-time polymerase chain reaction for the identification of *Culicoides obsoletus sensu strictu*, *C. scoticus* and *C. montanus* in the Obsoletus Complex in Italy. *Vet Ital* **46**(2): 209-214, 215-220.
- Morag N, Klement E, Saroya Y, Lensky I, Gottlieb Y (2012). Prevalence of the symbiont *Cardinium* in *Culicoides* (Diptera: Ceratopogonidae) vector species is associated with land surface temperature. *FASEB J* **26**(10): 4025-4034.
- Muir LE, Kay BH (1998). *Aedes aegypti* survival and dispersal estimated by mark-release-recapture in northern Australia. *Am J Trop Med Hyg* **58**(3): 277-282.

- Mullens BA, Gerry AC, Monteys VS, Pinna M, Gonzalez A (2010). Field studies on *Culicoides* (Diptera: Ceratopogonidae) activity and response to deltamethrin applications to sheep in northeastern Spain. *J Med Entomol* **47**(1): 106-110.
- Mullens BA, McDermott EG, Gerry AC (2015). Progress and knowledge gaps in *Culicoides* ecology and control. *Vet Ital* **51**(4): 313-323.
- Murphy MA, Dezzani R, Pilliod DS, Storfer A (2010). Landscape genetics of high mountain frog metapopulations. *Mol Ecol* **19**(17): 3634-3649.
- Murray MD (1987). Local Dispersal of the Biting-Midge *Culicoides-Brevitarsis* Kieffer (Diptera, Ceratopogonidae) in Southeastern Australia. *Aust J Zool* **35**(6): 559-573.
- Murray MD, Kirkland PD (1995). Bluetongue and Douglas virus activity in New South Wales in 1989 : further evidence for long-distance dispersal of the biting midge *Culicoides brevitarsis*. *Australian Veterinary Journal* **Vol. 72**.
- Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* **8**(7): e1002764.
- Nei M, Maruyama T, Chakraborty R (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution* **29**(1): 1-10.
- Neil A, David M, Jolyon m, Kate S, Willy W (2014). *A first attempt at modelling roe deer (Capreolus capreolus) distributions over Europe*.
- Newton RG, Spurrell DJ (1967). Examples of the Use of Elements for Clarifying Regression Analyses. *Applied Statistics* **16**(2): 165-172.
- Nimon K, Lewis M, Kane R, Haynes RM (2008). An R package to compute commonality coefficients in the multiple regression case: an introduction to the package and a practical example. *Behav Res Methods* **40**(2): 457-466.
- Nolan DV, Carpenter S, Barber J, Mellor PS, Dallas JF, Mordue Luntz AJ *et al* (2007). Rapid diagnostic PCR assays for members of the *Culicoides obsoletus* and *Culicoides pulicaris* species complexes, implicated vectors of bluetongue virus in Europe. *Vet Microbiol* **124**(1-2): 82-94.
- Nolan DV, Dallas JF, Piertney SB, Mordue Luntz AJ (2008). Incursion and range expansion in the bluetongue vector *Culicoides imicola* in the Mediterranean basin: a phylogeographic analysis. *Med Vet Entomol* **22**(4): 340-351.
- Nychka D, Furrer R, Paige J, Sain S (2017). *Fields: Tools for spatial data*.
- O'Brien TG, Pollock KH, Davidson WR, Kellogg FE (1985). A Comparison of Capture-Recapture with Capture-Removal for Quail Populations. *The Journal of Wildlife Management* **49**(4): 1062-1066.

- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB *et al* (2013). Vegan: Community Ecology Package. R Package Version. 2.0-10. *CRAN*.
- Onyango MG, Beebe NW, Gopurenko D, Bellis G, Nicholas A, Ogugo M *et al* (2015a). Assessment of population genetic structure in the arbovirus vector midge, *Culicoides brevitarsis* (Diptera: Ceratopogonidae), using multi-locus DNA microsatellites. *Vet Res* **46**: 108.
- Onyango MG, Michuki GN, Ogugo M, Venter GJ, Miranda MA, Elissa N *et al* (2015b). Delineation of the population genetic structure of *Culicoides imicola* in East and South Africa. *Parasit Vectors* **8**: 660.
- Ostfeld RS, Glass GE, Keesing F (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends Ecol Evol* **20**(6): 328-336.
- Paetkau D, Slade R, Burden M, Estoup A (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* **13**(1): 55-65.
- Pages N, Munoz-Munoz F, Talavera S, Sarto V, Lorca C, Nunez JI (2009). Identification of cryptic species of *Culicoides* (Diptera: Ceratopogonidae) in the subgenus *Culicoides* and development of species-specific PCR assays based on barcode regions. *Vet Parasitol* **165**(3-4): 298-310.
- Pages N, Munoz-Munoz F, Verdun M, Pujol N, Talavera S (2017). First detection of *Wolbachia*-infected *Culicoides* (Diptera: Ceratopogonidae) in Europe: *Wolbachia* and *Cardinium* infection across *Culicoides* communities revealed in Spain. *Parasit Vectors* **10**(1): 582.
- Pages N, Sarto IMV (2005). Differentiation of *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae) based on mitochondrial cytochrome oxidase subunit I. *J Med Entomol* **42**(6): 1026-1034.
- Parry HR, Macfadyen S, Hopkinson JE, Bianchi FJJA, Zalucki MP, Bourne A *et al* (2015). Plant composition modulates arthropod pest and predator abundance: Evidence for culling exotics and planting natives. *Basic and Applied Ecology* **16**(6): 531-543.
- Pascual M, Chapuis MP, Mestres F, Balanya J, Huey RB, Gilchrist GW *et al* (2007). Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Mol Ecol* **16**(15): 3069-3083.
- Pentinsaari M, Vos R, Mutanen M (2017). Algorithmic single-locus species delimitation: effects of sampling effort, variation and nonmonophyly in four methods and 1870 species of beetles. *Mol Ecol Resour* **17**(3): 393-404.
- Petit R, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R *et al* (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**(5625): 1563-1565.

- Petkova D, Novembre J, Stephens M (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nat Genet* **48**(1): 94-100.
- Pfluger FJ, Balkenhol N (2014). A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Mol Ecol* **23**(9): 2146-2156.
- Phillips BL, Brown GP, Shine R (2010). Life-history evolution in range-shifting populations. *Ecology* **91**(6): 1617-1627.
- Phillipsen IC, Kirk EH, Bogan MT, Mims MC, Olden JD, Lytle DA (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Mol Ecol* **24**(1): 54-69.
- Pilgrim J, Ander M, Garros C, Baylis M, Hurst GDD, Siozios S (2017). Torix group *Rickettsia* are widespread in *Culicoides* biting midges (Diptera: Ceratopogonidae), reach high frequency and carry unique genomic features. *Environ Microbiol* **19**(10): 4238-4255.
- Pilgrim J, Siozios S, Baylis M, Hurst GDD (2020). Tissue Tropisms and Transstadial Transmission of a *Rickettsia* Endosymbiont in the Highland Midge, *Culicoides impunctatus* (Diptera: Ceratopogonidae). *Appl Environ Microbiol* **86**(20): e01492-01420.
- Pili E, Carcangiu L, Oppo M, Marchi A (2010). Genetic structure and population dynamics of the biting midges *Culicoides obsoletus* and *Culicoides scoticus*: implications for the transmission and maintenance of bluetongue. *Med Vet Entomol* **24**(4): 441-448.
- Pioz M, Guis H, Crespin L, Gay E, Calavas D, Durand B *et al* (2012). Why did bluetongue spread the way it did? Environmental factors influencing the velocity of bluetongue virus serotype 8 epizootic wave in France. *PLoS One* **7**(8): e43360.
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J Hered* **95**(6): 536-539.
- Piry S, Chapuis MP, Gauffre B, Papaix J, Cruaud A, Berthier K (2016). Mapping Averaged Pairwise Information (MAPI): a new exploratory tool to uncover spatial structure. *Methods Ecol Evol* **7**(12): 1463-1475.
- Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S *et al* (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* **55**(4): 595-609.
- Pons O, Petit R (1996). Measuring and testing genetic differentiation with ordered vs. unordered alleles. *Genetics* **144**: 1237-1245.
- Pons O, Petit RJ (1995). Estimation, variance and optimal sampling of gene diversity : I. Haploid locus. *Theor Appl Genet* **90**(3-4): 462-470.

- Pringle JM, Blakeslee AM, Byers JE, Roman J (2011). Asymmetric dispersal allows an upstream region to control population structure throughout a species' range. *Proc Natl Acad Sci U S A* **108**(37): 15288-15293.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2): 945-959.
- Provan J, Bennett KD (2008). Phylogeographic insights into cryptic glacial refugia. *Trends Ecol Evol* **23**(10): 564-571.
- Prugnolle F, de Meeus T (2002). Inferring sex-biased dispersal from population genetic tools: a review. *Heredity (Edinb)* **88**(3): 161-165.
- Prugnolle F, Liu H, de Meeus T, Balloux F (2005). Population genetics of complex life-cycle parasites: an illustration with trematodes. *Int J Parasitol* **35**(3): 255-263.
- Prunier JG, Colyn M, Legendre X, Flamand MC (2017). Regression commonality analyses on hierarchical genetic distances. *Ecography* **40**(12): 1412-1425.
- Prunier JG, Colyn M, Legendre X, Nimon KF, Flamand MC (2015). Multicollinearity in spatial genetics: separating the wheat from the chaff using commonality analyses. *Mol Ecol* **24**(2): 263-283.
- Prunier JG, Kaufmann B, Fenet S, Picard D, Pompanon F, Joly P *et al* (2013a). Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: towards a better assessment of functional connectivity using an individual-based sampling scheme. *Mol Ecol* **22**(22): 5516-5530.
- Prunier JG, Kaufmann B, Léna J-P, Fenet S, Pompanon F, Joly P (2013b). A 40-year-old divided highway does not prevent gene flow in the alpine newt *Ichthyosaura alpestris*. *Conserv Genet* **15**(2): 453-468.
- Pudar D, Petric D, Allene X, Alten B, Ayhan N, Cvetkovikj A *et al* (2018). An update of the *Culicoides* (Diptera: Ceratopogonidae) checklist for the Balkans. *Parasit Vectors* **11**(1): 462.
- Puillandre N, Lambert A, Brouillet S, Achaz G (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol Ecol* **21**(8): 1864-1877.
- Pulliam HR (1988). Sources, Sinks, and Population Regulation. *Am Nat* **132**(5): 652-661.
- Purse BV, Carpenter S, Venter GJ, Bellis G, Mullens BA (2015). Bionomics of temperate and tropical *Culicoides* midges: knowledge gaps and consequences for transmission of *Culicoides*-borne viruses. *Annu Rev Entomol* **60**: 373-392.
- Purse BV, Mellor PS, Rogers DJ, Samuel AR, Mertens PP, Baylis M (2005). Climate change and the recent emergence of bluetongue in Europe. *Nat Rev Microbiol* **3**(2): 171-181.
- Rainey RC (1973). Airborne Pests and the Atmospheric Environment*. *Weather* **28**(6): 224-239.

- Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**(12): 2092-2100.
- Rannala B, Mountain JL (1997). Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci U S A* **94**(17): 9197-9201.
- Ranta E, Kaitala V, Lindstrom J, Helle E (1997). The Moran effect and synchrony in population dynamics. *Oikos* **78**(1): 136-142.
- Rasic G, Schama R, Powell R, Maciel-de Freitas R, Endersby-Harshman NM, Filipovic I *et al* (2015). Contrasting genetic structure between mitochondrial and nuclear markers in the dengue fever mosquito from Rio de Janeiro: implications for vector control. *Evol Appl* **8**(9): 901-915.
- Reid NM, Carstens BC (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol Biol* **12**: 196.
- Reynolds AM, Reynolds DR (2009). Aphid aerial density profiles are consistent with turbulent advection amplifying flight behaviours: abandoning the epithet 'passive'. *Proc Biol Sci* **276**(1654): 137-143.
- Reynolds DR, Chapman JW, Harrington R (2006). The migration of insect vectors of plant and animal viruses. *Adv Virus Res* **67**: 453-517.
- Rezza G, Nicoletti L, Angelini R, Romi R, Finarelli AC, Panning M *et al* (2007). Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet* **370**(9602): 1840-1846.
- Rheindt FE, Edwards SV (2011). Genetic Introgression: An Integral but Neglected Component of Speciation in Birds. *Auk* **128**(4): 620-632.
- Rice WR (1989). Analysing tables of statistic tests. *Evolution* **43**(1): 223-225.
- Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol Lett* **9**(8): 981-993.
- Richardson DM, Pysek P, Rejmanek M, Barbour MG, Panetta FD, West CJ (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity* **6**(2): 93-107.
- Richardson JL, Brady SP, Wang IJ, Spear SF (2016). Navigating the pitfalls and promise of landscape genetics. *Mol Ecol* **25**(4): 849-863.
- Rieux A, Halkett F, de Lapeyre de Bellaire L, Zapater MF, Rousset F, Ravigne V *et al* (2011). Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. *Mol Ecol* **20**(8): 1661-1674.

- Rieux A, Soubeyrand S, Bonnot F, Klein EK, Ngando JE, Mehl A *et al* (2014). Long-distance wind-dispersal of spores in a fungal plant pathogen: estimation of anisotropic dispersal kernels from an extensive field experiment. *PLoS One* **9**(8): e103225.
- Rio RVM, Attardo GM, Weiss BL (2016). Grandeur Alliances: Symbiont Metabolic Integration and Obligate Arthropod Hematophagy. *Trends Parasitol* **32**(9): 739-749.
- Rodhain F, Bruce-Chwatt L, Perez C (1985). *Précis d'entomologie médicale et vétérinaire : notions d'épidémiologie des maladies à vecteurs*. Paris : Maloine.
- Rodríguez-Sánchez F, Hampe A, Jordano P, Arroyo J (2010). Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: A review. *Rev Palaeobot Palynol* **162**(3): 507-521.
- Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* **9**(3): 552-569.
- Ronce O (2007). How Does It Feel to Be Like a Rolling Stone? Ten Questions About Dispersal Evolution. *Annu Rev Ecol Evol Syst* **38**(1): 231-253.
- Ronce O, Clobert J (2012). Dispersal syndromes *Dispersal Ecology and Evolution*, pp 119-138.
- Rosenberg MS (2010). The Bearing Correlogram: A New Method of Analyzing Directional Spatial Autocorrelation. *Geographical Analysis* **32**(3): 267-278.
- Rougemont Q, Bernatchez L (2018). The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations. *Evolution* **72**(6): 1261-1277.
- Rousset F (2000). Genetic differentiation between individuals. *J Evol Biol* **13**(1): 58-62.
- Ruiz-Fons F, Sanchez-Matamoros A, Gortazar C, Sanchez-Vizcaino JM (2014). The role of wildlife in bluetongue virus maintenance in Europe: lessons learned after the natural infection in Spain. *Virus Res* **182**: 50-58.
- Russell RC, Webb CE, Williams CR, Ritchie SA (2005). Mark-release-recapture study to measure dispersal of the mosquito *Aedes aegypti* in Cairns, Queensland, Australia. *Med Vet Entomol* **19**(4): 451-457.
- Saastamoinen M, Bocedi G, Cote J, Legrand D, Guillaume F, Wheat CW *et al* (2018). Genetics of dispersal. *Biol Rev Camb Philos Soc* **93**(1): 574-599.
- Sailleau C, Breard E, Viarouge C, Vitour D, Romey A, Garnier A *et al* (2017). Re-Emergence of Bluetongue Virus Serotype 8 in France, 2015. *Transbound Emerg Dis* **64**(3): 998-1000.
- Sanders CJ, Carpenter S (2014). Assessment of an immunomarking technique for the study of dispersal of *Culicoides* biting midges. *Infect Genet Evol* **28**: 583-587.

- Sanders CJ, Harrup LE, Tugwell LA, Brugman VA, England M, Carpenter S (2017). Quantification of within- and between-farm dispersal of *Culicoides* biting midges using an immunomarking technique. *J Appl Ecol* **54**(5): 1429-1439.
- Sanders CJ, Shortall CR, Gubbins S, Burgin L, Gloster J, Harrington R *et al* (2011). Influence of season and meteorological parameters on flight activity of *Culicoides* biting midges. *J Appl Ecol* **48**(6): 1355-1364.
- Sanogo ZL, Yaro AS, Dao A, Diallo M, Yossi O, Samake D *et al* (2020). The Effects of High-Altitude Windborne Migration on Survival, Oviposition, and Blood-Feeding of the African Malaria Mosquito, *Anopheles gambiae s.l.* (Diptera: Culicidae). *J Med Entomol*.
- Santman-Berends IM, van Wuijckhuise L, Vellema P, van Rijn PA (2010). Vertical transmission of bluetongue virus serotype 8 virus in Dutch dairy herds in 2007. *Vet Microbiol* **141**(1-2): 31-35.
- Sappington T, Siegfried B, Guillemaud T (2006). Coordinated Diabrotica Genetics Research: Accelerating Progress on an Urgent Insect Pest Problem. *Am Entomol* **52**: 90-97.
- Sarvasova A, Goffredo M, Sopoliga I, Savini G, Kocisova A (2014). *Culicoides* midges (diptera: ceratopogonidae) as vectors of orbiviruses in Slovakia. *Vet Ital* **50**(3): 203-212.
- Savini G, MacLachlan NJ, Sanchez-Vizcaino JM, Zientara S (2008). Vaccines against bluetongue in Europe. *Comp Immunol Microbiol Infect Dis* **31**(2-3): 101-120.
- Schmitt T, Varga Z (2012). Extra-Mediterranean refugia: The rule and not the exception? *Front Zool* **9**(1): 22.
- Schulz C, Breard E, Sailleau C, Jenckel M, Viarouge C, Vitour D *et al* (2016). Bluetongue virus serotype 27: detection and characterization of two novel variants in Corsica, France. *J Gen Virol* **97**(9): 2073-2083.
- Schwartz-Cornil I, Mertens PP, Contreras V, Hemati B, Pascale F, Breard E *et al* (2008). Bluetongue virus: virology, pathogenesis and immunity. *Vet Res* **39**(5): 46.
- Schwenkenbecher JM, Mordue AJ, Piertney SB (2009). Phylogenetic analysis indicates that *Culicoides dewulfi* should not be considered part of the *Culicoides obsoletus* complex. *Bull Entomol Res* **99**(4): 371-375.
- Searle KR, Barber J, Stubbins F, Labuschagne K, Carpenter S, Butler A *et al* (2014). Environmental drivers of *Culicoides* phenology: how important is species-specific variation when determining disease policy? *PLoS One* **9**(11): e111876.
- Sedda L, Brown HE, Purse BV, Burgin L, Gloster J, Rogers DJ (2012). A new algorithm quantifies the roles of wind and midge flight activity in the bluetongue epizootic in northwest Europe. *Proc Biol Sci* **279**(1737): 2354-2362.
- Sedda L, Morley D, Brown HE (2015). Characteristics of Wind-Infective Farms of the 2006 Bluetongue Serotype 8 Epidemic in Northern Europe. *EcoHealth* **12**(3): 461-467.

- Sedda L, Rogers DJ (2013). The influence of the wind in the Schmallenberg virus outbreak in Europe. *Sci Rep* **3**: 3361.
- Segelbacher G, Cushman SA, Epperson BK, Fortin MJ, Francois O, Hardy OJ *et al* (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conserv Genet* **11**(2): 375-385.
- Sellers RF (1980). Weather, host and vector--their interplay in the spread of insect-borne animal virus diseases. *J Hyg (Lond)* **85**(1): 65-102.
- Sellers RF, Maarouf AR (1993). Weather factors in the prediction of western equine encephalitis epidemics in Manitoba. *Epidemiol Infect* **111**(2): 373-390.
- Sellers RF, Pedgley DE, Tucker MR (1982). Rift Valley fever, Egypt 1977: disease spread by windborne insect vectors? *Vet Rec* **110**(4): 73-77.
- Semenza JC, Suk JE (2018). Vector-borne diseases and climate change: a European perspective. *FEMS Microbiol Lett* **365**(2): fnx244.
- Semenza JC, Tran A, Espinosa L, Sudre B, Domanovic D, Paz S (2016). Climate change projections of West Nile virus infections in Europe: implications for blood safety practices. *Environ Health* **15 Suppl 1**(1): 28.
- Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009). Evolution and Ecology of Species Range Limits. *Annual Review of Ecology Evolution and Systematics* **40**(1): 415-436.
- Sghaier S, Hammami S, Goffredo M, Hammami M, Portanti O, Lorusso A *et al* (2017). New species of the genus *Culicoides* (Diptera Ceratopogonidae) for Tunisia, with detection of Bluetongue viruses in vectors. *Vet Ital* **53**(4): 357-366.
- Shao ZY, Mao HX, Fu WJ, Ono M, Wang DS, Bonizzoni M *et al* (2004). Genetic structure of Asian populations of *Bombus ignitus* (Hymenoptera: Apidae). *J Hered* **95**(1): 46-52.
- She JX, Autem M, Kotulas G, Pasteur N, Bonhomme F (1987). Multivariate-Analysis of Genetic Exchanges between *Solea-Aegyptiaca* and *Solea-Senegalensis* (Teleosts, Soleidae). *Biol J Linn Soc* **32**(4): 357-371.
- Shirk AJ, Landguth EL, Cushman SA (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Mol Ecol Resour* **17**(6): 1308-1317.
- Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI (2010). Inferring landscape effects on gene flow: a new model selection framework. *Mol Ecol* **19**(17): 3603-3619.
- Shults P, Ho A, Martin EM, McGregor BL, Vargo EL (2020). Genetic Diversity of *Culicoides stellifer* (Diptera: Ceratopogonidae) in the Southeastern United States Compared With Sequences From Ontario, Canada. *J Med Entomol* **57**(4): 1324-1327.
- Sick F, Beer M, Kampen H, Wernike K (2019). *Culicoides* Biting Midges-Underestimated Vectors for Arboviruses of Public Health and Veterinary Importance. *Viruses* **11**(4).

- Simo G, Njiokou F, Tume C, Lueong S, De Meeus T, Cuny G *et al* (2010). Population genetic structure of Central African *Trypanosoma brucei gambiense* isolates using microsatellite DNA markers. *Infect Genet Evol* **10**(1): 68-76.
- Simpson SD, Harrison HB, Claereboudt MR, Planes S (2014). Long-distance dispersal via ocean currents connects Omani clownfish populations throughout entire species range. *PLoS One* **9**(9): e107610.
- Siozios S, Pilgrim J, Darby AC, Baylis M, Hurst GDD (2019). The draft genome of strain cCpun from biting midges confirms insect *Cardinium* are not a monophyletic group and reveals a novel gene family expansion in a symbiont. *PeerJ* **7**: e6448.
- Slade NA, Blair SM (2000). An empirical test of using counts of individuals captured as indices of population size. *J Mammal* **81**(4): 1035-1045.
- Slarkin M (1985). Gene Flow in Natural Populations. *Annual Review of Ecology and Systematics* **16**(1): 393-430.
- Slatkin M (1993). Isolation by Distance in Equilibrium and Non-Equilibrium Populations. *Evolution* **47**(1): 264-279.
- Smouse PE, Long JC, Sokal RR (1986). Multiple-Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence. *Syst Zool* **35**(4): 627-632.
- Sokal RR, Jacquez GM, Oden NL, DiGiovanni D, Falsetti AB, McGee E *et al* (1993). Genetic relationships of European populations reflect their ethnohistorical affinities. *Am J Phys Anthropol* **91**(1): 55-70.
- Solano P, Kaba D, Ravel S, Dyer NA, Sall B, Vreysen MJ *et al* (2010). Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. *PLoS Negl Trop Dis* **4**(5): e692.
- Solano P, Ravel S, Bouyer J, Camara M, Kagbadouno MS, Dyer N *et al* (2009). The population structure of *Glossina palpalis gambiensis* from island and continental locations in Coastal Guinea. *PLoS Negl Trop Dis* **3**(3): e392.
- Solecki AM, Skevington JH, Buddle CM, Wheeler TA (2019). Phylogeography of higher Diptera in glacial and postglacial grasslands in western North America. *BMC Ecol* **19**(1): 53.
- Sommer RS, Zachos FE (2009). Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. *J Biogeogr* **36**(11): 2013-2020.
- Sork VL (2016). Gene flow and natural selection shape spatial patterns of genes in tree populations: implications for evolutionary processes and applications. *Evol Appl* **9**(1): 291-310.
- Spear ER, Coley PD, Kursar TA, Thrall P (2015a). Do pathogens limit the distributions of tropical trees across a rainfall gradient? *J Ecol* **103**(1): 165-174.

- Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Mol Ecol* **19**(17): 3576-3591.
- Spear SF, Cushman SA, McRae BH (2015b). Resistance Surface Modeling in Landscape Genetics *Landscape Genetics*, pp 129-148.
- Sperlova A, Zendulkova D (2011). Bluetongue: a review. *Vet Med (Praha)* **56**(9): 430-452.
- Sprygin AV, Fiodorova OA, Babin YY, Elatkin NP, Mathieu B, England ME *et al* (2014). *Culicoides* biting midges (Diptera, Ceratopogonidae) in various climatic zones of Russia and adjacent lands. *J Vector Ecol* **39**(2): 306-315.
- Stewart JR, Lister AM (2001). Cryptic northern refugia and the origins of the modern biota. *Trends Ecol Evol* **16**(11): 608-613.
- Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF *et al* (2007). Putting the "landscape" in landscape genetics. *Heredity (Edinb)* **98**(3): 128-142.
- Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010). Landscape genetics: where are we now? *Mol Ecol* **19**(17): 3496-3514.
- Stouthamer R, Breeuwer JA, Hurst GD (1999). *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. *Annu Rev Microbiol* **53**: 71-102.
- Sudakaran S, Kost C, Kaltenpoth M (2017). Symbiont Acquisition and Replacement as a Source of Ecological Innovation. *Trends Microbiol* **25**(5): 375-390.
- Suk JE, Semenza JC (2011). Future infectious disease threats to Europe. *Am J Public Health* **101**(11): 2068-2079.
- Sumner T, Orton RJ, Green DM, Kao RR, Gubbins S (2017). Quantifying the roles of host movement and vector dispersal in the transmission of vector-borne diseases of livestock. *PLoS Comput Biol* **13**(4): e1005470.
- Sun EC, Huang LP, Xu QY, Wang HX, Xue XM, Lu P *et al* (2016). Emergence of a Novel Bluetongue Virus Serotype, China 2014. *Transbound Emerg Dis* **63**(6): 585-589.
- Tabachnick WJ (2004). *Culicoides* and the global epidemiology of bluetongue virus infection. *Vet Ital* **40**(3): 144-150.
- Taberlet P, Cheddadi R (2002). Quaternary refugia and persistence of biodiversity. *Science* **297**.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* **7**(4): 453-464.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**(3): 585-595.

- Takezaki N, Nei M (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**(1): 389-399.
- Takken W, Verhulst N, Scholte EJ, Jacobs F, Jongema Y, van Lammeren R (2008). The phenology and population dynamics of *Culicoides* spp. in different ecosystems in The Netherlands. *Prev Vet Med* **87**(1-2): 41-54.
- Talavera G, Dinca V, Vila R (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods Ecol Evol* **4**(12): 1101-1110.
- Talavera S, Munoz-Munoz F, Duran M, Verdun M, Soler-Membrives A, Oleaga A *et al* (2015). *Culicoides* Species Communities Associated with Wild Ruminant Ecosystems in Spain: Tracking the Way to Determine Potential Bridge Vectors for Arboviruses. *PLoS One* **10**(10): e0141667.
- Talavera S, Munoz-Munoz F, Verdun M, Pages N (2017). Morphology and DNA barcoding reveal three species in one: description of *Culicoides cryptipulicaris* sp. nov. and *Culicoides quasipulicaris* sp. nov. in the subgenus *Culicoides*. *Med Vet Entomol* **31**(2): 178-191.
- Talbot B, Garant D, Rioux Paquette S, Mainguy J, Pelletier F (2012). Lack of genetic structure and female-specific effect of dispersal barriers in a rabies vector, the striped skunk (*Mephitis mephitis*). *PLoS One* **7**(11): e49736.
- Tatem AJ, Hay SI, Rogers DJ (2006). Global traffic and disease vector dispersal. *Proc Natl Acad Sci U S A* **103**(16): 6242-6247.
- Taylor PD, Fahrig L, Henein K, Merriam G (1993). Connectivity Is a Vital Element of Landscape Structure. *Oikos* **68**(3): 571.
- Temizel EM, Yesilbag K, Batten C, Senturk S, Maan NS, Mertens PPC *et al* (2009). Epizootic hemorrhagic disease in cattle, Western Turkey. *Emerg Infect Dis* **15**(2): 317-319.
- Tibayrenc M (1998). Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. *Int J Parasitol* **28**(1): 85-104.
- Tibayrenc M, Ayala FJ (2002). The clonal theory of parasitic protozoa: 12 years on. *Trends Parasitol* **18**(9): 405-410.
- Torina A, Caracappa S, Mellor PS, Baylis M, Purse BV (2004). Spatial distribution of bluetongue virus and its *Culicoides* vectors in Sicily. *Med Vet Entomol* **18**(2): 81-89.
- Tsutsui T, Hayama Y, Yamakawa M, Shirafuji H, Yanase T (2011). Flight behavior of adult *Culicoides oxystoma* and *Culicoides maculatus* under different temperatures in the laboratory. *Parasitol Res* **108**(6): 1575-1578.
- Turner MG (2005). Landscape Ecology: What Is the State of the Science? *Annu Rev Ecol Evol Syst* **36**(1): 319-344.

- Tzedakis PC, Emerson BC, Hewitt GM (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends Ecol Evol* **28**(12): 696-704.
- Uezu A, Metzger JP, Vielliard JME (2005). Effects of structural and functional connectivity and patch size on the abundance of seven Atlantic Forest bird species. *Biol Conserv* **123**(4): 507-519.
- Valade R, Kenis M, Hernandez-Lopez A, Augustin S, Mari Mena N, Magnoux E *et al* (2009). Mitochondrial and microsatellite DNA markers reveal a Balkan origin for the highly invasive horse-chestnut leaf miner *Cameraria ohridella* (Lepidoptera, Gracillariidae). *Mol Ecol* **18**(16): 3458-3470.
- Van Strien MJ, Keller D, Holderegger R (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* **21**(16): 4010-4023.
- Venail R (2014). Sensibilité aux insecticides et évaluation préliminaire des méthodes de lutte antivectorielle disponibles contre les *Culicoides* (Diptera : Ceratopogonidae) paléarctiques, vecteurs de virus émergents d'intérêt en santé animale. Université De Montpellier II Sciences Et Techniques Du Languedoc.
- Venail R, Balenghien T, Guis H, Tran A, Setier-Rio ML, Delécolle JC *et al* (2012). Assessing Diversity and Abundance of Vector Populations at a National Scale: Example of *Culicoides* Surveillance in France After Bluetongue Virus Emergence *Arthropods as Vectors of Emerging Diseases*. Vol. 3, pp 77-102.
- Venturi G, Di Luca M, Fortuna C, Remoli ME, Riccardo F, Severini F *et al* (2017). Detection of a chikungunya outbreak in Central Italy, August to September 2017. *Euro Surveill* **22**(39): 17-00646.
- Veronesi E, Henstock M, Gubbins S, Batten C, Manley R, Barber J *et al* (2013). Implicating *Culicoides* biting midges as vectors of Schmallenberg virus using semi-quantitative RT-PCR. *PLoS One* **8**(3): e57747.
- Viennet E, Garros C, Gardes L, Rakotoarivony I, Allene X, Lancelot R *et al* (2013). Host preferences of Palaearctic *Culicoides* biting midges: implications for transmission of orbiviruses. *Med Vet Entomol* **27**(3): 255-266.
- Vignal A, Milan D, SanCristobal M, Eggen A (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol* **34**(3): 275-305.
- Votypka J, Synek P, Svobodova M (2009). Endophagy of biting midges attacking cavity-nesting birds. *Med Vet Entomol* **23**(3): 277-280.
- Vreysen MJ, Balenghien T, Saleh KM, Maiga S, Koudougou Z, Cecchi G *et al* (2013). Release-recapture studies confirm dispersal of *Glossina palpalis gambiensis* between river basins in Mali. *PLoS Negl Trop Dis* **7**(4): e2022.
- Wagner HH, Fortin MJ (2005). Spatial analysis of landscapes: Concepts and statistics. *Ecology* **86**(8): 1975-1987.

- Wagner HH, Fortin MJ (2015). Basics of Spatial Data Analysis: Linking Landscape and Genetic Data for Landscape Genetic Studies. *Landscape Genetics: Concepts, Methods, Applications*: 77-98.
- Walker AR (1977). Seasonal Fluctuations of *Culicoides* Species (Diptera-Ceratopogonidae) in Kenya. *Bull Entomol Res* **67**(2): 217-233.
- Walton TE (2004). The history of bluetongue and a current global overview. *Vet Ital* **40**(3): 31-38.
- Wan ZM (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens Environ* **140**: 36-45.
- Wang IJ (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution* **67**(12): 3403-3411.
- Waples RS, Do C (2008). Ldne: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* **8**(4): 753-756.
- Waples RS, Gaggiotti O (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* **15**(6): 1419-1439.
- Watson TM, Saul A, Kay BH (2000). *Aedes notoscriptus* (Diptera: Culicidae) survival and dispersal estimated by mark-release-recapture in Brisbane, Queensland, Australia. *J Med Entomol* **37**(3): 380-384.
- Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol Ecol* **7**(11): 1569-1573.
- Weir BS, Cockerham CC (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38**(6): 1358-1370.
- Weir BS, Goudet J (2017). A Unified Characterization of Population Structure and Relatedness. *Genetics* **206**(4): 2085-2103.
- Wenk CE, Kaufmann C, Schaffner F, Mathis A (2012). Molecular characterization of Swiss Ceratopogonidae (Diptera) and evaluation of real-time PCR assays for the identification of *Culicoides* biting midges. *Vet Parasitol* **184**(2-4): 258-266.
- Werner D, Groschupp S, Bauer C, Kampen H (2020). Breeding Habitat Preferences of Major *Culicoides* Species (Diptera: Ceratopogonidae) in Germany. *Int J Environ Res Public Health* **17**(14).
- Werren JH, Baldo L, Clark ME (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol* **6**(10): 741-751.

- Westbrook JK, Isard SA (1999). Atmospheric scales of biotic dispersal. *Agricultural and Forest Meteorology* **97**(4): 263-274.
- White BP, Pilgrim EM, Boykin LM, Stein ED, Mazor RD (2014). Comparison of four species-delimitation methods applied to a DNA barcode data set of insect larvae for use in routine bioassessment. *Freshwater Science* **33**(1): 338-348.
- White TA, Perkins SE, Heckel G, Searle JB (2013). Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. *Mol Ecol* **22**(11): 2971-2985.
- Whittaker RJ, Field R (2000). Tree species richness modelling: an approach of global applicability? *Oikos* **89**(2): 399-402.
- Widmer A, Lexer C (2001). Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends Ecol Evol* **16**(6).
- Widmer A, Schmid-Hempel P, Estoup A, Scholl A (1998). Population genetic structure and colonization history of *Bombus terrestris* s.l. (Hymenoptera : Apidae) from the Canary Islands and Madeira. *Heredity* **81**: 563-572.
- Wiens JA, Stenseth NC, Vanhorne B, Ims RA (1993). Ecological Mechanisms and Landscape Ecology. *Oikos* **66**(3): 369-380.
- Wilson AJ, Mellor PS (2009). Bluetongue in Europe: past, present and future. *Philos Trans R Soc Lond B Biol Sci* **364**(1530): 2669-2681.
- Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**(3): 1177-1191.
- Winter DJ (2012). MMOD: an R library for the calculation of population differentiation statistics. *Mol Ecol Resour* **12**(6): 1158-1160.
- Wittmann EJ, Mello PS, Baylis M (2002). Effect of temperature on the transmission of orbiviruses by the biting midge, *Culicoides sonorensis*. *Med Vet Entomol* **16**(2): 147-156.
- Wold S, Esbensen K, Geladi P (1987). Principal Component Analysis. *Chemometrics Intellig Lab Syst* **2**(1-3): 37-52.
- Wotton KR, Gao B, Menz MHM, Morris RKA, Ball SG, Lim KS *et al* (2019). Mass Seasonal Migrations of Hoverflies Provide Extensive Pollination and Crop Protection Services. *Curr Biol* **29**(13): 2167-2173 e2165.
- Wright IM (2014). Serological and genetic characterisation of putative new serotypes of bluetongue virus and epizootic haemorrhagic disease virus isolated from an alpaca MSc Thesis thesis, University Potchefstroom.
- Wright S (1931). Evolution in Mendelian Populations. *Genetics* **16**(2): 97-159.

- Wright S (1943). Isolation by Distance. *Genetics* **28**(2): 114-138.
- Yang RC (1998). Estimating Hierarchical F-Statistics. *Evolution* **52**(4): 950-956.
- Yang ZH (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology* **61**(5): 854-865.
- Yeruham I, Perl S, Braverman Y (2004). Seasonal allergic dermatitis in sheep associated with *Ctenocephalides* and *Culicoides* bites. *Vet Dermatol* **15**(6): 377-380.
- Young SG, Carrel M, Kitchen A, Malanson GP, Tamerius J, Ali M *et al* (2017). How's the Flu Getting Through? Landscape genetics suggests both humans and birds spread H5N1 in Egypt. *Infect Genet Evol* **49**: 293-299.
- Zeller KA, McGarigal K, Whiteley AR (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology* **27**(6): 777-797.
- Zhang DX, Hewitt GM (1997). Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology* **25**(2): 99-120.
- Zhang DX, Hewitt GM (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Mol Ecol* **12**(3): 563-584.
- Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29**(22): 2869-2876.
- Zientara S, Sailleau C, Viarouge C, Hoper D, Beer M, Jenckel M *et al* (2014). Novel bluetongue virus in goats, Corsica, France, 2014. *Emerg Infect Dis* **20**(12): 2123-2125.
- Zientara S, Sánchez-Vizcaíno JM. (2013). *Veterinary Microbiology, Vol. 165*.
- Zimmer J-y, Haubruge É, Francis F (2014a). Synthèse bibliographique: l'écologie larvaire des *Culicoides* (Diptera: Ceratopogonidae). *Biotechnol Agron Soc Environ* **18**: 301-312.
- Zimmer JY, Brostaux Y, Haubruge E, Francis F (2014b). Larval development sites of the main *Culicoides* species (Diptera: Ceratopogonidae) in northern Europe and distribution of coprophilic species larvae in Belgian pastures. *Vet Parasitol* **205**(3-4): 676-686.

Résumé

Depuis 2006, des épizooties importantes de fièvre catarrhale ovine et de virus (FCO) de Schmallenberg (SB) sont observées dans la région ouest paléarctique, ces deux virus étant transmis par des moucheron hémato-phages autochtones du genre *Culicoides* (Diptera : Ceratopogonidae). Ces virus affectent les bovins, les ovins, les caprins et les ruminants sauvages. La dispersion des *Culicoides* correspond à ce que les écologues appellent dispersion stratifiée ou « *stratified dispersal* ». Elle résulte de la combinaison de dispersion active de proche en proche, sur de courtes distances, et de sauts sur de longues distances via le vent de manière passive ou semi-passive, permettant des sauts de dispersion de plusieurs centaines de kilomètres. *Culicoides obsoletus* est une espèce vectrice importante pour les virus de la FCO et de SB et abondante en région paléarctique. Si la dispersion des *Culicoides* par le vent au-dessus des masses d'eaux est bien établie, il est aussi important d'évaluer la dispersion au-dessus des terres et de déterminer l'influence des facteurs biotiques et abiotiques influant sur celle-ci. L'objectif de ce travail de thèse est de déterminer par l'observation des patrons de distribution des fréquences alléliques comment la dispersion a pu façonner l'arrangement spatial de la diversité génétique et comment elle est influencée par les paysages et par les caractéristiques des habitats qui peuvent favoriser ou limiter les mouvements des *Culicoides*. Pour cela, deux plans d'échantillonnage ont été utilisés : un échantillonnage populationnel à l'échelle de l'Europe (de la Scandinavie aux îles Canaries, à la Turquie), et un échantillonnage individuel à l'échelle française. L'utilisation complémentaire de marqueurs mitochondriaux et nucléaires à ces différentes échelles spatiales nous ont permis (i) de clarifier la distribution et de réviser le schéma taxonomique du Complexe Obsoletus/Scoticus, (ii) de mettre en évidence des flux de gènes important permettant d'homogénéiser la structure génétique des populations à l'échelle d'un pays et d'un continent, (iii) de quantifier l'impact d'un large éventail de facteurs environnementaux et météorologiques sur la dispersion des *Culicoides* et de mettre en évidence l'importance de la densité d'hôtes, (iv) et de proposer un scénario présentant l'histoire évolutive récente des populations européennes de *C. obsoletus*. L'hypothèse retenue est une recolonisation post-glaciaire récente depuis le refuge glaciaire des Balkans et un refuge plus au nord-est. Ces populations auraient recolonisé ensemble l'ouest de l'Europe, engendrant une admixture des populations. Nos résultats démontrent une capacité de dispersion de *C. obsoletus* à l'intérieur des terres très importante, qui doit être prise en compte pour les travaux à venir sur la compétence des vecteurs et la modélisation épidémiologique de la transmission des virus transmis.

Mots-clés : *Culicoides*, dispersion, génétique des populations, microsatellites, gènes mitochondriaux.

Emerging diseases and animal health: genetic diversity and dispersal of a biting midge species vector of bluetongue and Schmallenberg viruses in the Palearctic region, *Culicoides obsoletus* (Diptera: Ceratopogonidae)

Since 2006, important outbreaks of bluetongue virus (BTV) and Schmallenberg virus (SBV) have been reported in the West Palearctic region. Both viruses are transmitted by hematophagous biting midges of the genus *Culicoides* (Diptera: Ceratopogonidae). These viruses affect wild cattle, sheep, goats and wild ruminants. The dispersion of *Culicoides* corresponds to what ecologists call stratified dispersal. It results from the combination of active dispersal from near to near, over short distances, and jumps over long distances via the wind in a passive or semi-passive way, allowing dispersal jumps of several hundred kilometers. *Culicoides obsoletus* is a recognized vector species for BTV and SBV, abundant in the Palearctic region. While the dispersal of *Culicoides* by wind over water bodies is well established, it is also important to assess the dispersion over land and determine the influence of biotic and abiotic factors influencing it. The objective of this thesis is to determine by observing the pattern of distributions of allelic frequencies how dispersion may have shaped the spatial arrangement of genetic diversity and how it is influenced by landscapes and habitat characteristics that can promote or limit the movements of *Culicoides*. Two sampling strategies were used: an Europe-wide population sampling (from Scandinavia to the Canary Islands, to Turkey), and an individual sampling at the French level. The complementary use of mitochondrial and nuclear markers at these different spatial scales has allowed us (i) to clarify the distribution and revise the taxonomic scheme of the Obsoletus/Scoticus Complex, (ii) to highlight important gene flows that homogenize the genetic structure of populations at the country and continent level, (iii) to quantify the impact of a wide range of environmental and meteorological factors on *Culicoides* dispersal and highlight the importance of host density, (iv) and to propose a scenario for the recent evolutionary history of European populations of *C. obsoletus*. The hypothesis is a recent post-glacial recolonization from the Balkan ice refuge and a refuge further northeast. These populations would have recolonized together western Europe, resulting in admixture of populations. Our results demonstrate a very important inland dispersal capability of *C. obsoletus*, which must be considered for future works on vector competence and epidemiological modelling of transmitted virus transmission.

Key-words : *Culicoides*, dispersal, population genetics, microsatellites, mitochondrial genes.