



HAL
open science

Evolution of genetic and phenotypic variation during an invasion : the case of *Drosophila Suzukii*

Antoine Fraimout

► **To cite this version:**

Antoine Fraimout. Evolution of genetic and phenotypic variation during an invasion : the case of *Drosophila Suzukii*. Systematics, Phylogenetics and taxonomy. Museum national d'histoire naturelle - MNHN PARIS, 2016. English. NNT : 2016MNHN0019 . tel-04785901

HAL Id: tel-04785901

<https://theses.hal.science/tel-04785901v1>

Submitted on 15 Nov 2024

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.

MUSEUM NATIONAL



D'HISTOIRE NATURELLE

École Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2016

N°attribué par la bibliothèque

□□□□□□□□□□

THESE

Pour obtenir le grade de

DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : « Biologie Évolutive »
Présentée et soutenue publiquement par

Antoine Fraimout

Le 9 décembre 2016

Évolution de la variation génétique et phénotypique au cours d'une invasion : le cas de *Drosophila suzukii*

Sous la direction de : M. Vincent Debat, Maître de Conférences du Muséum et de Mlle
Céline Teplitsky, Chargée de Recherche au CNRS

JURY :

Mme Porcher, Emmanuelle
M. Debat, Vincent
Mlle Teplitsky, Céline
M. Claude, Julien
M. Guillemaud, Thomas
M. Merilä, Juha

Professeur, Muséum national d'histoire naturelle, Paris (France)
Docteur, Muséum national d'histoire naturelle, Paris (France)
Docteur, Chargée de recherches au CNRS, Montpellier (France)
Docteur, Université Montpellier II, Montpellier (France)
Docteur, INRA, Nice (France)
Professeur, University of Helsinki (Finlande)

Présidente
Directeur de Thèse
Directeur de Thèse
Rapporteur
Rapporteur
Examineur

Remerciements

Comment se souvenir de toutes les personnes qui ont, de près ou de loin, apporté leur aide, leur soutien, ou leur participation à cette thèse? Peut être qu'il faut simplement reprendre le fil de l'histoire depuis le début.

Merci à mes directeurs. Vincent et Céline, vous qui m'avez fait confiance pour ce projet et qui m'avez accompagné tout au long de son déroulement. Je me souviendrai toujours de cette première rencontre au Muséum pour l'entretien qui allait déterminer si oui, ou non j'allais pouvoir passer 3 ans à travailler avec vous, et me dire en sortant «pas sur que ça ce soit bien passé...dommage ils avaient l'air vraiment sympas ». Faut croire que j'avais tort, et tant mieux.

Vincent **merci** de m'avoir fait faire le tour du monde (ça sonne tellement romantique!). Ces souvenirs de terrains resteront dans ma mémoire pour un bon bout de temps. Comment on dit **merci** en Japonais déjà..itadakimas... ? Mais surtout **merci** de m'avoir transmis ta science avec autant de passion, **merci** de m'avoir fait rencontrer autant de grands chercheurs, et chercheuses, et **merci** de m'avoir conforter dans l'idée de continuer ce métier.

Céline de m'avoir introduit au monde merveilleux de la génétique quantitative ! Maintenant je peux l'avouer mais je n'y croyais pas. Ça me paraissait être un épreuve insurmontable que de se lancer dans une discipline aussi riche et aussi complexe, comme ça d'un coup, et de n'avoir que 3 ans pour le faire. Et même si je suis encore (très) loin d'avoir fini mon apprentissage, grâce à toi, à ta patience et à ta bonne humeur j'ai l'impression d'avoir franchi un cap.

Merci à toute cette équipe formidable avec qui j'ai travaillé pendant ces trois ans. **Merci** Violaine, Marianne, Mathieu, Florence, Sylvain, Raph, Yann, Nico, Claire, Barbara, Monica. Je ne pourrais pas souhaiter à quelqu'un de meilleures personnes avec qui passer du temps au labo, discuter de papillons, de star wars, de black metal, de café, de tout et de rien...mais surtout de tout.

Merci à Simon grâce à qui je me suis fait embarqué dans la grande aventure Montpelliéraine de suzuki. Mais surtout **merci** pour tous les bons moments sur le terrain (god bless you ! comme dirait quelqu'un qu'on connaît), pour tous les bons moments à discuter, refaire le monde et (essayer) de refaire la science les pieds dans le sable au bord du Pacifique.

Merci à tous les collègues du CBGP : Anne X. et Anne L., Antoine, Raph, Christophe, Julien, **merci** à tous les meilleurs joueurs de volley-ball de la recherche Française, Mathieu, Jeff, Laure, et **merci** Arnaud d'avoir été comme presque-directeur-adoptif pendant tout ce temps passé là bas !

Merci à Patricia et Nicolas d'avoir accepté d'être des membres éminents de mon comité de thèse ! **Merci** pour votre aide pour tout ce travail.

Merci à tous les gens du muséum que j'ai côtoyé pendant ces trois ans : Vinz, Julien, Camille, tous les membres du BDEM, Gilles, Marion, et tous les autres.

Et puis bien sur merci à tous ceux qui ont toujours été là.

Merci à mes amis, mes meilleurs amis : Charlie, Raph, Paul, Keke, Benjouz, Mitou, Belben, et Flobe. Sans vous cette thèse aurait eu une tout autre saveur.

Merci à celle qui a été à mes côtés dans ces derniers moments, pas les plus faciles, et qui m'a supporté, soutenu et encouragé tous les jours, mä rakastan sua Susanna.

Merci à ma famille : Nico, Lulu, Cléo & Arthur, j'ai hâte de vous revoir vite.

Merci à mes parents : je leur dédie cette thèse.

Table des matières

Introduction générale.....	3
I – Variations et réponse(s) à la sélection.....	4
I – 1. Sources de variation.....	4
I – 2. Réponse à la sélection.....	5
I – 3. Réponse univariée.....	5
I – 4. Réponse multivariée.....	6
I – 5. Plasticité phénotypique	12
I – 6. Conclusion sur la partie I.....	13
II – Invasions biologiques : introduction et considérations évolutives.....	14
II – 1. Introduction	14
II – 2. Les invasions biologiques comme laboratoires naturels.....	18
II – 3. Histoire de l'invasion.....	18
II – 3 – a. Goulots d'étranglements et dérive génétique.....	18
II – 4. Potentiel adaptatif et succès invasif.....	21
II – 4 – a. Variation génétique et succès invasif.....	21
II – 4 – b. Plasticité phénotypique et succès invasif.....	22
II – 5. Conséquences sur les trajectoires évolutives.....	23
II – 5 – a. Attendus sur la matrice G.....	23
II – 5 – b. Attendus sur la divergence phénotypique.....	24
II – 6. Conclusion sur la partie II.....	24
III – Modèle d'étude : La Drosophile à ailes tachetées, <i>Drosophila suzukii</i>	24
IV – Grands axes de la thèse.....	28
IV – 1. Echantillonnage et maintien des populations.....	28
IV – 2. Histoire de l'invasion de <i>D. suzukii</i>	28
IV – 3. Evolution de la matrice G en contexte invasif.....	30
IV – 4. Plasticité phénotypique et performances de vol en réponse à la température de développement.....	32
IV – 5. Evolution de la niche climatique de <i>D. suzukii</i>	33
Chapitre 1 : Génétique des populations et histoire de l'invasion de <i>D. suzukii</i>	35
Article 1.	35
Fraitout <i>et al.</i> (2015). New set of microsatellite markers for the spotted-wing <i>Drosophila suzukii</i> (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest. <i>Eur. J. of Entomol.</i>	35
Article 2.	41
Fraitout <i>et al.</i> Deciphering the invasion routes of <i>Drosophila suzukii</i> by means of ABC random forest. <i>Mol. Biol. Evol.</i> submitted.....	41

Chapitre 2 : Génétique quantitative, évolution de la matrice G et divergence phénotypique.....	78
Article 3.	78
Fraimout <i>et al.</i> Quantitative genetics of <i>Drosophila suzukii</i> wing shape: early phases of divergence during a worldwide invasion. in prep.....	78
Chapitre 3 : Plasticité phénotypique et performance de vol chez <i>Drosophila suzukii</i>	104
Article 4.	104
Fraimout <i>et al.</i> Phenotypic plasticity of <i>Drosophila suzukii</i> wing to developmental temperature: implications for flight performance. in prep.....	104
Chapitre 4 : Modélisation de la niche climatique de <i>Drosophila suzukii</i>	125
Article 5.	125
Fraimout, A., Monnet, A.C. Global invasion of the spotted-wing <i>Drosophila suzukii</i> : new insights from climatic niche modeling. in prep.....	125
Conclusion générale.....	148
Résultats principaux et perspectives.....	148
Apports théoriques et méthodologiques de la thèse.....	150
Chapitre 1 : génétique des populations.....	150
Chapitre 2 : génétique quantitative.....	151
Chapitre 3 : plasticité phénotypique.....	151
Chapitre 4 : modélisation de niche.....	152
Conclusions personnelles.....	152
Références bibliographiques.....	153
Annexes.....	176
Annexe du Chapitre 1: Supplementary material for the manuscript „Deciphering the routes of invasion of <i>Drosophila suzukii</i> by means of ABC-Random Forest“.....	176
Annexe du Chapitre 3: Supplementary material for the manuscript „Phenotypic plasticity of <i>Drosophila suzukii</i> wing to developmental temperature: implications for flight performance“...	176
Article en collaboration : <i>Wolbachia</i> in European Populations of the Invasive Pest <i>Drosophila suzukii</i> : Regional Variation in Infection Frequencies.....	176

Introduction générale

"Aucun individu ne ressemble parfaitement à un autre, aucune espèce n'existe sans un grand nombre de variétés". Par cette simple phrase, Georges-Louis Leclerc de Buffon dans son Histoire Naturelle (1749), souligne une réalité naturelle fondamentale: la variation dans les systèmes biologiques est omniprésente, et chaque niveau d'organisation du vivant est caractérisé par une diversité des nombres, des formes, des couleurs et des structures. A l'échelle des espèces, la variation morphologique est depuis longtemps utilisée pour classer le vivant, et les travaux historiques de systématique du XVIII^e siècle (e.g. Linné 1758) se basaient sur une approche typologique¹ de l'espèce mais n'accordaient que peu d'importance à la variabilité entre les individus d'un même groupe (Bowler 2005), considérant cette variation comme une véritable nuisance à l'exercice de classification (e.g. Mayr 1959). L'importance de cette variabilité intra-spécifique fut mise au premier plan de la biologie évolutive par Darwin, qui en proposa une réinterprétation capitale, en particulier dans *L'origine des espèces* (1859) et dans son ouvrage *De la variation des animaux et des plantes sous l'action de la domestication* (1868), lui donnant un rôle central dans sa théorie de l'évolution par sélection naturelle. Selon lui, à la base de l'évolution se trouve la lutte entre les variants individuels d'une même espèce pour l'accès aux ressources de l'environnement. Cette lutte (*the struggle for existence*) résultera en la survie et la reproduction de certains individus au détriment d'autres, qui transmettront alors leurs caractères aux générations suivantes, augmentant ainsi la fréquence de ces caractères dans l'espèce (Endler 1986). Le matériel de base de l'évolution est donc la variation : sans variation il ne peut pas y avoir de sélection (Lewontin 1970). Le modèle d'évolution Darwinienne place donc la variation comme matériau de base à la sélection naturelle.

Au sein des espèces (i.e. à l'échelle des populations) plusieurs questions se posent sur l'origine de cette variation et son rôle dans l'évolution. Plus particulièrement, comment l'interaction entre mécanismes évolutifs – i.e mutation, sélection, migration et dérive (Bell 2008) – et variation individuelle influence-t-elle les directions de divergences entre populations ? L'évolution se fait-elle dans toutes les directions de variation possibles ? Certaines sont-elles plus souvent utilisées que d'autres?

Qu'ils soient phénotypiques ou génétiques, adaptatifs ou neutres au regard de l'évolution, les patrons de variation intra-spécifique sont à la fois le produit de l'évolution passée et le matériau de base de l'évolution future. Leur analyse fine est nécessaire pour répondre à ces questions.

¹ L'unité de classification (ici l'espèce) est fondée sur la description d'un phénotype unique et représentatif de l'ensemble du groupe à l'intérieur des limites définies par le descripteur, et servant de référence objective à la communauté (ICZN International code of zoological nomenclature).

I – Variations et réponse(s) à la sélection

I – 1. Sources de variation

Le phénotype, qui inclut morphologie, physiologie et comportement, est le produit de l'interaction constante entre les gènes et l'environnement de chaque individu (West-Eberhard 2003). La variation phénotypique au sein d'une population peut donc résulter d'effets génétiques, si par exemple les individus expriment des phénotypes différents parce qu'ils ne possèdent pas le(s) même(s) allèle(s) (la variation de couleur des yeux chez l'Homme est un exemple classique où le phénotype dépend directement du génotype, Duffy *et al.* 2007). La variation phénotypique peut aussi provenir d'effets environnementaux, si les conditions du milieu sont différentes entre individus (e.g. la variation du système immunitaire chez l'Homme est en grande partie dépendante des pathogènes rencontrés au cours de la vie, Brodin *et al.* 2015). On parle alors de plasticité phénotypique (Bradshaw 1965, Schlichting 1986, West-Eberhard 1989, 2003), définie comme la capacité d'un génotype à produire plusieurs phénotypes dans différents environnements (voir par exemple Debat & David (2001) pour une discussion des définitions). Enfin, une part de la variabilité phénotypique peut provenir de l'interaction entre gènes et environnement (e.g. chez *Drosophila melanogaster*, ~15% des transcrits exprimés varient selon l'environnement, Zhou *et al.* 2012 revu dans Des Marais *et al.* 2013), c'est à dire que différents génotypes ont des réponses différentes à l'environnement. Cette partition de la variation phénotypique est résumée dans l'équation (1):

$$V_P = V_G + V_E + V_{G \times E} \quad (1)$$

où la variation phénotypique V_P est la somme de la variation génétique V_G , de la variation environnementale V_E et de leur interaction $V_{G \times E}$. La distribution de cette variation phénotypique dans une population est différente selon les types de traits. Pour les traits Mendéliens qualitatifs (e.g. couleur, présence ou absence du caractère etc.) elle est représentée par la fréquence des différents génotypes (Conner & Hartl 2004). Pour les caractères quantitatifs, qui contrairement aux caractères Mendéliens se mesurent sur une échelle de variation continue, les modèles de génétique quantitative font l'hypothèse que ces traits dépendent d'un très grand nombre de gènes à effets faibles et additifs (i.e. modèle infinitésimal, Fisher 1930, et voir Barton & Turelli 1989) et la variance phénotypique résulte de l'ensemble des effets de ces gènes (Orr 2005). C'est cette variation phénotypique qui passera au crible de la sélection naturelle et l'adaptation peut donc être génétique (évolution de la fréquence des gènes) ou environnementale (plasticité adaptative).

I – 2. Réponse à la sélection

L'équation (1) est à la base de la génétique quantitative, qui tend à identifier la part de cette variation qui sera transmise aux générations suivantes (Falconer 1960). La variation génétique peut être plus finement décomposée en part héritable, la variation additive (V_A), et non héritable, la variation de dominance (interactions entre allèles au sein d'un locus) et d'épistasie (interactions entre allèles entre loci ; respectivement V_D et V_I) (Lynch & Walsh 1998). Cette variation héritable (V_A) est responsable de la ressemblance entre parents et descendants (et globalement entre individus apparentés) et correspond aux effets additifs des gènes sur le phénotype. Estimée à partir de mesures effectuées sur des familles, elle permet de déterminer l'héritabilité d'un caractère donné, c'est à dire la part de variation phénotypique de ce caractère expliquée par la variation génétique additive (Lynch & Walsh 1998). L'héritabilité au sens strict² d'un trait unique est donc :

$$h^2 = V_A/V_P \quad (2)$$

et peut varier entre 0 (absence de variance génétique) et 1 (absence de variance environnementale, de dominance ou d'épistasie) selon le trait d'intérêt. Par exemple, il est généralement considéré que les traits morphologiques sont plus héritables (i.e. possèdent plus de variance additive) que les traits d'histoire de vie (Mousseau & Roff 1987, Price & Schluter 1991, Hoffmann *et al.* 2016, mais voir Houle 1992 pour discussion). L'héritabilité occupe une place centrale dans les modèles prédictifs de l'évolution des caractères quantitatifs.

I – 3. Réponse univariée

L'héritabilité intervient dans l'équation de réponse à la sélection (ou équation du sélectionneur, Lush 1937):

$$R = h^2S$$

où R , le changement moyen du trait d'intérêt dans la population en une génération (i.e. la réponse à la sélection) est égal au produit de son héritabilité h^2 et du différentiel de sélection S , qui correspond à la différence entre la moyenne du trait de la population et celle des parents sélectionnés (et mesure donc l'intensité de la sélection ; Falconer & MacKay 1996). L'héritabilité a de ce fait souvent été utilisée en génétique quantitative pour estimer la capacité de réponse à la sélection d'une population – son potentiel évolutif – pour un trait phénotypique, et prédire la vitesse d'évolution de ce trait (e.g. Prout & Barker 1989). L'importance évolutive et la généralité de son utilisation comme estimateur du potentiel évolutif ont toutefois été très discutées (Houle 1992, Morissey *et al.* 2010), en particulier du fait que l'héritabilité est variable selon les conditions

² En opposition à l'héritabilité au sens large H^2 , telle que $H^2 = V_G/V_P$

environnementales (Hoffmann & Merilä 1999, Charmantier & Garant 2005). Cette utilisation n'est valide que sous certaines conditions et s'avère être plus appropriée pour étudier la sélection dans un cadre expérimental – où la sélection est manipulée et peut être estimée – plutôt que pour prédire l'évolution des populations naturelles (Merilä *et al.* 2001, Morissey *et al.* 2010, et voir Houle 1992 pour discussion).

I – 4. Réponse multivariée

Décrire la sélection comme agissant sur des caractères isolés est une « immense simplification » (Lande & Arnold 1983). Les phénotypes sont en effet le plus souvent composites – constitués d'ensembles de traits dont les corrélations doivent être prises en compte : la réponse à la sélection (i.e. l'adaptation) est donc par définition multivariée (Lande & Arnold 1983, Phillips & Arnold 1989, Roff 2000, Merilä *et al.* 2001, Blows 2007, Walsh & Blows 2009). Ces paramètres (variances et covariances des caractères entre autres) ont été mis en avant par les biologistes de l'évolution des années 1980, et ont permis de formaliser une description mathématique multivariée des éléments de la sélection (Brodie *et al.* 1995). Plus particulièrement, la version multivariée de l'équation du sélectionneur est introduite par Lande en 1979 telle que :

$$R = GP^{-1}S$$

où **G** et **P** correspondent aux matrices de variance-covariance génétique (additive) et phénotypique respectivement (Roff 2000), et **S** au vecteur de différentiel de sélection. La matrice **G** résume l'information sur la quantité de variance génétique additive pour un ensemble de traits d'une population (sur sa diagonale), ainsi que l'intensité et la direction de leurs associations (i.e. leurs covariances, hors diagonale). Au même titre que l'héritabilité, **G** renseigne donc sur la variation génétique disponible pour chacun des traits et leur capacité de réponse à la sélection. Les éléments hors-diagonale de **G** apportent toutefois une information supplémentaire fondamentale: celle des covariances entre traits comme possibles contraintes à la réponse à la sélection.

La covariance entre traits peut être due à une base génétique commune (i.e. pléiotropie, Cheverud 1996) ou à l'association non-aléatoire d'allèles à différents loci (i.e. déséquilibre de liaison, Hartl & Clark 1997). Le signe de ces covariances peut être positif, ou négatif. Par exemple, il existe chez les mâles du Guppy *Poecilia reticulata*, une corrélation génétique négative entre l'attrait (i.e. la propension à plaire aux femelles) et le taux de survie des descendants (Brooks 2000). Inversement, la résistance aux nématodes parasites chez le mouton de Soay (*Ovis aries*) est positivement corrélée à la taille du corps (Coltman *et al.* 2001). La corrélation entre les traits peut avoir un effet sur la réponse à la sélection. En effet, un trait possédant suffisamment de

variation génétique devrait répondre plus facilement à la sélection. Toutefois, s'il existe une corrélation forte entre ce trait, et un ou plusieurs autres, alors l'interaction entre la direction de ces corrélations (i.e. le signe des covariances), et la direction de la sélection pourra faciliter ou contraindre l'évolution (Agrawal & Stinchcombe 2009). Un exemple particulièrement probant de l'influence des corrélations génétiques sur la réponse à la sélection se trouve dans la domestication du renard gris (*Vulpes vulpes*). Chez cette espèce, une forte sélection artificielle pour la docilité a été associée à une réponse corrélée de plusieurs autres traits tels que la couleur du pelage ou le cycle reproducteur (Trut *et al.* 2009). Dans ce cas, la réponse d'un trait sous sélection (ici la docilité) a entraîné une réponse d'un trait corrélé (e.g. la couleur du pelage), sans que la sélection y soit directement appliquée. On retrouve de telles réponses corrélées par exemple dans l'adaptation à un nouvel environnement chez la Salicaire commune (*Lythrum salicaria*). Il a été montré qu'une forte sélection pour un raccourcissement du temps de floraison – liée à l'expansion géographique de l'espèce vers le Nord (Montague *et al.* 2008) – aboutit à un changement des dates de floraison, mais également à une diminution de la taille et de la fécondité (Colautti *et al.* 2010) : les individus à floraison précoce sont plus petits et moins féconds que la moyenne des individus de la population originale (i.e. avant sélection). Cette baisse de fécondité, liée à la corrélation positive entre temps de floraison, taille et fécondité, va donc tendre à s'opposer à la sélection sur le temps de floraison. C'est un exemple clair de corrélation génétique agissant comme une contrainte évolutive.

L'information sur les covariances génétiques entre traits, telle que résumée dans la matrice **G** est donc essentielle à la prédiction de la réponse à la sélection. Graphiquement, la réponse théorique d'une population à la sélection peut se comprendre en représentant le positionnement de la matrice **G** sur un paysage adaptatif (Wright 1932, Arnold *et al.* 2001, Stepan *et al.* 2002) (Fig. 1). Dans ce paysage, la matrice **G** est représentée par une ellipse centrée sur la moyenne du phénotype de la population. Sa géométrie est caractérisée par trois paramètres : sa taille, sa forme (qui définissent sa structure) et son orientation. La taille de **G** (le volume de l'ellipse) correspond à la quantité de variance génétique disponible dans la population, toutes directions confondues. Ainsi, une population présentant beaucoup de variance aura une matrice de grande taille, et pourra, en théorie, répondre rapidement à la sélection. A l'inverse, une population aura plus de difficulté à répondre à la sélection si elle présente une matrice de petite taille. La forme de **G** correspond à la distribution de la variance entre caractères et combinaisons de caractères, combinaisons qui définissent les axes principaux de l'ellipse (on parle aussi de vecteurs propres ou eigen-vecteurs). Si la variance est répartie de façon homogène entre les axes (les traits considérés sont également variables et corrélés), alors la matrice peut être représentée par un cercle (en 2D), ou une sphère (en 3D) ; si au contraire, une des directions contient plus de variance que les autres (certaines combinaisons de traits sont génétiquement plus variables), la matrice aura plutôt une forme d'ellipse (en 2D), ou ellipsoïde. L'orientation de la matrice **G** est

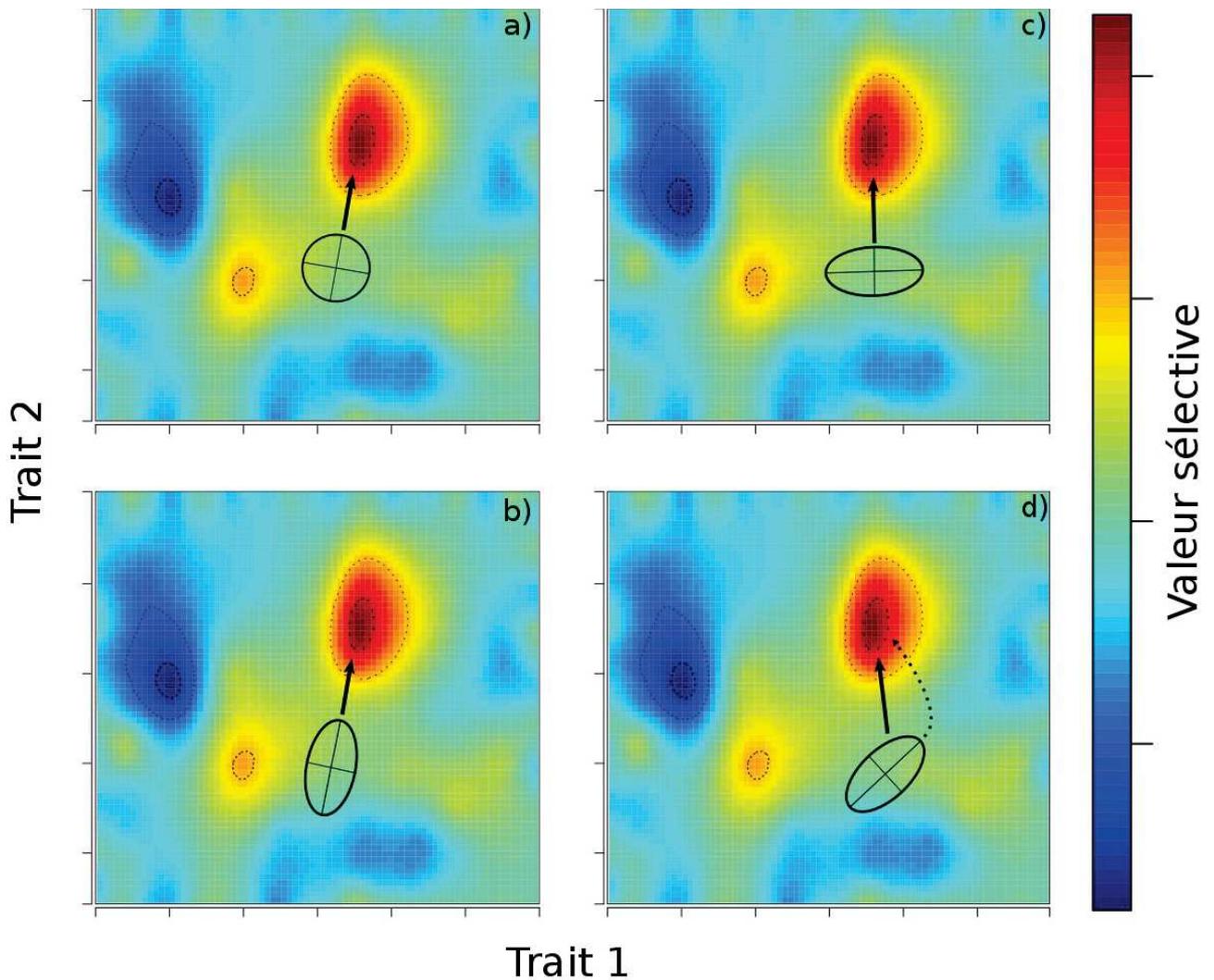


Figure 1. Représentation graphique de la matrice G sur un paysage adaptatif.

Note : Le paysage adaptatif représente un espace phénotypique théoriquement multidimensionnel (mais représenté en 2D ou 3D) dans lequel certaines combinaisons de traits correspondent à des optimums de valeur sélective dans l'environnement (pics adaptatifs, en rouge) et des combinaisons éloignées de cet optimum (vallées adaptatives, en bleu). L'évolution vers un pic adaptatif dépend de l'alignement de G (ellipse centrée sur le phénotype moyen de la population) avec la direction de la sélection (flèche pleine). Si la matrice est ronde (a) les variances génétiques sont également distribuées dans toutes les directions. Si elle est elliptique (b, c, d) certaines combinaisons de traits sont génétiquement plus variables et l'axe majeur de G , g_{max} , peut être aligné ou non avec la réponse à la sélection. L'évolution se fera plus rapidement si le vecteur de sélection est aligné avec g_{max} (b) et plus lentement si l'axe mineur de G est aligné avec la sélection (c). L'évolution suivra une trajectoire courbe si G et la sélection ne sont pas alignés (d).

déterminée relativement à la position de la population par rapport à l'optimum adaptatif (i.e. la combinaison de traits associée à une haute valeur sélective), qui définit d'ailleurs également la direction de la sélection. La réponse de la population vers cet optimum dépendra de la structure de **G** (taille et forme) ainsi que de son orientation (Fig. 1, et voir Arnold *et al.* 2001, Steppan *et al.* 2002, Bacigalupe 2011). Cette réponse se fait le long d'un axe de variation de **G** et dans la direction d'un pic sur le paysage adaptatif (Fig. 1a). La vitesse de déplacement de **G** vers l'optimum adaptatif dépend de l'alignement de **G** avec la direction de la sélection. Si la direction de la sélection est alignée avec l'axe majeur de **G** (i.e. **g**_{max}, possédant le plus de variation génétique) alors la réponse à la sélection sera rapide (Fig. 1b). Inversement la réponse sera lente si la direction de la sélection est alignée avec l'axe mineur de **G** (Fig. 1c). Dans ces deux cas, la trajectoire de **G** sur le paysage adaptative est linéaire et orientée directement vers l'optimum adaptatif. Si sélection et **g**_{max} ne sont pas alignés, c'est à dire si la sélection ne se fait pas dans la direction de variation génétique maximale, (Fig. 1d) alors **G** suivra une trajectoire courbe vers l'optimum : cette courbure illustre l'inertie de **G** et son effet contraignant sur la réponse à la sélection.

Une question centrale se pose alors sur les conséquences évolutives de **G** sur la divergence phénotypique : à quel point **G** influence-t-elle la divergence en réponse aux mécanismes évolutifs de sélection et de dérive génétique (e.g. Ackerman & Cheverud 2002)?

La prédiction en contexte de dérive est simple: **D**, la matrice de divergence entre populations dérivées (c'est à dire la matrice de covariance calculée à partir des moyennes des populations) issues d'une même population ancestrale devrait être proportionnelle à la matrice **G** de cette population ancestrale (Lande 1979, Hohenlohe & Arnold 2008). L'alternative, l'évolution adaptative, est plus complexe à prédire, en termes de patrons de divergence. On s'attend, à minima, à s'écarter de la proportionnalité entre **D** et **G** ancestrale, la divergence pouvant être contrainte dans sa direction et sa vitesse par l'axe majeur de **G** (Fig.1; Arnold *et al.* 2001). Cette prédiction simple a immédiatement été utilisée pour tenter d'interpréter à posteriori les patrons de divergence entre populations ou espèces (e.g. Ackerman & Cheverud 2004, Bégin & Roff 2004, Marroig & Cheverud 2005). Ackerman & Cheverud (2004) ont par exemple abordé cette question dans l'évolution des structures cranio-faciales au sein de la tribu des Hominini. Dans cette étude, les auteurs proposent de tester l'hypothèse d'une divergence neutre en comparant **D** et **G** ancestrale. Leurs résultats indiquent que la dérive seule ne peut pas expliquer tous les patrons de variations observés, certains étant attribuables à la sélection. La divergence phénotypique peut donc être le résultat de la sélection ou de la dérive, deux mécanismes alimentés par la variation et la covariation génétique (**G**) intra-spécifique (Ackerman & Cheverud 2002, 2004). Par conséquent, les trajectoires évolutives sont dépendantes de la stabilité de la matrice **G** (Encadré 1).

En résumé, l'évolution adaptative se fait donc plus facilement le long des directions génétiquement les plus variables (c'est à dire **g**_{max}; Blows & Hoffmann 2005, Hine & Blows 2006)

également appelées *lignes génétiques de moindre résistance* évolutives (Stebbins 1974, Futuyma *et al.* 1993, Schluter 1996).

Encadré 1. Stabilité et évolution de la matrice **G**

La stabilité de **G** est une question cruciale pour son utilisation en tant qu'outil d'exploration prédictif et rétrospectif de l'évolution (Turelli 1988, Arnold 1992, Björklund & Gustafsson 2013): si **G** varie rapidement, elle ne pourra être utilisée que pour des prédictions à court terme (à l'échelle de quelques générations) et le degré de conservation de **G** déterminera si elle permet de retracer l'histoire évolutive des populations.

Du point de vue théorique, tout changement des fréquences alléliques – par dérive, sélection ou autres effets démographiques (e.g. migration, admixture ; Nei 1977) – ou une association non-aléatoire d'allèles à différents loci (i.e. déséquilibre de liaison, Hartl & Clark 1997) devrait entraîner un changement de la structure de **G** (Turelli 1988, Carrière & Roff 1995, Scoville 2009). Prédire la nature de ces effets sur la géométrie de **G** s'avère toutefois très complexe. Jones *et al.* (2003, 2004) ont étudié l'effet de plusieurs mécanismes évolutifs – sélection, dérive, mutation – par une approche de modélisation. Brièvement, leurs résultats indiquent que les différents éléments de la géométrie de **G** ne sont pas toujours affectés de façon similaire par ces mécanismes et que la stabilité de **G** dépend également de paramètres démographiques (i.e. tailles de populations). Ces résultats renforcent l'importance – dans les analyses comparatives de **G** – de prendre en compte les différences entre tous les aspects de la géométrie de **G** (Philips & Arnold 1999).

Bien que stabilité de **G** soit une question empirique en soit (Revell 2007), des attendus théoriques sur l'évolution de **G** sont toutefois connus. Notamment, des effets potentiels de la sélection et de la dérive sur **G** ont été formulés (e.g. Roff & Mousseau 1999, Roff 2000, Philips *et al.* 2001, Whitlock *et al.* 2002)

1. Attendus théoriques sur l'évolution de **G**

a) Dérive

Si des processus neutres dominent (i.e. absence ou équivalence de sélection), plusieurs prédictions peuvent être faites sur la géométrie de **G**. La réduction de quantité de variance génétique dans la population par fixation aléatoire d'allèles, devrait se traduire par une réduction du volume de **G** sans en modifier la forme ni l'orientation (Roff & Mousseau 1999, Roff 2000). Ce changement résulterait, en théorie, en une matrice **G** proportionnellement plus petite que la matrice ancestrale. Si **G** évolue sur un paysage adaptatif plat (*drift landscape*, Arnold *et al.* 2001) alors sa trajectoire sous l'effet de la dérive sera chaotique, et difficilement prédictible (Arnold *et al.* 2001, Jones *et al.* 2003, 2004)

b) Sélection

L'effet de la sélection sur **G** est plus difficile à prédire. Il semble que globalement, la sélection, et plus particulièrement la sélection corrélative puisse stabiliser l'orientation de **G** (Jones *et al.* 2003, 2004, Revell 2007). Si la sélection est suffisamment forte pour altérer la fréquence des gènes du ou des traits sous sélection, alors la matrice **G** devrait être modifiée en conséquence, mais ce changement est dépendant de la direction et de l'intensité de la sélection (Roff 2000).

La stabilité de **G** n'est pas acquise en soi et résulte de la mise en jeu de nombreux paramètres sur différents aspects de **G**, qu'il est important de considérer en contexte empirique (Phillips & Arnold 1999).

2. Observations empiriques

a) Dérive

Empiriquement, l'effet de la dérive sur la matrice **G** a été estimé sur des lignées consanguines de *Drosophila melanogaster* maintenues en conditions de laboratoire pendant plusieurs générations (Philips *et al.* 2001, Whitlock *et al.* 2002). Deux résultats principaux ont émergé de ces études : i) conformément à la prédiction, la dérive génétique peut entraîner des réductions proportionnelles du volume de **G** dans les populations dérivées ; et ii) certaines lignées isolées présentent également des trajectoires évolutives clairement différentes, et des matrices **G** ayant évolué à la fois dans leur structure et dans leur orientation. Sous l'hypothèse d'évolution neutre, on s'attend alors à ce que la matrice **G** puissent être affectée par des paramètres quantitatifs de la variance génétique (i.e. modifications du volume de **G**), mais des changements de forme et/ou d'orientation semblent plus difficile à prédire. De plus, une divergence non-proportionnelle entre matrices ne permet pas de rejeter l'hypothèse de changement par dérive seule (Phillips *et al.* 2001) de part le caractère stochastique de ce mécanisme et de ces effets sur **G**.

b) Sélection

Les approches empiriques de l'effet de la sélection sur la structure et l'orientation de **G** sont encore assez rares et souvent menées dans un cadre expérimental (Wilkinson *et al.* 1990, Shaw *et al.* 1995, Doroszuk *et al.* 2008, Careau *et al.* 2015) potentiellement éloigné de la réalité biologique en conditions naturelles (Eroukhmanoff 2009, Jamniczky & Hallgrimsson 2009).

Comment alors estimer l'effet de la sélection sur la stabilité de **G** dans les populations naturelles ? Certaines études empiriques se sont intéressées à cette question (e.g. Cano *et al.* 2004, Eroukhmanoff & Svensson 2011) et partagent une méthodologie commune: la comparaison de matrices **G** entre populations distribuées dans des environnements différents. En effet, un changement d'environnement devrait (dans certains cas) s'accompagner d'un changement de pressions de sélections. Ainsi, on s'attend à ce que des populations ayant colonisé des habitats différents, présentent des matrices **G** différentes du fait, entre autres, de la sélection.

Cette approche comparative a par exemple révélé dans des populations divergentes de l'isopode aquatique *Asellus aquaticus* un changement dans l'orientation et la structure de **G** chez les populations expérimentant de nouvelles conditions environnementales (Eroukhmanoff & Svensson 2011). Dans ce dernier cas, l'approche semble particulièrement pertinente puisque l'histoire de la population est connue. Ainsi, il est possible de comparer directement la structure de la matrice **G** ancestrale avant divergence (i.e. celle de la population source) à la structure d'une ou plusieurs matrices dérivées. Il est donc possible dans ces conditions, d'inférer les effets de l'histoire évolutive sur **G** (cette histoire étant connue) mais également de déterminer l'effet de la structure de **G** sur la divergence. Un système idéal serait donc un cas dans lequel il serait possible de comparer **G** entre populations ancestrales et populations dérivées ; que le temps de divergence soit suffisamment court pour qu'il soit possible d'estimer **G** dans la population ancestrale ; que cette divergence soit le résultat d'un changement d'environnement et donc potentiellement, de mécanismes évolutifs neutres ou sélectifs. Ce cas de figure est typiquement retrouvé dans les cas d'invasions biologiques, et est particulièrement prometteur dans l'analyse de l'évolution de **G** (Eroukhmanoff 2009, Bacigalupe 2011).

I – 5. Plasticité phénotypique

Les populations ne possédant pas de variation génétique dans la direction de la sélection sont-elles alors condamnées à l'extinction ? Comme le formulent Arnold & Lande (1983) « *la sélection agit sur les phénotypes, sans prendre en compte leurs bases génétiques [...]* ». Quelle qu'en soit l'origine, génétique ou non, un changement de phénotype en direction de l'optimum adaptatif, sera avantageux. Si ce phénotype est produit de façon plastique alors la plasticité phénotypique peut être considérée comme un mécanisme adaptatif. Bien que longtemps négligée par les théoriciens de l'évolution (Pigliucci 2005), la plasticité phénotypique est désormais un mécanisme évolutif intégré dans la synthèse moderne (e.g. West-Eberhard 2003 ; Pigliucci 2005, 2009, Merilä & Hendry 2014, Levis & Pfennig 2016). La plasticité n'apporte cependant pas toujours d'augmentation de la valeur sélective : elle peut être neutre, voire mal-adaptative (e.g. Ghalambor *et al.* 2007). C'est ce qui est prédit en particulier en réponse à un stimulus environnemental nouveau dans l'histoire évolutive de l'espèce (i.e. qui n'a jamais été rencontré par les populations). C'est le cas notamment de la réponse plastique aux pesticides (Levis *et al.* 2016). Chez la salamandre maculée (*Ambystoma maculatum*) l'exposition de juvéniles à différents pesticides produit chez les adultes des variations de morphologie (e.g forme de la tête et de la queue) qui n'ont pas d'effets (dans ce cas fonctionnels) sur les individus plastiques par rapport aux individus contrôles.

Le caractère adaptatif de la plasticité a néanmoins été suggéré dans de nombreuses études, et souvent de façon convaincante (e.g. Pfennig *et al.* 2006, Brakefield *et al.* 2007, Ledón-Rettig *et al.* 2010 et voir Beldade *et al.* 2011 pour revue). Certaines populations de la mésange charbonnière (*Parus major*) ont par exemple ajusté leur comportement de ponte de façon plastique (Charmantier *et al.* 2008) en réponse à un changement du climat saisonnier au cours du temps. Ce phénotype plastique est dans ce cas un avantage dans le nouvel environnement (i.e. le nouveau climat) puisqu'il coïncide avec l'émergence des larves de la Phalène brumeuse (*Operophtera brumata*), une ressource essentielle à l'alimentation des jeunes chez *P. major*.

Le rôle de la plasticité dans la réponse à la sélection est depuis longtemps sujet à controverse: dans quelle mesure la plasticité phénotypique s'oppose-t-elle ou au contraire, facilite l'évolution adaptative (Ghalambor *et al.* 2007, Chevin *et al.* 2010, Merilä 2015) ? Le rôle évolutif de la plasticité a été en particulier discuté dans le contexte des changements de pics adaptatifs. Comment une population peut-elle évoluer d'un pic à l'autre, c'est à dire traverser les vallées de faible valeur sélective sans être ramenée à son point de départ par la sélection ? Il a été suggéré que la plasticité adaptative puisse permettre cette traversée (Price *et al.* 2003, Crispo 2007, Ghalambor, 2007): soumise à un changement d'environnement, si une population a dans son histoire, rencontré des conditions environnementales similaires, elle peut avoir conservé une plasticité sélectionnée dans ces conditions (plasticité adaptative). La réponse plastique induite par

les nouvelles conditions pourrait alors rapprocher suffisamment la population du nouveau pic adaptatif et ainsi permettre à la sélection vers ce pic de prendre le relais. La plasticité aura ainsi facilité la réponse à la sélection, d'autant plus si le changement phénotypique plastique est « imparfait », et nécessite un ajustement par sélection de variants génétiques pour se rapprocher du nouveau pic.

Une question connexe est de savoir comment une réponse plastique peut-elle être adaptative dans un environnement totalement nouveau³ (De Jong 2005, Ghalambor *et al.* 2007)? Il a été suggéré que cela pourrait être le cas si les conditions changent de manière graduelle (Chevin & Hoffmann, in press). Si les conditions du nouvel environnement sont très différentes de l'environnement historique, les populations peuvent être soumises à un stress environnemental. Ces conditions peuvent entraîner un accroissement de la variance phénotypique par réponse plastique au delà de la norme de réaction historique de l'espèce (Hoffmann & Hercus 2000, Ghalambor *et al.* 2007). Cette variation se retrouvera alors exposée à la sélection et si certains des traits exprimés sont avantageux ils seront sélectionnés. L'augmentation de cette variance phénotypique par un stress environnemental a été proposée comme étant le résultat de l'expression de variation génétique cryptique (e.g. Rutherford 2003, Gibson & Dworkin 2004, Leddon-Retig *et al.* 2010). Cette variation génétique, n'ayant jamais été exposée à la sélection (car non exprimée dans l'environnement normal) s'est accumulée de façon neutre dans l'histoire des populations (Ghalambor *et al.* 2007). Exprimée dans le nouvel environnement, si elle est par chance associée à un phénotype adaptatif, alors elle pourra être sélectionnée et suivre un modèle d'évolution Darwinienne (Waddington, 1942, 1953, Levis & Pfennig 2016).

I – 6. Conclusion sur la partie I

L'évolution adaptative met donc en jeu des processus génétiques et plastiques complexes de réponse à la sélection. Qu'il s'agisse de déterminer les mécanismes évolutifs à l'origine de la diversification espèces (e.g. Ackerman & Cheverud 2004), de comprendre les dynamiques contemporaines de colonisation d'habitats (e.g. Lee 2002), ou d'anticiper les réponses des populations aux changements climatiques (Charmantier *et al.* 2008), l'étude des patrons de variation génétique et phénotypique est un enjeu fondamental de la biologie évolutive.

Une limitation des études cherchant à interpréter les patrons de divergence *a posteriori* est – outre qu'elle utilisent le plus souvent **P**, la matrice de covariance phénotypique comme approximation de **G** – qu'elles ne disposent généralement pas de la matrice **G** ancestrale. Cette dernière est alors inférée à partir des données de la matrice **P** intra-populationnelle commune (*pooled within-group matrix*), soit de façon plus subtile en tenant compte de la phylogénie.

Un contexte particulièrement propice pour tester les hypothèses de proportionnalité entre **G**

³ Conditions environnementales en dehors de la gamme historiquement rencontrée par la population.

et **D**, implique donc de connaître populations ancestrales et populations dérivées. Ce contexte se retrouve dans les invasions biologiques, qui ont été déjà il y a plus de 40 ans, proposés comme des contextes particulièrement favorables à l'étude de l'évolution (Baker & Stebbins 1965)

II – Invasions biologiques : introduction et considérations évolutives

II – 1. Introduction

Une invasion biologique correspond à la prolifération d'individus d'une espèce hors de son aire de répartition historique (Richardson *et al.* 2000). Elle peut être définie comme la suite d'événements regroupés en trois étapes principales (Coalautti & Maclsaac 2004, Valéry *et al.* 2008) (Fig. 2) : La première étape, l'introduction, correspond au déplacement géographique d'une population depuis son aire de répartition d'origine (ou aire native ; Fig 2). L'introduction se caractérise par trois événements principaux. i) L'échantillonnage aléatoire d'un groupe d'individus dans l'aire native, qui constitueront les fondateurs de la population envahissante ; ii) la dispersion, qui correspond au déplacement (aidé ou non) de ces individus vers la nouvelle aire; iii) le relâchement des individus fondateurs dans la nouvelle aire. La seconde étape est l'établissement (également appelée « naturalisation », Richardson *et al.* 2000), qui correspond à la capacité des fondateurs introduits à survivre et à se reproduire dans leur nouvel environnement (Fig. 2). Enfin, la troisième étape est l'expansion démographique et spatiale de la population dans sa nouvelle aire géographique (Fig 2). Toutes les introductions d'espèces n'aboutissent néanmoins pas systématiquement à l'invasion de la nouvelle aire (Zenni *et al.* 2013). Un nombre insuffisant de fondateurs, des conditions environnementales défavorables ou encore la présence d'un compétiteur ou d'un prédateur dans l'écosystème de relâchement, sont autant de facteurs qui peuvent compliquer voire empêcher l'invasion. Toutefois, lorsqu'elles arrivent à terme, les invasions biologiques sont largement considérées comme l'un des composants majeur du changement environnemental global (Vitousek *et al.* 1997) et sont le plus souvent associées à des conséquences néfastes dans les zones géographiques envahies (Encadré 2).

Comme le suggèrent justement Bock *et al.* (2015), il existe un véritable rapport « amour-haine » entre biologistes et espèces envahissantes. Les lignes de conduite principales des autorités publiques en matière d'invasions biologiques sont entre autres : empêcher l'accroissement du nombre de nouvelles introductions d'espèces exogènes, le confinement des populations récemment introduites et l'éradication des invasions déjà établies (IUCN, 2012). Un grand nombre de recherches ont donc été menées dans le but de faciliter ces directives (Wittenberg & Cock 2001, Carlton 2003). Toutefois, une autre catégorie de recherches, plus fondamentales, notamment dans le domaine de la biologie évolutive, ont utilisé les invasions biologiques comme une opportunité d'étudier des changements écologiques et évolutifs (e.g. compétition, extinction, adaptation, dérive) en conditions naturelles.

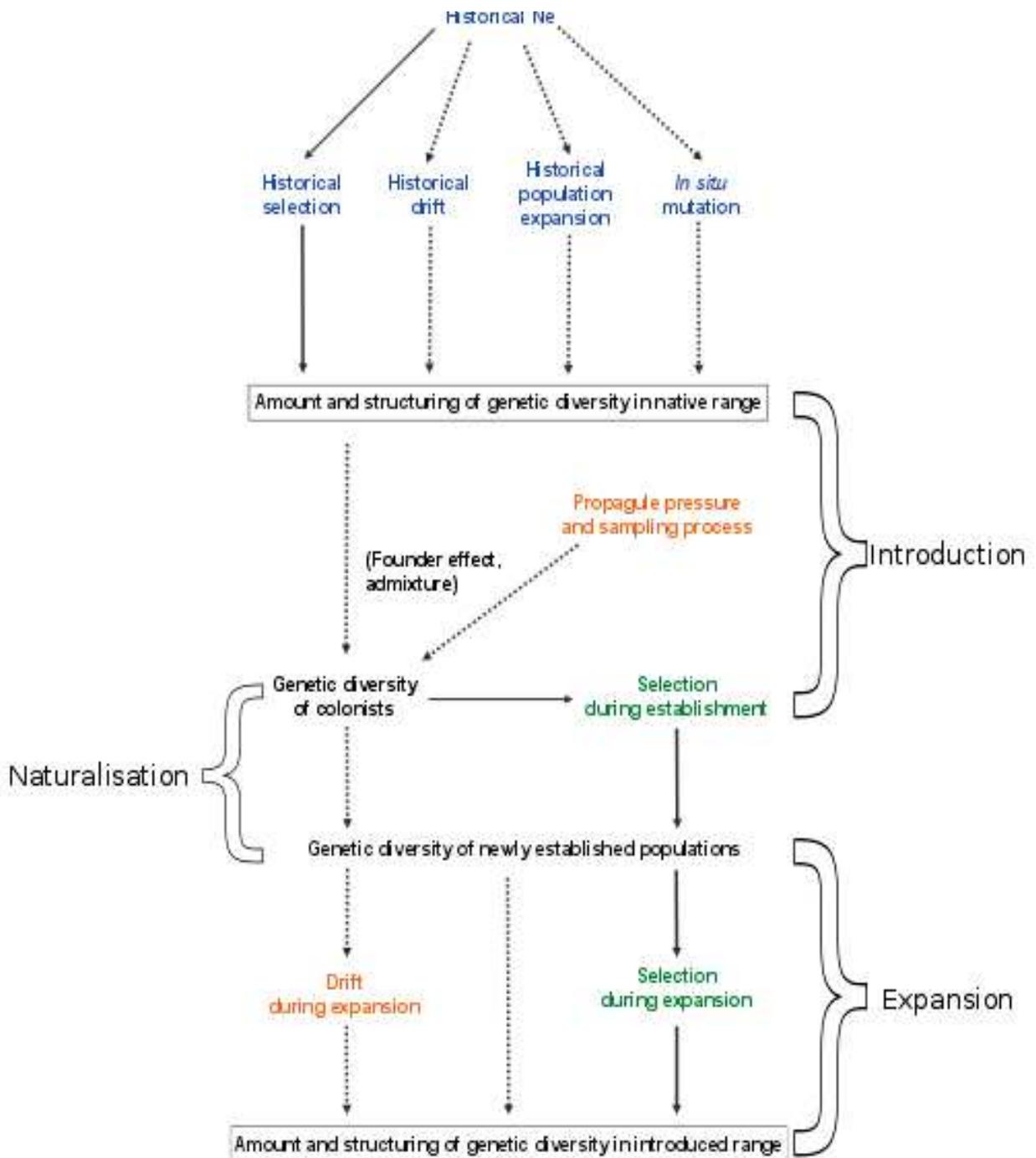


Figure 2. Représentation schématique d'une invasion biologique (adaptée de Keller & Taylor 2008)

La diversité génétique des populations natives est le résultat de l'histoire évolutive de l'espèce (en bleu) et implique des mécanismes aléatoires de dérive (orange) et sélectifs (en vert). Le long des chemins empruntés par les populations envahissantes (flèches) la variation génétique va évoluer selon les effets de la sélection (flèches pleines) et/ou des différents événements démo-génétiques au cours de l'invasion (e.g. goulots d'étranglements, admixture etc., flèches pointillées). Globalement, l'invasion peut être divisée en trois étapes principales (Introduction, naturalisation et expansion) et les études empiriques de génétiques quantitatives ou des populations peuvent se baser sur les patrons de variations à certains moments clés de l'invasion (encadrés noirs)

Encadré 2. Les invasions biologiques dans le contexte de globalisation

La part des activités humaines dans la dispersion d'espèces à grande échelle, qu'elle soit spatiale (dispersion sur de grandes distances) ou temporelle (multiplication des événements d'introduction) est devenue prépondérante dans la période récente en relation avec le développement des échanges commerciaux et des flux migratoires depuis la deuxième moitié du XX^e siècle (Perrings *et al.* 2005, Leprieur *et al.* 2008). Le nombre d'introductions en Europe s'élevait à 10 nouveaux établissements d'espèces par an à la fin du XX^e siècle (Hulme *et al.* 2009). Beaucoup de ces introductions ont causé, et causent encore, des dégâts considérables sur les écosystèmes naturels, les agro-systèmes, et sur l'économie.

Impacts sur les écosystèmes

Bien qu'il soit dans certains cas difficile d'estimer de façon précise les coûts écosystémiques de l'introduction d'espèces (Pereira *et al.* 2013), les invasions biologiques sont considérées comme une cause majeure du déclin des écosystèmes et plus particulièrement de réduction de la biodiversité mondiale (Clavero *et al.* 2009). Dans une récente revue, Maxwell *et al.* (2016) placent les invasions biologiques en quatrième position des menaces principales sur la biodiversité et plus particulièrement sur les espèces classées dans la liste rouge de l'IUCN (iucnredlist.org). Comment les invasions impactent-elles la biodiversité autochtone ? Par effets directs, les espèces envahissantes peuvent être responsable de la transmission de pathogènes dans l'écosystème envahi. Pour l'organisme envahissant vecteur, la dissémination d'un ou plusieurs pathogènes suivant l'étape d'introduction peut conférer un avantage compétitif dans la nouvelle aire, si des espèces autochtones partageant la même niche écologique ne possèdent pas les défenses immunitaires adéquates pour lutter contre l'infection (Vilcinskis 2015). Un exemple connu de ce phénomène se retrouve dans la disparition de l'écureuil roux (*Sciurus vulgaris*) au Royaume-Uni suite à l'introduction de l'écureuil gris (*Sciurus carolinensis*) originaire des États-Unis. Ce dernier, porteur d'un virus mortel pour l'espèce Européenne (*parapoxvirus*, Rushton *et al.* 2000) mais relativement bénin pour lui, s'est substitué à l'écureuil autochtone en seulement deux décennies (Tompkins *et al.* 2003).

Les invasions peuvent également engendrer des dérèglements en cascade des écosystèmes et impacter plusieurs niveaux d'organisation de ces derniers. Par exemple, l'île Christmas au large de l'Australie, a connu un profond changement de son écosystème suite à l'introduction de la fourmi *Anoploepis gracilipes*. Cette invasion a impacté au moins trois niveaux trophiques de l'île, de l'écroulement des populations du crabe endémique *Gecarcoidea natalis*, jusqu'à la biodiversité du sol, et de la forêt primaire (O'Dowd 2003).

Impacts sur l'économie et les agrosystèmes

De nombreux rapports font état de pertes s'élevant à des dizaines de milliards de dollars annuels dans le monde, du seul fait des espèces envahissantes (Ricciardi *et al.* 2011). Une grande part de ces pertes correspond à des coûts de production des cultures agricoles, sur lesquelles les espèces invasives ont un impact significatif (Paini *et al.* 2016). Certaines espèces envahissantes (d'insectes notamment) sont appelées « ravageurs des cultures », du fait de leurs effets nuisibles sur ces cultures. Ces espèces peuvent impacter les productions en tant que consommateurs des produits agricoles (e.g. insectes phytophages). La punaise asiatique par exemple (ou punaise diabolique, *Halyomorpha halis*) est un ravageur polyphage particulièrement nuisible aux États-Unis et en Europe, qui se nourrit sur une grande variété d'espèces à valeur commerciale et dont l'impact économique a été évalué à plus de 20 milliards de dollars aux États-Unis depuis son introduction (Cesari *et al.* 2015). L'effet sur les agrosystèmes peut également être dû à l'introduction de pathogènes dans les cultures, indirectement (si l'espèce est vecteur du pathogène) ou directement

(si l'espèce est le pathogène). Il existe une grande diversité de ces pathogènes (e.g. Bactéries, virus, champignons etc.) et plus de la moitié des maladies affectant les plantes agricoles, sont dues à des introductions (Bebber *et al.* 2013).

En prenant en compte l'ensemble des biens économiques (e.g. agriculture, horticulture, tourisme etc.) les insectes envahissants coûtent au minimum 70 milliards de dollars annuels dans le monde, un chiffre toutefois probablement sous-estimé (Bradshaw *et al.* 2016).

Impacts sur la santé publique

Enfin, certaines invasions biologiques ont des conséquences directes sur la santé humaine. Les moustiques du genre *Aedes* sont des cas notoires de dispersion et de transmission de maladies graves pour l'Homme tels que la Dengue, le Chikungunya ou plus récemment l'épidémie du virus Zika en Europe (Bonizzoni *et al.* 2013, Akiner *et al.* 2016).

II – 2. Les invasions biologiques comme laboratoires naturels

Les invasions biologiques ont été proposées comme des modèles particulièrement utiles à l'étude de l'évolution phénotypique (Baker & Stebbins 1965, Sakai *et al.* 2001, Bacigalupe *et al.* 2009, Eroukhmanoff *et al.* 2009) du fait de l'importance présumée des mécanismes évolutifs – particulièrement de la sélection et de la dérive (Lee 2002) – au cours de ces événements. Cet attrait pour les invasions comme modèle évolutif est peut-être en partie lié à leur similitude avec le modèle de spéciation allopatrique (i.e. séparation dans des aires géographiques non superposées de plusieurs populations, aboutissant à leur divergence et *in fine* la formation d'espèces séparées, Mayr 1970). Le rôle des invasions comme facteur de spéciation a d'ailleurs souvent été proposé, mais rarement étudié empiriquement (e.g. Lee 2002 pour revue). La particularité des invasions par rapport au modèle de Mayr, est que la plupart nous sont contemporaines et représentent donc les prémices d'un tel modèle. Les invasions offrent donc la possibilité d'étudier la divergence entre populations dans un contexte naturel.

II – 3. Histoire de l'invasion

Un des avantages du modèle de l'invasion est qu'il est possible d'en retracer l'histoire de façon détaillée en analysant l'évolution de la variation génétique neutre. Bien que la variation neutre ne soit pas toujours un estimateur pertinent du potentiel évolutif (Sgrò *et al.* 2011 et voir Merilä & Crnokrak 2001 pour discussion), elle fournit des données essentielles quant à la compréhension des dynamiques démographiques et génétiques des invasions (Sgrò *et al.* 2011). Ces données peuvent alors aider à l'interprétation en aval, des patrons de divergence quantitative, notamment en permettant l'identification d'événements de goulots d'étranglements ou d'admixture génétique au cours de l'invasion (Estoup & Guillemaud 2010). Pour cela, il faut donc comprendre comment les niveaux de variations génétiques peuvent fluctuer au cours de l'invasion.

II – 3 – a. Goulots d'étranglements et dérive génétique

L'initiation de l'étape d'introduction par échantillonnage aléatoire d'un faible nombre d'individus dans l'aire native aura comme conséquence immédiate de réduire, parfois drastiquement, la variabilité génétique de la population introduite (Tsutsui *et al.* 2000, Sakai *et al.* 2001). En effet, lors de cet échantillonnage il existe une probabilité que certains allèles rares de la population ne soient pas échantillonnés ; probabilité qui sera d'autant plus forte que le nombre d'individus échantillonnés sera faible. De plus, dans ces populations de petite taille, la dérive sera d'autant plus rapide (Nei 1978), réduisant encore la variation génétique des colonisateurs. On retrouve de nombreux cas dans la littérature de populations envahissantes bien moins variables génétiquement que leurs homologues natifs (e.g. Puillandre *et al.* 2007, Schmid-Hempel *et al.*

2007, Dlugosch & Parker 2008). En France par exemple, l'étude de la variation génétique mitochondriale et nucléaire du frelon asiatique envahissant (*Vespa velutina*) indique un fort goulot d'étranglement au moment de l'introduction, la population Française ne présentant qu'un seul haplotype, et moins de la moitié de la richesse allélique par rapport aux populations Asiatiques natives (Arca *et al.* 2015). Lors de l'expansion géographique de l'invasion, la récurrence des effets fondateurs chez les populations du front de migration (i.e. à la marge de la distribution) peut également réduire la variation génétique (Schulte *et al.* 2013). C'est le cas par exemple, chez la demoiselle *Erythromma viridulum* en Europe, où une perte de diversité génétique significative a été clairement détectée au sein des populations situées en marge de l'expansion par rapport aux populations du cœur de la distribution (Watts *et al.* 2010).

II – 3 – b. Admixture génétique

Si l'invasion est le fruit non pas d'un événement unique mais d'introductions multiples, alors chaque nouvelle introduction aura le potentiel de réapprovisionner la variation génétique de la population introduite par admixture (Lavergne & Molofsky 2007, Dlugosch & Parker 2008, Keller & Taylor 2008). Ainsi la variation génétique des populations introduites peut être substantiellement supérieure à celle des populations natives (Kolbe *et al.* 2007). Si les introductions multiples se sont fait à différents endroits d'une aire géographique, la dispersion des individus à partir de chacun de ces points peut mener à la rencontre de plusieurs fronts d'invasion et à une admixture au niveau de la zone de contact (Lombaert *et al.* 2014). On observe ainsi dans les populations envahissantes Nord-américaines de la coccinelle Asiatique (*Harmonia axyridis*) une zone de contact dans les états de l'Utah et du Colorado, entre deux fronts d'invasions ayant pour sources deux introductions indépendantes (Lombaert *et al.* 2014). Le corollaire de l'admixture est que plus les différentes introductions proviennent de sources différenciées génétiquement, plus la variance génétique dans la population introduite sera élevée. L'introduction de propagules génétiquement différenciées augmente la probabilité de production de génotypes nouveaux, par recombinaison, accroissant encore la variation génétique dans l'aire envahie (Keller & Taylor 2008).

Goulots d'étranglements et admixture génétique sont donc deux événements fréquents lors de l'invasion, pouvant modeler la variation génétique des populations introduites. Tous deux ont été respectivement été proposés comme participant au succès des invasions, par exemple en purgeant certains allèles délétères (Facon *et al.* 2011) ou à l'inverse favoriser l'apparition d'allèles allèles avantageux dans la population (Lavergne & Molofsky 2007). Bien que les mécanismes biologiques du succès des espèces envahissantes soient toujours très discutés (e.g. Bock *et al.* 2015, Estoup *et al.* 2016 ; Encadré 3), il est clair que le potentiel adaptatif des espèces envahissantes en est un élément central.

Encadré 3. Exemples d'hypothèses sur le succès d'invasion.

Préadaptation

L'hypothèse dite de la préadaptation (Facon *et al.* 2006) suggère que le succès d'invasion peut être expliqué par la similitude des environnements, colonisés et natifs. Le nouvel environnement ne poserait alors pas de « défi adaptatif » aux populations introduites (Estoup *et al.* 2016) qui s'y développeraient aussi bien que dans leur aire native. Dans ce cas particulier, on ne s'attend donc pas à ce que l'architecture génétique des populations envahissantes soit différente de celle des populations natives (Bacigalupe *et al.* 2013). *A contrario* de nombreuses hypothèses ont suggéré que le potentiel adaptatif des populations envahissantes est déterminant dans le succès d'invasion (e.g. Rosecchi *et al.* 2001, Richards *et al.* 2006, Facon *et al.* 2008, voir Hayes & Barry 2008 pour une méta-analyse).

Taille de propagules

Parmi les hypothèses démographiques on retrouve celles liées au nombre de propagules comme déterminant du succès d'invasion. Dans ce cas, le succès de l'invasion est directement lié à la taille initiale de la population introduite dans le nouvel environnement qui, si elle est suffisamment large, assurera la viabilité des propagules à l'introduction. Par exemple, Ahlorth *et al.* 2003 ont expérimentalement simulé la colonisation d'un nouvel environnement par la punaise d'eau *Aquarius najas*. En faisant varier la taille et le nombre des groupes colonisateurs, ils ont observé une relation proportionnelle entre ces deux facteurs et le succès de colonisation (correspondant dans ce cas à la maturation des juvéniles dans le nouvel environnement).

Enemy Release Hypothesis

Une hypothèse basée sur le contexte écologique est celle de l'*Enemy Release Hypothesis* (ERH; Crawley 1997) qui stipule que les populations introduites bénéficient de l'absence de prédateurs ou de pathogènes naturels dans le nouvel environnement. Torchin *et al.* (2003) ont testé cette hypothèse en mesurant la teneur en parasites de 26 espèces envahissantes représentant 7 taxons, dans leurs aires natives et envahissantes, et on observé une quantité double de parasites chez les populations natives. Néanmoins cette hypothèse est dans certains considéré comme relativement simpliste (Colautti *et al.* 2004) et ne semble pas être applicable à de nombreux modèles d'invasions (Liu & Stiling 2006). Un contre-exemple notable se retrouve dans l'invasion de la guêpe *Vespula vulgaris*. Chez cette espèce les populations envahissantes possèdent un taux d'infection parasitaire égale à celui des populations natives (Lester *et al.* 2014).

Evolution of Increase Competitive Ability

L'influence du relâchement des interactions biotiques sur le succès d'invasion est également proposé par l'hypothèse d'*Evolution of Increase Competitive Ability* (EICA ; Blossey & Nötzold 1995). Cette hypothèse, proche de l'ERH, a été formulé afin d'expliquer à la fois le succès des invasions mais également une particularité de certaines populations envahissantes, une meilleure performance⁴ des individus de l'aire envahie par rapport à leurs homologues natifs. En effet, en réponse au relâchement de pression de prédation, de compétition ou de pathogénicité, certaines espèces sont capables de ré-allouer les ressources liés à la défense immunitaire par exemple, vers la reproduction ou le développement (Parker *et al.* 2013). C'est le cas notamment chez le crapaud buffle *Rhinella marina* (*Buffo marinus*) envahissant en Australie, où les individus du front de l'invasion sont plus performants pour disperser (Philips *et al.* 2006) mais moins résistants aux

⁴ Paramètres morphologiques ou d'histoire de vie tels que la taille, la fécondité ou l'abondance relative (Parker *et al.* 2013)

pathogènes (Llewellyn *et al.* 2012).

Anthropogenically Induced Adaptation to Invade

Cette hypothèse, proposée par Hufbauer *et al.* (2012), stipule que certaines invasions peuvent être facilitées par l'homogénéisation des environnements induites par l'Homme. Plus particulièrement, cette hypothèse suggère que les espèces adaptées à des environnements anthropisés (e.g. cultures agricoles) seront plus aptes à coloniser des environnements également anthropisés hors de leur aire native. Bien que cette hypothèse nécessite de plus amples confirmations empiriques, certains modèles candidats ont été proposés: le Doryphore *Leptinotarsa decemlineata* par exemple aurait bénéficié de la culture intensive de la pomme de terre pour élargir son aire naturelle de distribution du continent Américain (Mexique et USA), jusqu'en Europe.

II – 4. Potentiel adaptatif et succès invasif

Les espèces envahissantes étant soumises dans certains cas à de fortes pressions de sélection, de dérive, de pressions démographiques (e.g. petite taille de population) ou écologiques (e.g. pressions de prédatons), comment expliquer qu'autant d'entre elles soient aussi prolifiques ? Comment des espèces ayant par exemple subi une réduction de leur diversité génétique (via goulots d'étranglement et dérive) peuvent-elles coloniser un environnement auquel elles ne sont pas historiquement adaptées, et parfois même y remplacer les espèces autochtones ? Cette question est considérée par beaucoup d'auteurs (e.g. Sax & Brown 2000, Allendorf *et al.* 2003, Estoup *et al.* 2016) comme *le paradoxe des invasions*. Si les conditions abiotiques (e.g. température, précipitation, salinité) et biotiques (e.g. présence de prédateurs ou de compétiteurs) sont différentes dans l'environnement envahi le succès de l'invasion dépendra alors du niveau de variance génétique, et/ou de la plasticité phénotypique des populations pour répondre à ces pressions sélectives (i.e. leur potentiel adaptatif, Colautti & Lau 2015).

II – 4 – a. Variation génétique et succès invasif

Pendant toutes les étapes de l'invasion, la quantité de variation génétique additive est cruciale pour répondre à la sélection. Une variation génétique suffisante dans la population introduite augmenterait la probabilité de réponse à la sélection, par exemple lors de l'étape d'établissement (Lee *et al.* 2002). Cette variation génétique peut être d'ores et déjà présentes dans les populations introduites – i.e. des allèles présents dans la population ancestrale introduits dans la nouvelle aire – et on parle alors de *standing genetic variation* (Orr & Betancourt 2001, Barrett & Schluter 2008). Dans l'aire native, ces allèles peuvent être neutres face à la sélection, voire délétères, mais se retrouver sélectionnés dans l'aire envahie. Par exemple chez la jacinthe *Eichhornia paniculata*, la colonisation de l'Amérique centrale et des Caraïbes depuis l'aire native au Brésil, s'est

accompagnée de l'apparition d'un phénotype autogame absent des populations natives. Dans ce nouvel environnement, la capacité d'auto-fécondation est sélectionnée en l'absence des pollinisateurs naturels d'*E. paniculata* et a favorisé l'invasion de ce phénotype dans ces régions (Barrett *et al.* 2008). Des analyses moléculaires ont révélé que la variation génétique associée à ce phénotype (i.e. allèles régulateurs récessifs) était bien présente dans les populations natives, mais en très faible fréquence (Husband & Barrett 1993). Cette variation ancestrale⁵ n'est toutefois pas toujours présente pour induire une réponse rapide à la sélection. Le temps de réponse à la sélection a d'ailleurs été suggéré pour expliquer la latence souvent observé entre la première introduction et l'expansion géographique des populations (*lag phase*, Sakai *et al.* 2001). Cette période correspondrait alors au temps nécessaire à la population introduite pour se réapprovisionner en diversité génétique, par mutation *de novo* (Linnen *et al.* 2009, et voir Dlugosh *et al.* 2015 pour revue). L'admixture entre populations génétiquement différenciées a également été proposé comme mécanisme facilitant l'adaptation (De Carvalho *et al.* 2010). Dans ce cas, l'apport de nouveaux allèles peuvent être incorporé dans l'aire envahie, et/ou de nouveaux génotypes potentiellement avantageux peuvent être produit par recombinaison (Rius & Darling 2014). Par exemple chez *P. arundinacea*, Lavergne & Molofsky (2007) ont observé chez les populations envahissantes de plus forts niveaux de variance génétique, dus à de multiples introductions de différentes sources de l'aire native. Ces nouveaux génotypes sont associés chez les populations envahissantes à des phénotypes plus « agressifs » (e.g. accroissement de la biomasse, de la reproduction clonale etc.) qui ont assurés le succès de cette espèce dans sa nouvelle aire.

II – 4 – b. Plasticité phénotypique et succès invasif

Une hypothèse pouvant expliquer le succès des invasions biologiques est basée sur la plasticité phénotypique. Selon cette hypothèse la plasticité de certains traits permettrait aux populations envahissantes de répondre rapidement à la sélection dans le nouvel environnement, particulièrement en l'absence de variance génétique suffisante à l'adaptation locale (e.g. suite à un goulot d'étranglement ; Geng *et al.* 2007).

Lavergne & Molofsky (2007) ont par exemple observé chez *P. arundinacea* un plus grand niveau de plasticité pour plusieurs traits (e.g. nombre de feuilles, longueur de la tige etc.) en réponse à un gradient d'humidité chez des clones dérivés de populations envahissantes. La plasticité est dans ce cas suggérée comme adaptative puisqu'elle facilite l'acquisition des ressources dans l'aire envahie (les traits plastiques sont donc avantageux dans le nouvel environnement). Si la plasticité est adaptative, elle peut être sélectionnée positivement dans la zone envahie et faciliter l'évolution des populations envahissantes (Lande 2015).

⁵ Par soucis de simplification dans ce manuscrit, le terme „*variation ancestrale*“ sera utilisé pour référer à la variation génétique présente dans les populations natives et retrouvée également dans les populations envahissantes.

L'analyse des patrons de variation génétique neutre des populations envahissantes peut donc nous renseigner sur les dynamiques démographiques et l'histoire des invasions. Ces histoires peuvent être complexes et impliquer des mécanismes de sélection et de la dérive. Quels sont alors les attendus théoriques sur l'effet de ces mécanismes sur la variation quantitative ? Et quelles attendus peut on formuler quant à l'évolution de la matrice **G** dans un contexte d'invasion biologique ?

II – 5. Conséquences sur les trajectoires évolutives

II – 5 – a. Attendus sur la matrice **G**

L'étape d'introduction (Fig. 2) est généralement caractérisée par des populations de taille réduite, une faible diversité génétique et donc un fort effet de la dérive (Sakai *et al.* 2001). Dans ce contexte, la théorie (Avery & Hill 1977, Arnold *et al.* 2001, Jones *et al.* 2003, 2004⁶) prédit que les matrices **G** des populations envahissantes peuvent avoir divergé (dans leur structure et leur orientation ; Phillips *et al.* 2001) aléatoirement par rapport aux matrices des populations natives. Néanmoins un effet général de réduction du volume de **G** est attendu si la dérive est dominante (Encadré 1). Si au contraire des effets sélectifs dominent, l'évolution des matrices **G** dérivées dépend alors de la direction de cette sélection locale.

Si les populations sont pré adaptées à leur nouvel habitat (Encadré 3), la sélection devrait être globalement similaire (i.e. même direction) entre aires native et envahie, et donc stabiliser les matrices dans un état proche de leur configuration ancestrale (Hohenlohe & Arnold 2008). A l'inverse, des pressions de sélection différentes peuvent être à l'œuvre entre populations envahissantes et natives. Dans ce deuxième cas la théorie prédit que la matrice **G** dérivée s'alignera à terme avec la direction de cette sélection (Arnold *et al.* 2008, Eroukhmanoff 2009 ; Fig. 1) et sa divergence de la matrice ancestrale dépendra donc de la direction de cette sélection.

Ce schéma peut bien entendu se complexifier dans les étapes suivantes de l'invasion, du fait par exemple, d'introductions multiples et d'événements d'admixture : la structure et l'orientation de **G** peuvent être impactées par l'introgession d'allèles nouveaux dans la population. Par exemple, l'axe majeur de **G** peut se ré-orienter dans la direction de la divergence dans le cas d'une forte migration entre une population source et une population dérivée et son volume augmenter en réponse à l'augmentation du niveau de variance génétique (Guillaume & Whitlock 2007). Durant l'étape de dispersion, un changement des fréquences alléliques est attendu du fait d'effets fondateurs répétés, de dérive (Olivieri 2009), de sélection du fait d'un environnement hétérogène (Parker *et al.* 2003, Killkenny & Galloway 2012) ou d'admixture génétique entre plusieurs fronts d'invasions (Lombaert *et al.* 2014), tout autant d'événements qui peuvent impacter la géométrie de

⁶ Bien que le modèle de Jones *et al.* (2003, 2004) s'intéresse à l'évolution de **G** dans le temps (i.e. d'une même population), nous transposons ici certains de leurs attendus théoriques dans le cas de la divergence entre populations envahissantes et natives.

G.

En contexte invasif, une divergence des matrices **G** est donc attendue du fait de la complexité probable des événements démographiques et génétiques mis en jeu au cours de l'invasion. Une connaissance préalable de l'histoire de l'invasion du modèle biologique d'intérêt est donc essentiel afin de formuler des hypothèses sur l'évolution de **G**.

II – 5 – b. Attendus sur la divergence phénotypique

Comme il a été discuté, des attendus relativement simples ont été formulés sur la divergence en contexte sélectif et de dérive, et invoquent une proportionnalité entre la matrice de divergence **D** et la matrice **G** ancestrale. En contexte d'invasion on s'attend donc, si la dérivé génétique domine (e.g. après un goulot d'étranglement au moment de l'introduction) à ce que la divergence soit proportionnelle à la matrice **G** de la population source de l'invasion (Holenlohe & Arnold 2007). Si l'invasion est associée à un changement de régime sélectif, **D** et **G** ancestrale peuvent ne pas être proportionnelles, la divergence pouvant être contrainte dans sa direction et sa vitesse par l'axe majeur de **G** (Fig.1; Arnold *et al.* 2001). Ces prédictions peuvent toutefois être compliquées par la multiplicité des paramètres pouvant, au cours de l'invasion, déstabiliser la matrice **G** (e.g. petites tailles de populations, Jones *et al.* 2003, 2004). L'interaction de ces paramètres sélectifs, stochastiques, démographiques, de la stabilité de **G** et des conséquences pour la divergence, est une question extrêmement complexe qui demande à être étudié empiriquement (McGuigan 2006, Revell 2007)

II – 6. Conclusion sur la partie II

Les attendus quantitatifs de l'évolution de **G** sont très dépendants du contexte démographique, génétique, et plus largement de l'histoire des populations. La prise en compte de l'histoire de l'invasion est donc essentielle à la formulation d'hypothèses sur l'évolution de **G** et son influence sur l'évolution phénotypique (Bacigalupe 2009) ainsi que dans la compréhension du potentiel évolutif des populations envahissantes, et de leur succès invasif. L'intégration des méthodes de génétique des populations et de génétique quantitative permet donc de tester ces hypothèses, à condition que le modèle biologique d'étude soit compatible avec l'approche expérimentale de ces méthodes.

III – Modèle d'étude : La Drosophile à ailes tachetées, *Drosophila suzukii*.

La drosophile à ailes tachetées (*Drosophila suzukii*, Matsumura 1931 ; Fig. 3) est une espèce du genre *Drosophila* appartenant au groupe *melanogaster*, historiquement décrite comme originaire d'Asie du sud-est. En Asie, sa très large distribution recouvre une grande partie de la Chine et des

pays continentaux voisins, ainsi que l'archipel Japonais (Ometto *et al.* 2013, Asplen *et al.* 2015). En 1980, l'espèce est observée pour la première fois hors de son aire de répartition native, sur l'archipel d'Hawaii (Kaneshiro 1983) mais aucune autre observation hors-Asie ne sera déclarée pendant près de 30 ans. En 2008, la présence de *D. suzukii* est documentée de façon quasi-simultanée en Europe et aux Etats-Unis, respectivement dans la région de Barcelone et dans l'état de Californie (Calabria *et al.* 2010, Hauser 2011). Dès 2009, l'espèce s'est déjà dispersée aux Etats-Unis vers l'état voisin de l'Oregon, au Nord, et au sud de la Californie dans la région de San Diego. En Europe au même moment, elle est identifiée dans le sud de la France dans la région de Montpellier. Aujourd'hui, l'espèce a colonisé la majeure partie des continents Nord-Américain et Européen et on la retrouve dans l'hémisphère sud au Brésil (Deprà *et al.* 2014) et sur l'île de la Réunion (Facon, pers. com.)



Figure 3. Femelle de *Drosophila suzukii* pondant sur une framboise (*Rubus* sp.)

Contrairement à la plupart des espèces de drosophiles, les femelles de *D. suzukii* utilisent leurs ovipositeur sclérifié pour percer la peau des baies et de certains fruits afin d'y déposer leur œufs. Les larves se développent alors à l'intérieur du fruit, cause des dégâts terribles sur les productions agricoles. Crédits photographie : Yann le Poul, Antoine Fraimout & Vincent Debat.

D. suzukii possède deux caractéristiques morphologiques la différenciant nettement de la plupart des espèces du groupe *melanogaster* : i) les mâles présentent des tâches de mélanine sur l'apex des ailes (qui confèrent à l'espèce son surnom de Drosophile à ailes tachetées); et ii) les femelles possèdent un ovipositeur sur-développé et sclérifié (Atallah *et al.* 2014). Cette dernière caractéristique – en connexion avec le caractère invasif de l'espèce – est à l'origine de l'intérêt porté à *D. suzukii*. En effet, alors que la plupart des espèces de drosophiles favorisent les substrats en décomposition pour la ponte, *D. suzukii* a une préférence pour les fruits en cours de maturation, ou mûrs (Keeseey *et al.* 2015). Grâce à leur ovipositeur, les femelles sont capables de transpercer la peau d'un grand nombre d'espèces de baies et de fruits rouges (Poyet 2015) afin d'y déposer leur œufs. Le développement des larves à l'intérieur du fruit entraîne la destruction partielle de ce dernier, et l'expose aux infections secondaires. *D. suzukii* est ainsi responsable de dégâts considérables sur les productions agricoles (on estime en 2011 un coût total annuel en Italie de 3 millions d'euros dû aux infestations de *D. suzukii* (Cini *et al.* 2012), et de près de 40% de perte de productions aux Etats-Unis, 1 an seulement après le début de l'invasion (EPPO).

Bien que l'espèce ait généré un nombre exponentiel d'études depuis ses premières observations en Europe et aux États-Unis, de nombreuses questions restent ouvertes concernant l'histoire et le succès de cette invasion à grande échelle. L'invasion de *D. suzukii* offre un modèle idéal pour l'étude de l'évolution rapide en conditions d'invasion. Sa proximité phylogénétique avec l'espèce modèle *Drosophila melanogaster* et son temps de génération court permettent d'établir des élevages en laboratoire et la mise en place de protocoles expérimentaux. De plus, cette invasion mondiale ayant été initiée il y a moins d'une décennie, elle permet de suivre les mécanismes évolutifs en temps réel.

Drosophila suzukii est un modèle d'étude idéal pour aborder les questions posées dans cette thèse. En plus de son caractère invasif qui, comme nous l'avons vu, est particulièrement pertinent du point de vue de la biologie évolutive, sa proximité phylogénétique avec le modèle *Drosophila melanogaster* (Kopp & True 2002) offre certains avantages méthodologiques. D'un point de vue purement biologique – et bien que leur écologie soit très différente (Ometto *et al.* 2003) – les innombrables connaissances générées au cours des cent dernières années sur *D. melanogaster* (Bellen *et al.* 2010) nous fournissent des attendus théoriques précieux sur la physiologie (e.g. réponse à la température), la génétique (e.g. paramètres mutationnels, Fraimout *et al.* 2016) ou encore le comportement (e.g. comportement de ponte en laboratoire). Du point de vue moléculaire par exemple, la multitude de ressources génomiques disponibles chez *D. melanogaster* (e.g. Adams *et al.* 2000) peut servir de référence au développement d'outils moléculaires chez *D. suzukii* (Fraimout *et al.* 2015). Pour l'analyse des traits quantitatifs, *D. suzukii* est également un modèle idéal, par exemple pour l'étude de l'évolution de l'aile de drosophile (e.g. Houle *et al.* 2003 ; Gidaszewski *et al.* 2009 ; Bolstad *et al.* 2015).

L'aile de drosophile a été étudiée en contexte macro- et micro-évolutif et il a été suggéré

qu'il existe un signal phylogénétique pour la forme de l'aile entre espèces (Houle *et al.* 2003). La divergence de cette forme est toutefois relativement faible et l'aile apparaît comme un caractère stable à l'échelle de la famille *drosophilidae* (Houle *et al.* 2003, Hansen & Houle 2004, 2008). Des études de sélection artificielle en laboratoire ont néanmoins montré que cette stabilité n'est pas due à des contraintes génétiques absolues⁷ (au moins, chez *D. melanogaster*; Weber 1992, Mezey & Houle 2005) et que l'aile peut évoluer dans d'autres directions que celles observées en nature. Ces différents résultats suggèrent donc que l'aile est un caractère sous sélection naturelle stabilisante (Hansen & Houle 2004, 2008). En conditions naturelles, des cas d'évolution de l'aile de plusieurs espèces de drosophiles ont néanmoins été documentés (Azevedo *et al.* 1998, Huey *et al.* 2000, Gilchrist *et al.* 2001, 2004) en particulier dans la mise en place de clines parallèles de latitude entre continents. Le caractère adaptatif de ces variations de formes reste encore sujet à discussion et de fait – en plus de la fonction évidente de l'aile dans le vol, et dans les comportements de cour dans le genre *Drosophila* (e.g. Greenspan & Ferveur 2000) – l'étude de l'évolution de l'aile dans le cadre de l'invasion de *D. sukii* est particulièrement intéressante pour nourrir ce débat.

Enfin, la forme de l'aile est un caractère intrinsèquement multivarié (Encadré 4), particulièrement approprié aux études de génétiques quantitative telles que présentées ici. En effet, selon le modèle infinitésimal (Fisher 1930) les traits quantitatifs varient de façon continue et sont l'expression d'une multitude de gène à effets faibles et additifs (Lande 1981). La variation quantitative peut donc être subtile et son analyse nécessite une approche tout aussi subtile dans l'extraction de l'information. Dans ce contexte, la morphométrie géométrique (i.e. l'analyse de la forme, Encadré 4) permet de mesurer finement la variation continue des caractères morphologiques ainsi que son évolution dans le temps, et dans l'espace (e.g. Bookstein 1996, Rohlf 1998, Adams *et al.* 2009). Cet ensemble de méthodes permet en outre d'analyser séparément taille et forme, mais également d'étudier leur covariation, c'est à dire l'allométrie, concept central en génétique quantitative (e.g. Lande *et al.* 1979, Voje *et al.* 2014; Bolstad *et al.* 2015).

L'invasion de *Drosophila sukii* offre donc la possibilité de tester des hypothèses relevant de nombreux domaines de la biologie évolutive et ce travail tente, par une approche aussi intégrative que possible, de répondre à ces questions selon plusieurs grands axes de recherche.

⁷ Le terme „contrainte génétique absolue“ définit l'absence de variation génétique pour un, ou plusieurs traits (i.e. combinaison de traits) quantitatifs, et indique une impossibilité de réponse à la sélection dans la direction de cette contrainte.

IV – Grands axes de la thèse

IV – 1. Echantillonnage et maintien des populations

L'échantillonnage est un élément crucial de toute étude cherchant à décrire des systèmes biologiques naturels (Albert *et al.* 2010). La stratégie d'échantillonnage est dictée par des considérations statistiques (i.e. nombre suffisant d'échantillons) mais également biologiques (i.e. échantillons pertinents pour répondre à la question posée). Ici, nous avons pour objectif de décrire l'évolution de la variation génétique neutre et quantitative de populations mondiales de *D. sukukii*. Cet objectif nécessite donc l'acquisition de deux types de matériel : i) des individus piégés en milieu naturel et conservés en alcool pour les analyses de génétique moléculaire (i.e. Chapitre 1) et ii) des individus vivants, également piégés en milieu naturel, afin d'établir des populations en laboratoire (i.e. Chapitres 2 & 3). Dans les deux cas, l'échantillonnage se devait d'être suffisamment exhaustif pour inclure des populations natives et des populations envahissantes. Nous avons donc échantillonné (Fig. 4) un total de 23 populations de *D. sukukii* (cf. Fig. 1, Fraimout *et al.* 2016) parmi lesquelles six ont été mises en élevage au laboratoire (cf. Chapitre 2). Ces populations d'élevage ont été maintenues dans des conditions contrôlée (i.e. température, nourriture etc.) tout au long du processus expérimental, et ont servi de base à l'établissement des protocoles de génétique quantitative, et de plasticité phénotypique.

IV – 2. Histoire de l'invasion de *D. sukukii*

Une approche comparative de l'architecture génétique entre populations natives et envahissantes semble particulièrement pertinente afin d'identifier les mécanismes évolutifs impliqués dans la divergence entre populations. Dans le cadre d'une analyse comparative de matrices **G** ancestrales et dérivées, il faut néanmoins être sûr de l'origine génétique des populations dérivées, et plus généralement, du lien de parenté entre les populations étudiées. Les populations peuvent être structurées génétiquement, même à faible échelle géographique (Slatyer *et al.* 2014) et dans le cas d'une invasion biologique, une telle structuration dans l'aire native peut impliquer la contribution de plusieurs sources génétiques différentes dans l'aire envahie (Estoup & Guillemaud 2010). De plus, certaines populations envahissantes peuvent être dérivées non pas de l'aire native, mais d'autres populations envahissantes établies plus tôt dans le processus d'invasion (on parle alors de populations « têtes de pont », *bridgehead populations*, Lombaert *et al.* 2011).



Figure 4. Echantillonnage de *Drosophila suzukii* au Japon en collaboration avec Hideyuki Mitsui.

L'histoire d'invasion de *Drosophila suzukii* est potentiellement caractérisée par des flux migratoires importants entre continents, ainsi qu'au sein de chaque pays, du fait du lien fort entre la biologie de l'espèce et la production agricole. Une étape initiale essentielle est donc de retracer l'histoire des populations envahissantes et d'identifier leur(s) origine(s) génétique(s). Pour cela, la génétique des populations dispose d'outils permettant de caractériser la variation génétique moléculaire des populations afin d'établir les relations entre les populations.

Le premier objectif est le développement de marqueurs moléculaires permettant de caractériser la variation neutre dans les populations natives et envahissantes de *D. suzukii*. Le choix du type de marqueur est une décision préalable essentielle, qui dépend de l'objectif de l'étude (Avisé 2012). Ici, la quantité de données moléculaires doit être proportionnelle aux nombres de populations impliquées dans l'invasion. Autrement dit, les échantillons récoltés doivent être suffisamment représentatif de l'ensemble de la distribution mondiale de *D. suzukii*. Ce grand nombre d'échantillon requiert donc l'utilisation de marqueurs à la fois précis dans leur détection des allèles de la population, mais ne demandant pas un effort de génotypage prohibitif. Les marqueurs microsatellites sont particulièrement pertinents pour cette approche (Vignal *et al.* 2002). Une fois ces marqueurs développés spécifiquement pour *D. suzukii* (Fraimout *et al.* 2015), le second objectif est de génotyper l'ensemble des populations naturelles échantillonnées afin de produire les données nécessaires à la reconstruction des routes d'invasion de *D. suzukii* (Fraimout *et al.* 2016).

IV – 3. Evolution de la matrice **G en contexte invasif**

Les travaux théoriques sur la stabilité de **G** et son effet sur la divergence entre populations font appel à des considérations complexes dont il est parfois difficile de tirer des prédictions claires. En condition d'invasion, l'évolution de certains aspects de la matrice **G** est toutefois attendue (e.g. taille de la matrice), et l'invasion de *D. suzukii* est un contexte idéal pour mesurer cette évolution.

A partir de populations naturelles de *D. suzukii* échantillonnées dans trois pays représentatifs de son aire native (Japon) et envahissante (USA et France), nous avons pu établir en laboratoire des populations stocks, utilisables dans le cadre d'une approche de génétique quantitative. L'estimation des paramètres de variance génétique à partir de traits quantitatifs nécessite la mise en place de familles au sein de chaque population d'intérêt (Lynch & Walsh 1998). Ici, chaque population de *D. suzukii* a été maintenue sous forme de lignées isofemelles (Hoffmann & Parson 1988, David *et al.* 2005) établies à partir de femelles gravides et dont les descendants ont été amplifiés (i.e. reproduits et maintenus au sein de chaque lignée) pendant 5 générations en laboratoire. Toutes les populations étant maintenues dans un environnement commun (i.e. même température, même milieu nutritif) la variance environnementale est réduite au minimum. La mesure de la variance phénotypique entre lignées nous permet donc d'estimer la

variance génétique de la population. A partir de données phénotypiques sur la forme de l'aile (Encadré 4) l'objectif est d'estimer les matrices **G** de six populations de *D. suzukii* (2 populations Françaises, 2 populations Nord-Américaines et 2 populations Japonaises) afin de comparer leur géométrie (taille, forme et orientation).

De plus, nous proposons grâce à notre stratégie d'échantillonnage de tester si des différences géographiques entre populations sud et nord sont détectables et potentiellement répétées dans les différentes zones géographiques, comme c'est le cas chez d'autres espèces de drosophiles envahissantes (Azevedo *et al.* 1998, Huey *et al.* 2000, Gilchrist *et al.* 2001, 2004). Dans ces exemples, la mise en place de clines latitudinaux a pris plusieurs décennies (20 ans chez *D. subobscura*, Gilchrist *et al.* 2004), soit une échelle de temps nettement supérieure à celle de notre étude. Il était donc particulièrement intéressant de documenter les premières phases de l'installation de *D. suzukii*. Plus particulièrement, nous testons l'hypothèse selon laquelle des différences affecteraient la géométrie de **G** entre continents. En effet, si les clines de latitudes pour l'aile de drosophile sont adaptatives, et associées aux conditions environnementales, nous nous attendons à une forte similitude entre les matrices estimées sur des latitudes équivalentes entre continents (cf. Eroukhanoff & Svensson 2011).

Encadré 4. Analyse de la forme : Morphométrie géométrique.

La génétique des populations atteint ses limites lorsqu'il s'agit d'expliquer les bases génétiques de la variation continue (David & Samadi 2004). Pourtant la plupart des traits quantitatifs en population naturelle varient de façon continue et ne sont pas le résultat de l'expression d'un seul gène mais plutôt d'une multitude de gènes à effets faibles et additifs (Fisher 1918, Lande 1981). Pour mesurer cette variation subtile sur des caractères quantitatifs, il faut donc être capable d'extraire l'information de façon aussi subtile. Pour cela, l'analyse de la variation de forme des caractères morphologiques – la morphométrie – offre des outils méthodologiques particulièrement adéquats (e.g. Bookstein 1996, Rohlf 1998, Adams *et al.* 2009). L'analyse de la forme des objets biologiques peut se faire de plusieurs façons, mais nous nous intéresserons ici son approche moderne, la morphométrie géométrique, basée sur l'analyse des coordonnées de points de repères. L'analyse de forme se fait à partir d'images du trait d'intérêt obtenues par différentes méthodes (e.g. photographie, scanner etc.). Il s'agit ensuite de placer sur ces images des points de repère homologues permettant de caractériser la structure globale du trait. Ces points homologues, appelés landmarks (qui correspondent sur la figure à l'intersection des veines de l'aile de drosophile) sont communs à tous les individus, mais varient dans leur position. La forme de l'objet est indépendante de sa taille, sa position ou son orientation (Klingenberg 2001) et ces informations doivent alors être extraites. Pour cela, la morphométrie géométrique dispose d'une procédure nommée Superposition Procrustes (Dryden & Mardia 1998). Cette procédure se fait en trois étapes successives à partir des coordonnées originales des landmarks : i) Mise à l'échelle : cette étape permet de retirer la variation de taille entre les objets tels qu'ils ont été photographiés afin de les mettre à une échelle de taille standardisée, la taille centroïde. ii) Translation: cette étape permet de retirer la variation de position des objets, en plaçant tous les centres de gravité de ces derniers au centre d'un système de coordonnées x, y. iii) Rotation : toutes les configurations d'objets sont pivotées par rapport à leur centre de gravité commun afin de minimiser la distance

entre les landmarks (i.e. la somme de leur déviation au carrée). La variation de forme peut alors être mesurée comme la variation entre les coordonnées des landmarks après superposition Procrustes. Cette variation peut par la suite être analysée de façon multivariée et visualisée le long de n'importe quelle direction de l'espace des formes (e.g. composantes principales, axes discriminants, droites de régression etc.). La morphométrie géométrique a souvent été utilisée dans un contexte de génétique quantitative (e.g. Klingenberg *et al.* 2004, Martinez-Abadias *et al.* 2006 et voir Polly 2008 pour revue) et est une méthode particulièrement pertinente pour l'analyse de l'évolution des caractères quantitatifs.

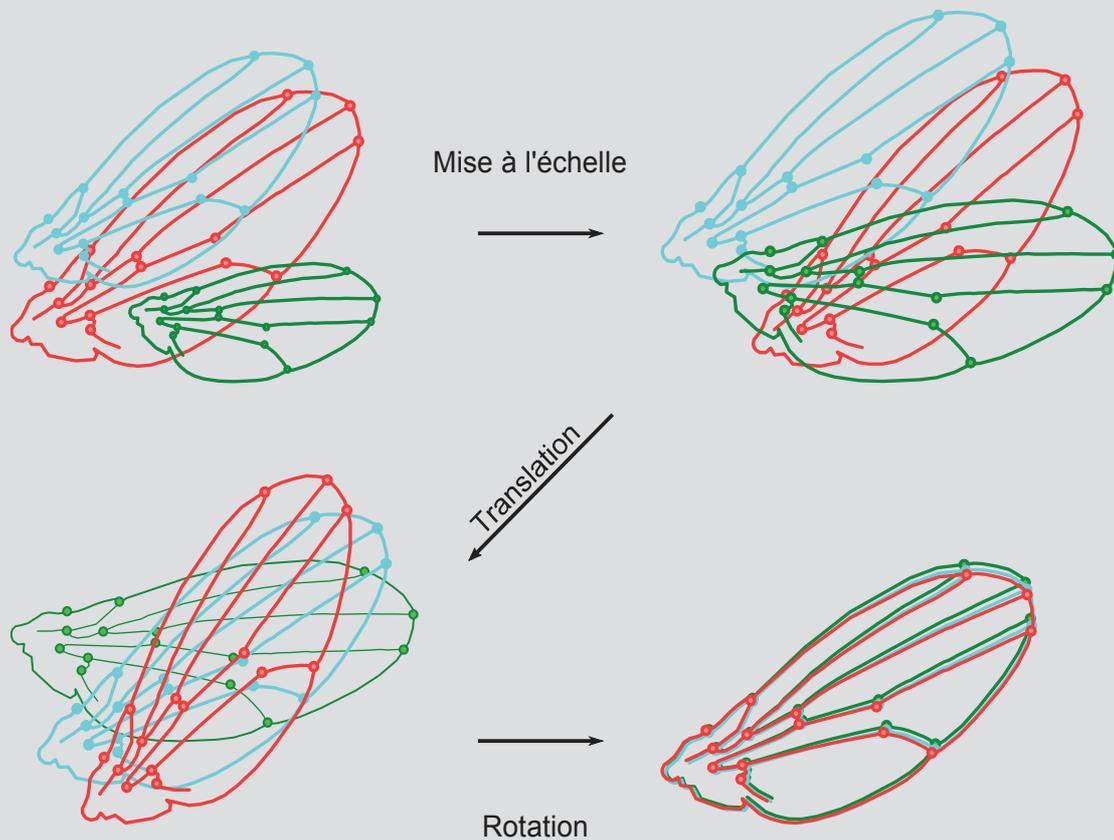


Figure 5. Superposition Procrustes Les configurations (ici visualisées sur des ailes de *Drosophiles*) sont mises à l'échelle, soumises à une translation puis une rotation qui minimisent la distance entre points de repères.

IV – 4. Plasticité phénotypique et performances de vol en réponse à la température de développement

Une hypothèse communément admise en biologie des invasions est que les espèces envahissantes sont plus plastiques, cette plasticité les rendant plus aptes à coloniser un nouvel habitat (Yeh & Price 2004, Davidson *et al.* 2011). De même, au sein de ces espèces, les populations envahissantes seraient également plus plastiques que leurs homologues natifs (e.g. Blossey & Nötzold 1995, Bossdorf *et al.* 2005, voir Parker *et al.* 2013 pour revue). Si la plasticité est adaptative on s'attend en effet à ce que le phénotypique plastique soit avantageux dans le nouvel environnement, mais le lien entre phénotype et fonction du trait (i.e. la mesure dans

laquelle il peut être avantageux) n'est pas toujours évalué de façon directe. Afin de tester ces hypothèses, il est donc nécessaire de comparer les niveaux de plasticité entre populations envahissantes et natives, et d'étudier le lien entre la réponse plastique du trait, et sa fonction (i.e. la performance).

L'objectif de ce chapitre est de mesurer et comparer la plasticité phénotypique entre populations natives et envahissantes de *D. suzukii*. Pour cela, nous utilisons l'aile comme trait focal et la température comme stimulus environnemental. La température est un facteur majeur dans l'écologie des organismes ectothermes, qui influence fortement leur distribution géographique. Une forte plasticité phénotypique à la température est typiquement observée chez les drosophiles (e.g. Gibert *et al.* 2000). En particulier, la taille du corps mais également la taille et la forme de l'aile sont sensibles aux conditions thermiques de développement. La température est en outre une variable relativement aisée à manipuler en laboratoire, notamment dans le cadre de l'élevage de drosophiles (i.e. en incubateurs). Ici, nous mesurons l'effet de trois températures de développement⁸ sur la morphologie alaire de *D. suzukii* et comparons cet effet entre populations envahissantes (Française et Américaine) et une population native (Japonaise). Les populations envahissantes sont-elles plus plastiques ? Le second objectif de ce chapitre est de mesurer l'association entre la forme de l'aile et les performances de vol. Plus particulièrement nous proposons de tester l'effet attendu de changements de forme de l'aile en réponse à la température sur le vol. Une association entre plasticité phénotypique de l'aile et performances de vol est insuffisante pour démontrer la valeur adaptative de cette plasticité, mais elle en constitue une condition *sine qua none* : l'absence d'effet suggérerait au contraire que la plasticité de l'aile n'est pas adaptative, du moins en termes de vol.

IV – 5. Evolution de la niche climatique de *D. suzukii*

En moins d'une décennie, *Drosophila suzukii* a colonisé plus de 20 pays distribués sur les deux hémisphères (Asplen *et al.* 2015). L'extrême rapidité avec laquelle cette espèce a pu coloniser une telle diversité d'environnements soulève plusieurs questions quand à la tolérance de cette espèce aux conditions climatiques. En effet, bien que des événements d'adaptation locale soient souvent observés en contexte d'invasion (e.g. Colautti & Barrett 2013), la probabilité que des populations envahissantes se soient indépendamment adaptées (et de façon répétée) à l'environnement de chacune des zones envahies semble assez faible. Une hypothèse plus parcimonieuse peut être formulée selon laquelle, *D. suzukii* est une espèce généraliste, c'est à dire possédant une grande tolérance physiologique aux conditions environnementales, et capable de se développer dans une grande diversité de climats. Cette hypothèse peut être testée en répondant à la question : La niche écologique (*sensu* Hutchinson 1957) de *D. suzukii* dans son aire de répartition d'origine est-elle

⁸ Température appliquée tout au long du développement des individus, de l'oeuf, à l'adulte.

suffisamment large pour permettre à l'espèce de coloniser d'autres régions du monde ? En d'autres termes : les niches envahissantes de *D. sukuzii* sont-elles incluses dans sa niche native ?

A partir de données de présence de l'espèce récoltées dans la littérature et sur le terrain, il est possible par une approche de modélisation de niche climatique (i.e. Species Distribution Modeling ; SDM, Miller 2010) d'estimer l'enveloppe climatique ancestrale (i.e. native) de *D. sukuzii* pour plusieurs variables environnementales (e.g. variations de température, précipitations etc.). A partir de cette information, les SDMs peuvent comparer cette enveloppe climatique à celles des régions envahies, et ainsi déterminer les probabilités de présence de l'espèce dans ces régions (Tingley *et al.* 2014). Cette approche permet de mesurer le degré de différenciation entre les environnements natifs et envahissants. Dans le cas des invasions biologiques, il est donc possible de déterminer si les populations envahissantes ont subi un *shift* de niche, et se sont établies dans un environnement différent de leur environnement natif (Broennimann *et al.* 2007, Medley 2010).

Une forte variance environnementale entre aires natives et envahies est un *a priori* important sur lequel se basent de nombreuses hypothèses de biologie des invasions mais qui n'est pas toujours évalué formellement. Ce dernier chapitre a donc pour objectif principal de déterminer si *D. sukuzii* a rencontré au cours de son invasion des environnements très différents, et si cette variance environnementale peut contribuer à expliquer les résultats obtenus dans les chapitres précédents.

Chapitre 1 : Génétique des populations et histoire de l'invasion de *D. suzukii*

Article 1.

Frainout, A., Loiseau, A., Price, D.K., Xuereb, A., Martin, J.-F., Vitalis, R., Fellous, S., Debat, V., and Estoup, A. (2015). New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest. *Eur. J. of Entomol.*

New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest

ANTOINE FRAIMOUT¹, ANNE LOISEAU², DONALD K. PRICE³, ANNE XUÉREB², JEAN-FRANÇOIS MARTIN²,
RENAUD VITALIS², SIMON FELLOUS^{2,*}, VINCENT DEBAT^{1,*} and ARNAUD ESTOUP^{2,*}

¹Institut de Systématique, Evolution et Biodiversité, ISYEB, UMR 7205, CNRS, UPMC, EPHE, Muséum National d'Histoire Naturelle, Sorbonne Universités, 45 rue Buffon, 75005 Paris, France; e-mails: antoine.fraimout@edu.mnhn.fr; debat@mnhn.fr

²INRA UMR, Centre de Biologie pour la Gestion des Populations (INRA/IRD/Cirad/Montpellier SupAgro), 755 avenue du Campus Agropolis, CS 30 016, 34988 Montferrier-sur-Lez cedex 17, France; e-mails: loiseau@supagro.inra.fr; xuereb@supagro.inra.fr; martinjf@supagro.inra.fr; vitalis@supagro.inra.fr; simon.fellous@supagro.inra.fr; estoup@supagro.inra.fr

³Department of Biology, and Tropical Conservation Biology and Environmental Science Graduate Program, University of Hawaii, 200 West Kawili Street, Hilo, HI 96720, USA; e-mail: donaldp@hawaii.edu

Key words. Diptera, Drosophilidae, *Drosophila suzukii*, population genetics, microsatellite, invasion routes, invasive species, genetic structure

Abstract. Historical and observational data for invasive species are often sparse and incomplete, so molecular genetic markers are increasingly used and have proved to be efficient tools to make inferences about invasion histories. Here, we report the development of 28 polymorphic microsatellite markers in the invasive spotted-wing drosophila, *Drosophila suzukii* (Matsumura, 1931) (Diptera: Drosophilidae) designed from recent genomics resources, and their cross-amplification in closely related *Drosophila* species of the *suzukii* subgroup. The markers, located on autosomal chromosomes, were characterized in two distinct invasive populations from France and Hawaii (USA), and in four sister species of the *suzukii* subgroup. They all showed substantial polymorphism as well as revealing strong genetic differentiation between the two genotyped populations. These markers represent a promising molecular tool for population genetic studies on *D. suzukii*, more especially in order to reconstruct the pathways and demographic processes of the world-wide invasion in this major insect pest.

INTRODUCTION

The spotted-wing drosophila, *Drosophila suzukii* (Matsumura, 1931) (Diptera: Drosophilidae) is an oriental species of the *melanogaster* group. It is widely distributed in south-eastern Asia, from India (Parshad & Paika, 1965) to Japan, where it was first described in 1931 (Hauser, 2011). *D. suzukii* females exhibit a large serrated ovipositor which allows them to lay eggs in ripening fruits, hence causing severe damages to stonefruit and red berry crops. The species has drawn particular attention during the past five years because of its status as a crop pest and its rapid invasion across Europe and the USA (reviewed in Hauser, 2011 and Calabria et al., 2012). When an invasive species presents immediate threats in a new environment (e.g. economical, ecological or for human health), most research plans legitimately focus on how to prevent further expansion or to reduce damages caused by introduced populations (Thresher et al., 2014). However, reconstructing the geographic pathways followed by propagules from their original source location to the newly invaded location (hereafter referred to as invasion routes) is often overlooked. In light of this, even though some crucial aspects of *D. suzukii* biology and ecology have been already investigated (Kimura, 2004), little is known about the species' invasion history and more especially the origins of the different introduced populations (but see Adrion et al., 2014 for a first approach). This is unfortunate, since robust information about the invasion routes of a species is crucial for at least two principle reasons: (i) it facilitates the design of strategies for controlling present or preventing future invasions;

and (ii) it is requirement necessary for defining and testing different hypotheses concerning the environmental and evolutionary factors responsible for biological invasions in general (Estoup & Guillemaud, 2010).

Most of our knowledge about the introduction routes of invasive species derives from historical and observational data, which are often sparse, incomplete and, sometimes, misleading. In the case of *D. suzukii*, this approach suggests there was a single introduction in Europe, located in France in between Spain and Italy, where the pest was first reported (Cini et al., 2014). An alternative approach is to use population genetics analyses based on polymorphic molecular markers, such as microsatellites, which have proved to be powerful tools for reconstructing invasion routes, highlighting the complexity and the often counter intuitive nature of the true story (Estoup et al., 2001; Estoup & Guillemaud, 2010). Here we report the development of a set of 28 polymorphic microsatellite markers in *D. suzukii* designed from recent genomics resources, as well as characterizing genetic variation at these particular loci in two distinct invasive populations along with their cross-amplification patterns in closely related *Drosophila* species.

MATERIAL AND METHODS

Sampling and DNA extraction

Adult *D. suzukii* were collected from the field at two distinct localities of the invasive range between October and November, 2013, and stored in ethanol. The Hawaii population consisted of

* Equal authorship.

33 adult flies caught at the Saddle Road locality (19°47'35.1", 155°30'54.3"W), the French population consisted of 27 adult flies caught near Montpellier (43°42'00"N, 3°51'56"E). Total genomic DNA was extracted using 10% BioRad Chelex (Estoup et al., 1996).

Primer design and PCR amplification

We used the recently assembled *D. suzukii* genome (Chiu et al., 2013; project accession number PRJNA213258) as an input for the QDD program (Méglec et al., 2014) to scan genomic data for microsatellite regions. Briefly, QDD uses four scripts (Sequence similarity, BLAST, microsatellite detection and markers design) to analyze raw data and provide a table summarizing, amongst other, the region of the genome containing a potential microsatellite sequence, the repeat motif and the number of repeats. We decided to further filter the results of QDD by choosing only continuous (i.e. non-interrupted) microsatellite sequences, with at least 10 repetitions of the motif and excluding (CG)_n and (AT)_n. To maximize the chances to obtain statistically unlinked loci, we excluded microsatellite loci belonging to the same *D. suzukii* scaffold and blasted the selected sequences to the *Drosophila melanogaster* (Meigen) genome to infer chromosomal location. We used the program PRIMER DESIGNER (version 2.0; Scientific & Educational Software 1990/1991) to design forward and reverse primers for the chosen microsatellite loci.

Each pair of primers was used for PCR amplification in a 10 µL final volume containing 1 X QIAGEN Multiplex Master Mix, 2 µM of each primer, and 2 µL of genomic DNA. PCR programs were set with an initial period of denaturation at 94°C (30 s) followed by 32 cycles of additional denaturation at 94°C (30 s), an annealing phase at 57°C (1 min 30 s), an elongation phase at 72°C (1 min), and ending with another extension phase at 72°C (30 min).

We first used 1.5% agarose gel for migration of the PCR products. Loci for which amplifications produced only one DNA band corresponding to the predicted size of the PCR product were then amplified following the protocol described above but with forward primers fluorophore-labelled on the 5' end and non-labelled reverse primers. These loci were finally sized using an ABI 3730 sequencer (Applied Biosystems, Montpellier, France) with the 500 LIZ™ GeneScan™ size standard and scored with the GENEMAPPER™ 5.0 software.

Genetic analyses

All loci were checked for null-alleles using FreeNa (Chapuis & Estoup, 2007; Chapuis et al., 2008) and MicroChecker (van Oosterhout et al., 2004) in the Hawaii and Montpellier populations.

Levels of polymorphism and departure from Hardy-Weinberg (H-W) equilibrium were assessed for both populations. Linkage disequilibrium (LD) was tested for each pair of loci using GENEPOP 4.2 (Raymond & Rousset, 1995). F_{ST} values were estimated using both GENEPOP 4.2 and FreeNa (with the ENA correction for null alleles implemented in the program). To test for loci under selection, we used the DetSel package (Vitalis et al., 2012) for R (R Development Core Team, 2008). DetSel identifies markers showing deviation from the neutral expectation in pairwise comparisons of diverging populations (Vitalis et al., 2001). We used the following parameters for our analysis (details about the user's parameter options can be found in Vitalis et al., 2012): number of simulation points = 1.10^6 , mutation rate = 5.10^{-5} , time of divergence = 5000 and effective size (N_e) = 500,000. We used a personal program to construct a Neighbor-Joining (NJ) dendrogram from individual genotypes using a variant of Chakraborty and Jin's allele-shared distance defined in Fournier et al. (2005).

We studied cross-species PCR amplification of our set of microsatellite loci, using the same PCR conditions as above, on individuals from four species belonging to the *suzukii* subgroup: *D. biarmipes* (Malloch), *D. subpulchrella* (Takamori & Watabe), *D. lucipennis* (Lin) and *D. mimetica* (Bock & Wheeler). Each sister-species was represented by a single isofemale line from which 12 individuals were genotyped (samples obtained from the UC San Diego *Drosophila* stock center).

RESULTS

A large number of regions containing continuous microsatellites located on different scaffolds of the *D. suzukii* genome assembly were identified using the QDD software. We ensured that the chosen loci were not aggregated on the *D. suzukii* genome by first BLAST searching them using the *D. melanogaster* genome. We then chose 32 loci matching two autosomal chromosomes (the third autosomal chromosome is of very small size in *Drosophila* species) and different regions within each chromosome, with a final distribution of 8 loci per chromosome arm. Thirty two pairs of primers, delimiting the 32 chosen loci, were designed for PCR amplifications. Twenty eight such markers were robustly sized and genotyped with GENEMAPPER™ 5.0 and used for further analyses.

Genetic diversity

Table 1 shows that all 28 microsatellite loci were polymorphic with allele numbers ranging from 2 (DS19, DS34, DS11) to 11 (DS32) in the pooled dataset (France + Hawaii). Observed heterozygosity (H_o) ranged from 0.19 (DS21) to 0.84 (DS14) in the Hawaii population and 0.12 (DS21) to 0.88 (DS05) in the French population. Only one locus (DS21) showed significant departure ($P < 0.05$) from H-W equilibrium in the Hawaii population and two loci (DS21 and DS25) in the French population. Significant LD was found for one pair of loci in the Montpellier population and for 14 pairs in the Hawaii population. All pairs of loci presenting significant LD were putatively located on the same chromosome arm. Only one pair (DS38–DS39) showed significant LD in both populations.

Testing for null alleles and selection

FreeNa pointed to five loci with null allele frequencies $> 10\%$ in at least one population (DS09, DS25, DS06, DS26 and DS21). The same loci were identified as bearing null alleles in at least one population by MicroChecker. Of these five loci, only one (DS21) showed high null allele frequencies in both populations and was therefore left out when testing for selection. DetSel did not detect any microsatellite locus that departed significantly from neutral expectations.

Population differentiation

We found a significant and substantial level of genetic differentiation between the populations from France and Hawaii ($F_{ST} = 0.206$). In agreement with this, the NJ tree using individuals as units shows a clear clustering of individuals in relation to their geographic origin (Fig. 1). Moreover genetic assignment tests significantly assigned all genotyped individuals to their population of origin (results not shown).

Cross-species PCR amplification

Table 2 shows that most loci (24 out of 28) successfully amplified in both *D. biarmipes* and *D. subpulchrella*. Although their profiles were generally similar to those observed with *D. suzukii*, most loci were homozygous and exhibited substantially different allele sizes compared to *D. suzukii*. Cross-amplification success was lower in *D. mimetica* and *D. lucipennis* (16 and 14 loci, respectively), which was expected according to the further ge-

TABLE 1. Primer sequences and diversity index for 28 microsatellite loci in two invasive populations of *Drosophila suzukii*. A – allele number, *Ho* – observed heterozygosity; *He* – expected heterozygosity, N – number of sampled individuals, * – significant departure from H-W equilibrium ($P < 0.05$), *D.mel.* – position of homologous loci on *D. melanogaster* chromosomes.

Locus	Repeat	<i>D.mel.</i>	Primer sequence 5'→3'	Dye	Size range	Hawaii (N = 33)			Montpellier (N = 27)		
						A	<i>Ho</i>	<i>He</i>	A	<i>Ho</i>	<i>He</i>
DS05	(TG)10	Chr 2L	AGGATAACGCGCAGCTTGAC TATGGAAGCTGGCAAGCAGA	Ned	260–300	6	0.8387	0.6967	4	0.8800	0.7584
DS06	(TG)11	Chr 2L	CGGTTTCGAGTGCTTGTTAGA ACACGTGGAGGACACCTTC	Fam	130–190	4	0.6061	0.6056	3	0.3333	0.4856
DS07	(CA)13	Chr 2L	AAGGCTGGAGTGGCAACAA GCTAAGGTTCTGTTCCGGCTG	Ned	160–210	3	0.4545	0.5147	7	0.8148	0.8278
DS08	(AG)10	Chr 2L	CGTTGTTGGCGGTGAGTAAG GGCCATCAATCAGTCAGTCA	Vic	110–170	9	0.6970	0.7222	4	0.8148	0.7565
DS09	(AC)15	Chr 2L	CACACATGGCGTATGCGTAT ACTTGTGAGCCGTCCTGG	Fam	190–250	4	0.4848	0.6772	8	0.6538	0.6649
DS10	(GT)11	Chr 2L	CGAGACTGTGCGAACGAGAG CATATGCTGACTGCCTCACA	Ned	270–330	5	0.3939	0.5124	6	0.6538	0.6294
DS11	(CA)11	Chr 2L	CGGTGACTCGTGCAGTTGTA GCCGACTCTGTCTAGAGCAA	Vic	230–290	2	0.4545	0.3512	3	0.4074	0.4753
DS12	(GT)19	Chr 2R	GCTGTTGCTGTTGCTATTGC AGAACCGTTAGCTGAGCGAG	Fam	320–380	6	0.8438	0.7866	6	0.6538	0.7167
DS14	(TG)10	Chr 2R	AAGAACC GCAACGAGCAA GAATTATCCAGCGACACGAC	Fam	180–220	7	0.8485	0.7374	6	0.5833	0.5755
DS15	(GT)11	Chr 2R	GGACAGCCGACATAAGAGG GAGTTGCTGGCTCGACACTT	Fam	260–320	5	0.5152	0.5505	7	0.8462	0.6553
DS16	(AC)13	Chr 2R	TTCGTATGTTAGGCGCCA CTGGCTGCTGACCTCAACTC	Pet	100–140	3	0.6061	0.6074	6	0.7308	0.7234
DS17	(GT)10	Chr 2R	CATCTCAGGCCACGAATG CTCCAGATTCTCGAGTGCAG	Fam	80–130	3	0.3939	0.3742	5	0.7407	0.7064
DS19	(CT)10	Chr 2R	CCGTTGGCATCTCTGAGTCT GCAGACGGAGAGCAGTTGTT	Vic	180–220	2	0.3636	0.3967	5	0.5556	0.5734
DS20	(AG)12	Chr 3L	CAGCCATATGCAATGCACTG ATATCCAGCGGAAGTCGAGA	Ned	210–270	6	0.6250	0.6973	4	0.7083	0.7405
DS21*	(AC)11	Chr 3L	GAGACGCGATGGTACCGTTA CCAATCGAGTGCAAGCGT	Vic	310–370	4	0.1935	0.4886*	3	0.1176	0.5519*
DS22	(GT)11	Chr 3L	TACAGATACCCGTCGGATT AAGACCAAGACGACGGACCT	Pet	290–360	6	0.5938	0.6948	5	0.6667	0.6979
DS23	(AC)10	Chr 3L	TGCCACTAAGCTCACACGGT CAGTTGCCACTTGCTGTGTA	Pet	237–300	6	0.6129	0.6145	3	0.7917	0.7283
DS25*	(CA)10	Chr 3L	CCTTCTGCTGCTCCTGCTTAT CGGTTCGAGCATCAAGTGAA	Vic	220–270	5	0.6667	0.6644	4	0.4444	0.7291*
DS26	(CA)10	Chr 3L	CCTGTGTGCATCTCAGTGTT TACAGCACTCCAGCACATGA	Vic	60–130	4	0.3548	0.6113	4	0.7037	0.7277
DS27	(GT)14	Chr 3L	CCAGCGACTGCAGAAGTGAC GCAATCCTCCACAACAAC	Ned	80–130	6	0.7576	0.7034	5	0.7037	0.6893
DS28	(TG)11	Chr 3L	TTAAGCTGACCTCCTCCTCG GCACTCGCACAGATAACAAGG	Pet	140–195	4	0.5758	0.5822	7	0.8400	0.7896
DS32	(TG)15	Chr 3R	CGGCGTGTGTCAGTTATTC ATGCACTGGTCGACATGACA	Fam	330–380	4	0.8333	0.8073	10	0.7813	0.6685
DS33	(AC)10	Chr 3R	GTTGGATGCCTGGAGGAATA TTGTGCGAGTATGTCAGCCA	Pet	150–210	3	0.6364	0.6318	5	0.8462	0.7655
DS34	(GA)12	Chr 3R	AACAACGACGCAGAAGCTCA CGACTGTTGCGCTCTGTAAT	Fam	240–300	5	0.4848	0.4444	2	0.5417	0.5634
DS35	(CA)11	Chr 3R	TCCGTATCCGTATCCGTGT GGAGTATGGCAGTGTGGCAG	Pet	198–240	8	0.4688	0.5171	3	0.8333	0.7925
DS36	(GT)13	Chr 3R	TTGGCAACGTGTGAAGCTG GAGACACTGCAATGCTGCCT	Vic	160–220	4	0.7879	0.7199	4	0.4074	0.5158
DS38	(AC)12	Chr 3R	CTGTCCGGCTACGCAATATTC CAACGGATTGTGGCTGATTG	Ned	130–180	7	0.5758	0.4835	3	0.8148	0.7833
DS39	(AC)12	Chr 3R	GCCAGGCAGCAACAACATC AGATCCGAGAGTTGCGAGTT	Ned	310–370	6	0.7188	0.7466	5	0.7407	0.7257

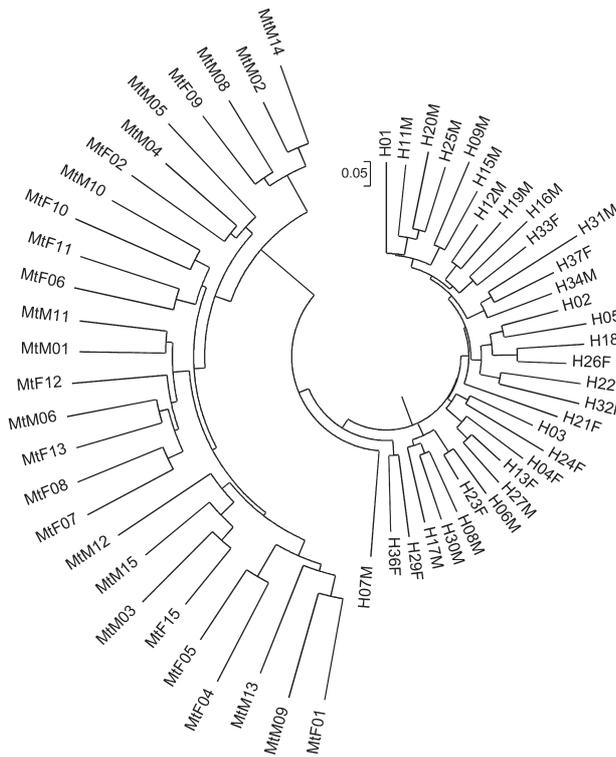


Fig. 1. NJ dendrogram of the microsatellite allele shared distances between *D. sukukii* individuals. Mt – French population; H – Hawaii population; M/F – male or female fly (when known).

netic distance of the two species from *D. sukukii* compared to *D. biarmipes* and *D. subpulchrella* (Kopp & True, 2002). Because the individuals used for cross-amplification tests originated from stock-center populations (with a single isofemale line for each species), a high level of inbreeding and adaptation to laboratory conditions is suspected for these flies, thereby questioning the reliability of the observed allelic pattern in the non-*sukukii* species.

DISCUSSION

The inferential methods used for retracing invasion routes necessitates that the studied populations display substantial polymorphism and significant genetic differentiation in order to discriminate among potential origins and make inferences about demographic processes (e.g. Estoup & Guillemaud, 2010). The microsatellite markers here developed for *D. sukukii* clearly fit these criteria, at least for the two tested populations. In a recent paper, Adrion et al. (2014) argued that at least 10^5 independent sequence loci would be needed to confidently discriminate among colonization histories in *D. sukukii* using ABC methods. We believe that this pessimistic conclusion should be taken cautiously due to potential methodological issues: (i) the estimation of such ABC power relies on the analysis of pseudo-observed datasets simulated by drawing parameters into prior distributions, hence not conditionally on the observed dataset, and (ii) the number of sequence loci needed to attain 90% power (10^5 loci) derives from a delicate extrapolation exercise based on pseudo-observed datasets simulated with a particularly small number of loci (comprising between 6 and 192 sequence loci, depending on the simulations). Interestingly, the same authors underline the fact that polymorphic markers such as microsatellites should better

TABLE 2. Cross-species amplification results for 28 *D. sukukii* microsatellite loci in four species of the *sukukii* subgroup. Cell values indicate allele numbers and size range (in brackets). N – number of individuals genotyped per species.

Locus	<i>D. biarmipes</i> (N = 12)	<i>D. subpulchrella</i> (N = 12)	<i>D. mimetica</i> (N = 12)	<i>D. lucipennis</i> (N = 12)
DS05	3 (276–282)	3 (276–282)	–	–
DS06	3 (154–189)	3 (154–189)	1 (165)	1 (201)
DS07	2 (179–181)	2 (179–181)	1 (177)	–
DS08	–	–	1 (128)	2 (140–157)
DS09	1 (211)	1 (211)	2 (225–231)	1 (244)
DS10	2 (292–302)	1 (292)	–	–
DS11	1 (251)	2 (251–253)	1 (239)	1 (256)
DS12	2 (326–338)	2 (326–338)	–	–
DS14	2 (189–195)	4 (189–207)	1 (200)	1 (201)
DS15	2 (260–270)	2 (260–270)	2 (316–323)	1 (244)
DS16	1 (111)	1 (111)	1 (119)	1 (100)
DS17	1 (99)	1 (99)	–	–
DS19	1 (139)	1 (139)	–	–
DS20	3 (232–236)	1 (236)	1 (243)	–
DS21	–	–	–	–
DS22	1 (329)	1 (329)	1 (330)	1 (317)
DS23	1 (256)	1 (256)	–	2 (241–271)
DS25	1 (241)	1 (241)	2 (225–231)	–
DS26	2 (82–89)	2 (83–89)	–	–
DS27	2 (84–98)	2 (84–98)	1 (93)	2 (85–87)
DS28	2 (149–178)	2 (149–178)	–	1 (159)
DS32	2 (335–337)	1 (335)	1 (335)	–
DS33	2 (159–165)	2 (159–165)	–	–
DS34	2 (253–257)	3 (251–255)	1 (257)	1 (268)
DS35	–	–	–	–
DS36	2 (175–182)	2 (175–182)	1 (177)	1 (201)
DS38	4 (133–148)	4 (133–148)	2 (126–128)	2 (113–119)
DS39	2 (311–334)	2 (311–334)	–	–

discriminate colonization models and that additional *D. sukuzii* population samples were needed to better discriminate such models. In agreement with this view, microsatellite markers (often analyzed using ABC methods) have proven invaluable for reconstructing invasion routes in many previous studies dealing with various biological models, and most of such studies were based on a larger set of population samples than those used in Adrion et al. (2014) (see for instance Lombaert et al., 2011 and 2014, and references therein). It is also worth noting that a lower number of microsatellite markers (compared to the set of 28 microsatellites we developed for *D. sukuzii*) were used in many of these other studies. We hence believe that our suite of 28 markers represents a molecular tool with high potential when attempting to reconstruct the pathways and demographic processes of invasion in this major dipterous pest, in complement to the sequence data recently produce for *D. sukuzii* populations as detailed in Adrion et al. (2014). We are currently using our suite of microsatellites to genotype a large number of population samples collected at various locations in the native and invasive range of *D. sukuzii* to further refine the invasion history of this species on a worldwide scale.

ACKNOWLEDGEMENTS. This work benefited from state funding by the Agence Nationale de la Recherche, through the LabEx ANR-10-LABX-0003-BCDiv, of the program “Investissements d’avenir” (ref. ANR-11-IDEX-0004-02). Funding was also provided by the French Institute of Agricultural Research (INRA-SPE). We thank two anonymous referees and H.D. Loxdale for their helpful editorial comments.

REFERENCES

- ADRION J.R., KOUSATHANAS A., PASCUAL M., BURRACK H.J., HADDAD N.M., BERGLAND A.O., MACHADO H., SACKTON T.B., SCHLENKE T.A., WATADA M., WEGMANN D. & SINGH N.D. 2014: *Drosophila sukuzii*: the genetic footprint of a recent, worldwide invasion. — *Mol. Biol. Evol.* **31**: 3148–3163.
- CALABRIA G., MÁCA J., BÄCHLI G., SERRA L. & PASCUAL M. 2012: First records of the potential pest species *Drosophila sukuzii* (Diptera: Drosophilidae) in Europe. — *J. Appl. Entomol.* **136**: 139–147.
- CHAPUIS M.P. & ESTOUP A. 2007: Microsatellite null alleles and estimation of population differentiation. — *Mol. Biol. Evol.* **24**: 621–631.
- CHAPUIS M.P., LECOQ M., MICHALAKIS Y., LOISEAU A., SWORD G.A., PIRY S. & ESTOUP A. 2008: Do outbreaks affect genetic population structure? A worldwide survey in *Locusta migratoria*, a pest plagued by microsatellite null alleles. — *Mol. Ecol.* **17**: 3640–3653.
- CHIU J.C., JIANG X., ZHAO L., HAMM C.A., CRIDLAND J.M., SAELAO P., HAMBY K.A., LEE E.K., KWOK R.S., ZHANG G., ZALOM F.G., WALTON V.M. & BEGUN D.J. 2013: Genome of *Drosophila sukuzii*, the Spotted Wing *Drosophila*. — *G3* **3**: 2257–2271.
- CINI A., ANFORA G., ESCUDERO-COLOMAR L.A., GRASSI A., SANTOSUOSSO U., SELJAK G. & PAPINI A. 2014: Tracking the invasion of the alien fruit pest *Drosophila sukuzii* in Europe. — *J. Pest Sci.* **87**: 559–566.
- ESTOUP A. & GUILLEMAUD T. 2010: Reconstructing routes of invasion using genetic data: why, how and so what? — *Mol. Ecol.* **19**: 4113–4130.
- ESTOUP A., LARGIADÈR C.R., PERROT E. & CHOURROUT D. 1996: One-tube rapid DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. — *Mol. Mar. Biol. Biotechnol.* **5**: 295–298.
- ESTOUP A., WILSON I.J., SULLIVAN C., CORNUET J.M. & MORITZ C. 2001: Inferring population history from microsatellite and enzyme data in serially introduced cane toads, *Bufo marinus*. — *Genetics* **159**: 1671–1687.
- FOURNIER D., ESTOUP A., ORIVEL J., FOUCAUD J., JOURDAN H., LE BRETON J. & KELLER L. 2005: Clonal reproduction by males and females in the little fire ant. — *Nature* **435**: 1230–1235.
- HAUSER M. 2011: A historic account of the invasion of *Drosophila sukuzii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. — *Pest Manag. Sci.* — **67**: 1352–1357.
- KIMURA M. 2004: Cold and heat tolerance of drosophilid flies with reference to their latitudinal distributions. — *Oecologia* **140**: 442–449.
- KOPP A. & TRUE J.R. 2002: Evolution of male sexual characters in the oriental *Drosophila melanogaster* species group. — *Evol. Devel.* **4**: 278–291.
- LOMBAERT E., GUILLEMAUD T., THOMAS C., LAWSON HANDLEY L.J., LI J., WANG S., PANG H., GORYACHEVA I.I., ZAKHAROV I.A., JOUSSELIN E., POLAND R., MIGEON A., VAN LENTEREN J.C., DE CLERCQ P., BERKVENNS N., JONES W. & ESTOUP A. 2011: Inferring the origin of populations introduced from a genetically structured native range by approximate Bayesian computation: case study of the invasive ladybird *Harmonia axyridis*. — *Mol. Ecol.* **20**: 4654–4670.
- LOMBAERT E., GUILLEMAUD T., LUNDGREN J., KOCH R., FACON B., GREZ A., LOOMANS A., MALAUSA T., NEDVED O., RHULE E., STAVERLOK A., STEENBERG T. & ESTOUP A. 2014: Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. — *Mol. Ecol.* **23**: 5931–6205.
- MATSUMURA S. 1931: *6000 Illustrated Insects of Japan-Empire*. Tokohshoin, Tokyo, 1497 pp. [in Japanese].
- MEGLÉCZ E., PECH N., GILLES A., DUBUT V., HINGAMP P., TRILLES A., GRENIER R. & MARTIN J.F. 2014: QDD version 3.1: a user-friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. — *Mol. Ecol. Res.* **14**: 1302–1313.
- PARSHAD R. & PAIKA I.J. 1965: Drosophilid survey of India II. Taxonomy and cytology of the subgenus *Sophophora* (*Drosophila*). — *Res. Bull. Panjab Univ.* **5**: 225–252.
- RAYMOND M. & ROUSSET F. 1995: GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. — *J. Hered.* **86**: 248–249.
- THRESHER R.E., HAYES K., BAX N.J., TEEM J., BENFEY T.J. & GOULD F. 2014: Genetic control of invasive fish: technological options and its role in integrated pest management. — *Biol. Invas.* **16**: 1201–1216.
- VAN OOSTERHOUT C., HUTCHINSON W.F., WILLS D.P.M. & SHIPLEY P. 2004: Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. — *Mol. Ecol. Notes* **4**: 535–538.
- VITALIS R., DAWSON K. & BOURSOT P. 2001: Interpretation of variation across marker loci as evidence of selection. — *Genetics* **158**: 1811–1823.
- VITALIS R. 2012: DetSel: a R-package to detect marker loci responding to selection. In Pompanon F. & Bonin A. (eds): *Data Production and Analysis in Population Genomics: Methods and Protocols*. Springer/Humana Press, New York, pp. 277–293.

Received December 3, 2014; revised and accepted April 21, 2015
 Prepublished online July 21, 2015

Article 2.

Frainout, A., Debat, V., Fellous, S., Hufbauer, R., Foucaud, J., Pudlo, P., Marin, J.-M., Price, D.K., Cattell, J., Chen, X., Deprá, M., Duyck, P.-F., Guedot, C., Loeb, G., Kenis, M., Kimura, M.T., Loiseau, A., Martinez-Sañudo, I., Pascual, M., Polihronakis Richmond, M., Shearer, P., Singh, N., Tamura, T., Xuéreb, A., Zhang, J. & Estoup, A. Deciphering the invasion routes of *Drosophila suzukii* by means of ABC random forest. *Mol. Biol. Evol.* submitted

Article

Methods

Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest

Antoine Fraimout¹, Vincent Debat¹, Simon Fellous², Ruth Hufbauer^{2,3}, Julien Foucaud², Pierre Pudlo⁴, Jean-Michel Marin⁵, Donald K. Price⁶, Julien Cattel⁷, Xiao Chen⁸, Marindia Deprá⁹, Pierre François Duyck¹⁰, Christelle Guedot¹¹, Gregory Loeb¹², Marc Kenis¹³, Masahito T. Kimura¹⁴, Isabel Martinez-Sañudo¹⁵, Marta Pascual¹⁶, Maxi Polihronakis Richmond¹⁷, Peter Shearer¹⁸, Nadia Singh¹⁹, Koichiro Tamura²⁰, Anne Xuéreb², Jinping Zhang²¹, Anne Loiseau², & Arnaud Estoup².

Corresponding author: Arnaud Estoup, UMR Centre de Biologie pour la Gestion des Populations (INRA / IRD / Cirad / Montpellier SupAgro), 755 avenue du Campus Agropolis, CS 30 016, 34988 Montferrier-sur-Lez cedex, France; email: estoup@supagro.inra.fr

Short title: Deciphering invasion routes using ABC random forest

Key words: *Drosophila suzukii*, invasion routes, random forest, approximate Bayesian computation, population genetics.

Abstract

Deciphering invasion routes from molecular data is crucial to understanding biological invasions, including identifying bottlenecks in population size and admixture among distinct populations. Approximate Bayesian computation (ABC) methods have improved reconstruction of invasion routes, but are computationally intensive. A new approach, ABC random forest (ABC-RF; Pudlo *et al.* 2016), compares invasion scenarios at lower computational cost, and is more flexible than standard ABC methods with respect to the summary statistics required. Here, we illustrate the use of ABC-RF, and compare it to a more standard ABC method, using a multi-locus microsatellite dataset (25 loci on 23 worldwide sampling locations) on the invasive pest *Drosophila suzukii*. We unravel the invasion routes of this species, and evaluate signatures of demographic bottlenecks and genetic admixture during the invasion. We show that the original invasion of North America by *D. suzukii* was distinct from that of Europe. Populations in western North America originate from both southeast China and Hawaii (the latter originating from Japan), and those in the east from western North America. European populations are genetically more homogeneous and originate from northeast China, with limited gene flow from the eastern US. All populations passed through bottlenecks, and analyses reveal five distinct admixture events. These findings can be used to develop and test rigorous hypotheses concerning this invasive species. Methodological comparisons indicate that ABC-RF considerably out-performs standard ABC when using a low number of simulated datasets, especially when analyzing complex settings (e.g. when comparing more than ten scenarios using more than 100 summary statistics).

Introduction

Biological invasions are a component of global change, and their impact on the communities and ecosystems they invade is substantial (Simberloff 2013). Invasive populations evolve rapidly, via both neutral and selective evolutionary processes, as they undergo dramatic range expansion, demographic reshuffling and experience new selection regimes (e.g. Keller & Taylor 2010; Dlugosch *et al.* 2015). While extensive literature has laid the foundations of an evolutionary framework of biological invasions (e.g. Sakai *et al.* 2001; Lee 2002; Facon *et al.* 2006; Dlugosch *et al.* 2015; Estoup *et al.* 2016), whether evolutionary processes are drivers of invasion success or outcomes of introduction processes remains unclear.

Two of the evolutionary processes hypothesized to play a role in invasions are demographic bottlenecks and genetic admixture (Ellstrand & Shierenbeck 2000; Dlugosch & Parker 2008; Rius & Darling 2014). Demographic bottlenecks, which correspond to transitory reductions in population size associated with invasions, can reduce genetic variation and thus may constrain invasion success (Edmonds *et al.* 2004, Dlugosch & Parker 2008, Peischl & Excoffier 2015). A diverse array of mechanisms allows bottlenecked populations to overcome the deleterious consequences of low genetic variation and adapt to their novel environments (reviewed in Estoup *et al.* 2016). Genetic admixture occurs when multiple introduction events derive from genetically differentiated native or invasive populations. Admixture can increase genetic diversity in introduced populations, potentially enhancing invasion success by increasing genetic variation on which selection can act (Kolbe *et al.* 2004; Lavergne & Molofsky 2007) or via producing entirely novel genotypes that may facilitate colonization of novel habitats (Dlugosch & Parker 2008, Rius & Darling 2014). It remains unclear, however, how often genetic admixture occurs in invasions and whether it acts as a true driver of invasion success (Uller & Leimu 2011, Rius & Darling 2014).

To evaluate hypotheses regarding bottlenecks and admixture, as well as adaptation or other processes that may occur during and after introductions, we must first identify the original native or invasive source(s) of the introductions. It is these source region(s) that provide the relevant benchmark against which genetic and evolutionary changes should be assessed (Dlugosch & Parker 2008, Keller & Taylor 2010). Typically, accurate historical data on introductions pathways are limited (Estoup & Guillemaud 2010). Traditional population genetic approaches can describe the genetic structure and relationships among sampled native and invasive populations (Boissin *et al.* 2012), but determining origins of introduced species from population genetic data is challenging. This challenge has prompted the development of new statistical methods, and in particular, approximate Bayesian computation (ABC; Beaumont *et al.* 2002). In ABC analysis, different introduction scenarios describe a set of possible introduction pathways from a given native or invaded area to an invaded region. The introduction pathways can be realistically complex. For example, rather than a species moving directly from the native to invasive range, it might first pass

through another region, experiencing bottlenecks and/or admixture along the way. These introduction scenarios are then compared quantitatively, taking a model-based Bayesian approach in which datasets matching different scenarios are generated by simulation, and the relative likelihoods of the models estimated (Bertorelle *et al.* 2010; Beaumont 2010). ABC has been applied successfully to a wide range of case studies, from pest invasion (e.g. *Harmonia axiridis*, Lombaert *et al.* 2014) to anthropological research (e.g. *Homo sapiens*, Verdu *et al.* 2009) and is now widely used by the population genetics community. However, and despite the many advantages of ABC, it relies on massive simulations that can render it computationally very costly. When multiple complex introduction scenarios are considered, the time and computational resources required to reliably choose among them can become prohibitive (Lombaert *et al.* 2014). To overcome this constraint, a new algorithm called ABC random forest (ABC-RF) has been developed (Pudlo *et al.* 2016). ABC-RF allows more efficient discrimination among scenarios than traditional ABC methods, and thus appears particularly suitable for distinguishing between complex invasion pathways (see section below *New approaches*).

Here, we evaluate the origins of invasive populations of spotted-wing *Drosophila suzukii* (Matsumura 1931), including where and whether they passed through bottlenecks in populations size, or experienced admixture between genetically distinct groups. We use this invasion to illustrate the use of ABC-RF, and as a case study to compare it to a more standard ABC method. *Drosophila suzukii* is a representative of the *melanogaster* group and is historically distributed in Southeast Asia, covering a large portion of China, Japan, Thailand and neighboring countries (Asplen *et al.* 2015). Unlike most *drosophilid* species, *D. suzukii* females display a large serrated ovipositor (Atallah *et al.* 2014) allowing them to lay eggs in ripening fruits, and are thus a major threat to the agricultural production of stone fruits and berries (Lee *et al.* 2011). Presence of the species outside of its native range was first recorded in the Hawaiian archipelago in the early 1980's (Kaneshiro 1983) but no further spread was observed until the late 2000's when *D. suzukii* was almost synchronously recorded in the southwest of the USA and southern Europe (Hauser 2011, Calabria *et al.* 2012). By 2015, the species had rapidly spread throughout most of the North-American and European continents and reached the southern hemisphere in Brazil (Deprá *et al.* 2014). *Drosophila suzukii* represents a particularly appealing biological model for gaining further insights into the evolutionary factors associated with invasion success. As a closely related species of the model insect *Drosophila melanogaster*, its short generation time and viability in laboratory conditions facilitate experimental approaches to study evolutionary events.

Evidence from DNA sequencing of six X-linked gene fragments suggests that North American and European populations represent separate invasion events (Adrion *et al.* 2014). The sources of these invasions and potential admixture among different regions remain unclear. The wide native and introduced distribution of *D. suzukii* imply a potentially large number of possible sources and hence a large number of possible introduction scenarios. By first grouping sample

size into genetic clusters, the number of populations treated independently can be reduced somewhat (e.g. Lombaert *et al.* 2014), but not fully. Additionally the temporal proximity of invasions in the US and Europe makes it difficult to a priori propose certain introduction scenarios over others. Together, these features of *D. sukuzii*'s invasion mean that a large number of scenarios will need to be compared, dramatically increasing computational requirements for traditional ABC methods, and hence making ABC-RF methods particularly appealing.

New approaches

A crucial step in determining the most probable introduction pathway is to formally describe a finite set of possible introduction scenarios as models and to compare them quantitatively (Estoup & Guillemaud 2010). Choosing among these models is thus the central statistical problem to be overcome when reconstructing routes of invasion from molecular data. Both theoretical arguments and simulation experiments indicate that posterior probabilities estimated from ABC analyses for the modeled introduction scenarios can be inaccurate, even though the models being compared can still be ranked appropriately using numerical approximation (Robert *et al.* 2011). To overcome this problem, Pudlo *et al.* (2016) developed a novel approach based on a machine learning tool named 'random forests' (RF; Breiman 2001), which selects among the complex introduction scenarios covered by ABC algorithms. This approach enables efficient discrimination among scenarios while being computationally less intensive. It hence is particularly appealing for invasions involving multiple populations and multiple potential sources, as is found in the recent worldwide spread of *D. sukuzii*.

In ABC random forest (ABC-RF), Pudlo *et al.* (2016) make two important advances regarding the use of summary statistics, and the identification of the most probable model. For the summary statistics, given a pool of different metrics available, rather than selecting from among them, ABC-RF optimally extracts information from the entire set of proposed statistics using machine-learning methodology (Breiman 2001). With respect to identifying the most probable model, ABC-RF uses a classification vote system rather than the posterior probabilities traditionally used in ABC analysis (which are known to have problems with accuracy; Robert *et al.* 2011). Specifically, in ABC-RF the model that best fits the observed dataset is determined by constructing a machine learning classifier from simulations recorded in a table called the reference table. The reference table includes a given number of datasets that have been simulated for different scenarios using parameter values drawn from prior distributions, each dataset being summarized with a pool of statistics. The outcome when applied to a given target dataset is a classification vote for each competing model. The classification vote represents the number of times a model is selected in a forest of n classification trees. The model with the highest number of classification votes corresponds to the model best suited to the target dataset among the set of competing

models. As a byproduct, this step also provides a measure of the classification error called the prior error rate. The prior error rate is calculated as the probability of choosing a wrong model when drawing model index and parameter values into priors. Only after the best model has been selected, its posterior probability is estimated through a secondary random forest that regresses the model selection error of the first-step random forest over the available summary statistics. See Pudlo *et al.* (2016) for detailed algorithms.

As compared to previous ABC methods for choosing among models, ABC-RF offers at least four advantages (Pudlo *et al.* 2016): (i) it significantly reduces the model classification error as measured by the prior error rate; (ii) it is robust to the number and choice of summary statistics, as RF can handle many superfluous and/or strongly correlated statistics with no impact on the performance of the method; (iii) the computing effort is considerably reduced as RF requires a much smaller reference table compared with alternative methods (i.e., a few thousands of simulated datasets versus hundreds of thousands to millions of simulations per compared model); and (iv) it provides more reliable estimation of posterior probability of the selected model (i.e., the model that best fit the observed dataset). To our knowledge, the present study is the first to use the ABC-RF method for an ensemble of model choice analyses with different levels of complexity (i.e., various number of compared models including various and sometimes large number of parameters, sampled populations and hence number of summary statistics) on real multi-locus microsatellite datasets.

In the present paper, we analyze molecular data at 25 microsatellite loci from 23 *D. sukukii* sample sites located across most of its native and introduced range. We conducted our population genetics study in five steps. 1/ We defined focal genetic groups, and used this information to formalize eleven sets of competing introduction scenarios. We then choose among them in a sequential ABC-RF analysis. 2/ We compared the performance of ABC-RF to a more standard ABC method (ABC-LDA; Estoup *et al.* (2012); see Materials & Methods). We choose ABC-LDA in our comparative study because it is considered as one of the most efficient standard ABC methods to discriminate among models (Pudlo *et al.* 2016) and it is widely used in population genetics studies (e.g. Lombaert *et al.* 2014 and reference therein). 3/ We refined the inferred worldwide invasion scenario by assessing the most likely origins of the primary introductions from the native area. 4/ We estimated the posterior distributions under the final invasion scenario of demographic parameters associated with bottleneck and genetic admixture events. 5/ Finally, we performed model-posterior checking analyses to insure that the final worldwide invasion scenario did not display major inadequacy with the observed dataset.

Results

Origins of invasive genetic groups

Prior to conducting any type of ABC analysis, focal genetic groups must be defined by characterizing genetic diversity and structure within and between all genotyped sample sites using traditional statistics and clustering methods. This is described in details in Appendix S1. We defined seven main genetic groups from our initial set of 23 sample sites (table S1): Asia (sample sites CN-Lan, CN-Lia, CN-Nin, CN-Shi, JP-Tok and JP-Sap), Hawaii (sample site US-Haw), western US (sample sites US-Wat, US-Sok and US-SD), eastern US (sample sites US-Col, US-NC, US-Wis and US-Gen), Europe (sample sites GE-Dos, FR-Par, FR-Bor, FR-Mon, SW-Del, SP-Bar and IT-Tre), Brazil (BR-PA) and La Réunion (FR-Reu).

This genetic grouping along with historical and geographical information (table S1) allowed us to formalize eleven nested sets of competing invasion scenarios that we analyzed sequentially using ABC model choice methodologies (table 1; see also table S2). The eleven sequential ABC analyses permit to reconstruct step-by-step the introduction history of *D. sukuzii* worldwide. Results for each analysis based on prior set 1 (which used bounded uniform distributions of parameters – see Materials and Methods) and a single set of sample sites representative of the genetic groups defined above are given in table 2. For all eleven analyses, similar ABC-RF results were obtained using the prior set 2 (which used more peaked distributions; table S3) and considering various sets of representative sample sites (table S4).

Following the dates of first observation available for *D. sukuzii* after the invasion of Hawaii, we first deciphered the invasion pathways of the genetically heterogeneous western US (analyses 1a to 1d, table 2). Considering the three western US sample sites in separated analyses, we found in each case that the best scenario included genetic admixture between Asia and Hawaii as sources (table 2; analysis 1a to 1c, mean $p = 0.999$). We subsequently tested if this common admixture pattern resulted from a single or multiple introduction events from Asia and/or Hawaii, including the possibility of local colonization events among western US sites (analysis 1d). The best invasion scenario for the western US included an initial admixture event with Asia and Hawaii as sources in northern California (site US-Wat; admixture event A1 in fig. 1), followed by (i) a local colonization southward into San Diego (US-SD) and northward into Oregon state (US-Sok), and (ii) a secondary introduction event from Hawaii to Oregon (US-Sok) (A2 event in fig. 1; $p = 0.690$). This result is supported by various genetic clustering results (fig. 1, fig. S1 and S2) as well as raw F -statistics values which indicate high differentiation level between US-Sok and other continental US sample sites and a substantially lower F_{ST} between US-Sok and Hawaii (table S5).

Once the invasion history was resolved for the western US group, we inferred the most likely source(s) of the invasive eastern US and European genetic groups, using two independent ABC analyses (analyses 2a and 2b). We found a single origin for the eastern US group corresponding to an intra-continental spread from San Diego ($p = 0.779$, table 2). The most likely origin for Europe was an independent introduction from the Asian native range ($p = 0.716$, table 2).

Model choice analysis	Number of compared scenarios	Tackled question	Potential source genetic group	Focal populations
1a 1b 1c	3	What are the origins of western US populations?	Asia, Hawaii	US-Wat US-Sok US-SD
1d	7	What are the relations among western US populations and their extra-continental sources?	Asia, Hawaii, US-Wat, US-Sok, US-SD	US-Wat + US-Sok + US-SD
2a	6	What are the origins of eastern US populations?	Asia, Hawaii, western US	eastern US
2b	6	What are the origins of European populations?	Asia, Hawaii, western US	Europe
3a	10	Is there asymmetrical gene flow from Europe to eastern US ?	Asia, Hawaii, western US, Europe	eastern US
3b	10	Is there asymmetrical gene flow from eastern US to Europe?	Asia, Hawaii, western US, eastern US	Europe
4	2	Does the admixture with eastern US genes in northern Europe result from a secondary Asian introduction?	Asia, Hawaii, western US, eastern US, southern Europe	northern Europe
5a	21	What are the origins of the Brazilian population?	Asia, Hawaii, western US, eastern US, southern Europe, northern Europe	Brazil
5b	21	What are the origins of La Reunion population?	Asia, Hawaii, western US, eastern US, southern Europe, northern Europe	La Reunion

Table 1. Formulation of the model choice analyses carried out successively to reconstruct invasion routes of *D. sukii* using ABC.

Note: Each numbered analysis is a comparison of a certain number of scenarios by ABC model choice. We summarize each analysis by stating the question that it addressed. A detailed verbal description of each compared scenario is given in table S2. The eleven analyses are nested in the sense that each subsequent analysis use the result obtained from the previous one. For example in Analysis 2a “What are the origins of eastern US populations” capitalizes on the history inferred for western US populations from analysis 1d. “Potential source genetic group” indicates all the potential source populations considered in the analyses for which one wants to identify the origin of the focal population (i.e. the target).

Analysis	Total n° of scenarios	Number of Sum. Stats.	prior error rate			Posterior probability of the best model		Origin of the focal population using either ABC-RF or ABC-LDA (i.e. best model)
			ABC-RF (s.d.)	ABC-LDA (large reference table)	ABC-LDA (small reference table)	ABC-RF (s.d.)	ABC-LDA (large reference table)	
1a	3	39	0.100 (\pm 0.002)	0.075	0.085	1.000 (\pm 0.001)	1.000 [1.000, 1.000]	Asia + Hawaii
1b	3		0.094 (\pm 0.001)	0.070	0.091	0.989 (\pm 0.005)	0.999 [0.999, 0.999]	
1c	3		0.096 (\pm 0.001)	0.079	0.087	0.998 (\pm 0.002)	0.999 [0.999, 0.999]	
1d	7	130	0.327 (\pm 0.001)	0.274	0.368	0.690 (\pm 0.018)	0.741 [0.714, 0.767]	Asia + Hawaii
2a	6	130	0.232 (\pm 0.001)	0.203	0.241	0.779 (\pm 0.019)	0.788 [0.764, 0.813]	western US
2b	6	130	0.232 (\pm 0.001)	0.179	0.393	0.716 (\pm 0.013)	0.604 [0.592, 0.616]	Asia
3a	10	204	0.328 (\pm 0.001)	0.265	0.364	0.744 (\pm 0.020)	0.844 [0.820, 0.868]	western US
3b	10	204	0.408 (\pm 0.001)	0.358	0.423	0.510 (\pm 0.014)	0.409 [0.391, 0.426]	Asia
4	2	301	0.121 (\pm 0.001)	0.111	0.212	1.000 (\pm 0.009)	0.999 [0.999, 0.999]	southern Europe + eastern US
5a	21	424	0.294 (\pm 0.001)	0.230	0.423	0.631 (\pm 0.034)	0.812 [0.788, 0.837]	western US + eastern US
5b	21	424	0.298 (\pm 0.001)	0.242	0.428	0.500 (\pm 0.027)	0.290 [0.256, 0.323]	northern Europe + southern Europe

Table 2. Results of model choice analyses using ABC-RF and ABC-LDA.

Note: All model choice analyses were carried out using the prior set 1 (table S8) and a single set of sample sites representative of the pre-defined genetic groups. Sample site JP-Tok for the native Asian group, US-Haw for the Hawaiian group, US-Wat, US-Sok and US-SD for the western US group, US-NC for the eastern US group, IT-Tre for the southern Europe group, GE-Dos for the northern Europe group, BR-PA for the Brazil group, and FR-Reu for the La Réunion group (fig. 1). Datasets were summarized using the whole set of summary statistics proposed by DIYABC (Cornuet *et al.* 2014). The total number of summary statistics (“Number of Sum. Stats.”) as well as the total number of compared scenarios (“Total n° of scenarios”) are indicated for each analysis. Prior error rates and posterior probabilities of the best model chosen using ABC-RF were averaged over ten replicate analyses. ABC-RF and ABC-LDA treatments yielded the same best model choice for all analyses and is denoted in the column “Origin of the focal population using either ABC-RF or ABC-LDA (i.e. best model)”. Admixture between two source populations are represented by a “+” sign. ABC-LDA posterior probabilities of the best models were estimated using “large reference tables” with 500,000 simulated datasets per scenario. ABC-LDA prior error rates were computed using reference tables of two different sizes: “large reference table” (i.e., 500,000 simulated datasets per scenario) and “small reference table” (i.e., 10,000 simulated datasets per scenario as for ABC-RF analyses). S.d. and CI stand for standard deviation and 95% confidence interval, respectively.

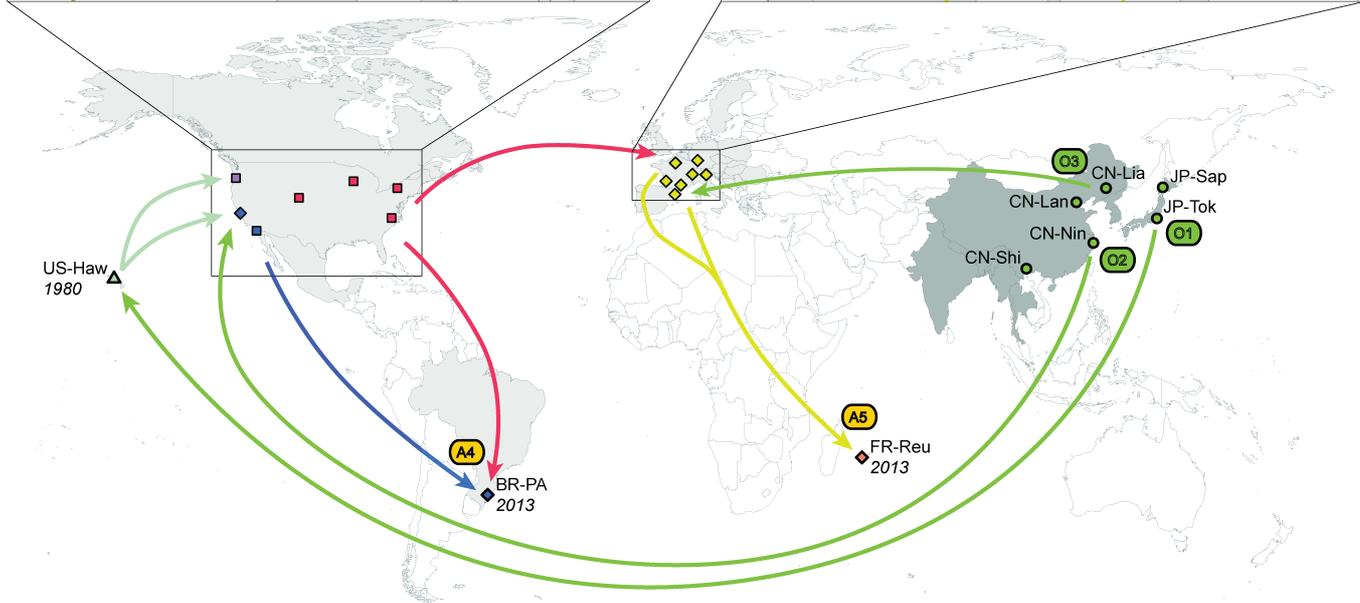
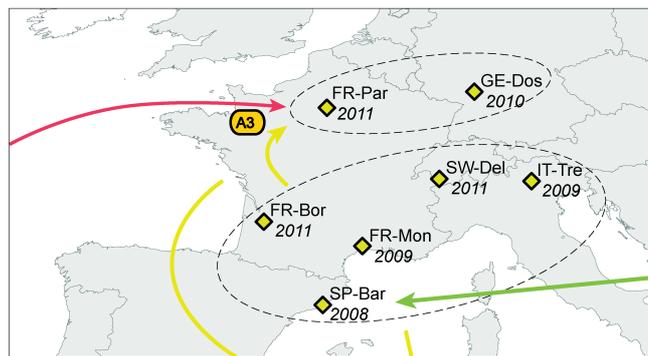
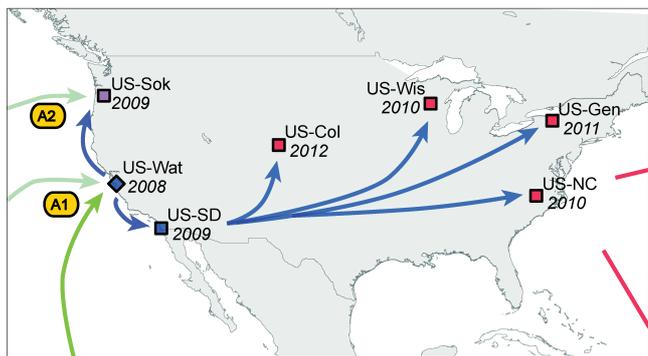


Fig. 1. Worldwide invasion scenario of *D. sukukii* inferred from microsatellite data and date of first observation.

Map and schematic showing sample sites and the invasion routes taken by *D. sukukii*, as reconstructed by ABC-RF (Pudlo *et al.* 2016) on a total of 685 individuals from 23 geographic locations genotyped at 25 microsatellite loci (see results and methods for details). The native range is in dark grey, and the invasive range is in light-gray (cf. delimitation from fig. 1 in Asplen *et al.* 2015). The year in which *D. sukukii* was first observed at each sample site is indicated in italics. The 23 geographical locations that were sampled are represented by circles (native range), and squares, diamonds and triangles (introduced range). Squares indicate populations that experienced weak bottlenecks (i.e. median value of bottleneck severity < 0.12, see main text and table 3), diamonds indicate moderate bottlenecks (0.12 < bottleneck severity < 0.22) and triangles indicate strong bottlenecks (i.e. bottleneck severity > 0.3). The colors of the symbol for the sample sites and arrows between them correspond to the different genetic groups obtained using the clustering method BAPS (Appendix S1). The arrows indicate the most likely invasion pathways. A1 to A5 indicate five separate admixture events between different sources. O1 to O3 indicate the most likely geographic locations of the origins within the native range for primary introductions. A1 = Hawaii + South-West China; A2 = Watsonville (western US) + Hawaii; A3 = southern Europe + eastern US; A4 = West US + East US; A5 = southern Europe + northern Europe. O1 = Japan; O2 = South-East China; O3 = North-East China.

We then focused our investigations on deciphering the potential relationships between the eastern US and European groups. Europe was never selected as a source for the eastern US (analysis 3a, $p = 0.744$), no matter which sets of sample sites were considered. By contrast, we found some evidence of asymmetrical gene flow from the eastern US into at least some European locations. When considering the northern European sample site from Germany (GE-Dos) as a target, we found that the best scenario included an admixed origin with Asia and the eastern US (table S4, analysis 3b, $p = 0.488$ and $p = 0.501$ for prior set 1 and 2, respectively). By contrast, the best scenario did not include such admixture event when considering the southern European sample site from Italy (IT-Tre) as a target (table 2, analysis 3b, $p = 0.510$). We therefore replicated analysis 3b on all possible combinations of European and eastern US sample sites to assess the robustness of this result and evaluate the possibility of a geographic pattern for admixture events. For each of the European sample sites, fig. 2 summarizes which replicate analyses selected a scenario including an admixture between eastern US and Asia or a scenario of a single introduction from Asia (and see fig. S3 for results using an alternative representative sample site for the native range). Results tended to indicate an uneven spatial distribution of the presence of genes of eastern US origin in Europe, following roughly a north-south gradient. More specifically eastern US alleles were present in northern Europe and absent, or at least not detected by our model choice method, in southern Europe. We then further tested whether the observed admixture in the north corresponded to the mixing of eastern US genes with southern European genes originating from the initial introduction event from Asia versus the mixing of eastern US genes with Asian genes introduced through a separate secondary introduction event from Asia in northern Europe (analysis 4). The best scenario was clearly that of an admixture between eastern US genes and southern European genes originating from the initial introduction event from Asia ($p = 0.998$; A3 event in fig. 1).

Finally, we found that the sampled population from Brazil originated from North America, with genetic admixture between *D. sukukii* individuals from the southwest and eastern regions of the US (analysis 5a; $p = 0.631$; A4 event in fig. 1). Regarding the most recent invasive population from La Réunion, we found that this island population originated from Europe, with admixture between individuals from northern and southern Europe (analysis 5b; $p = 0.500$; A5 event in fig. 1.).

Comparison of model choice analyses using ABC-RF versus ABC-LDA

For all eleven analyses, we found similar prior error rates for ABC-LDA using large size reference tables (i.e. 500,000 simulations per scenario) and ABC-RF (using only 10,000 simulations per scenario), with a trend towards slightly lower error rates for ABC-LDA (table 2, table S3). ABC-RF produced close or lower prior error rates than ABC-LDA when considering similarly small numbers



Best model = Admixture between Asia and eastern US



Best model = Single introduction from Asia

Source sample site of the eastern US group used for each replicated ABC-RF analyses

- 1: US-NC
- 2: US-Wis
- 3: US-Gen
- 4: US-Col

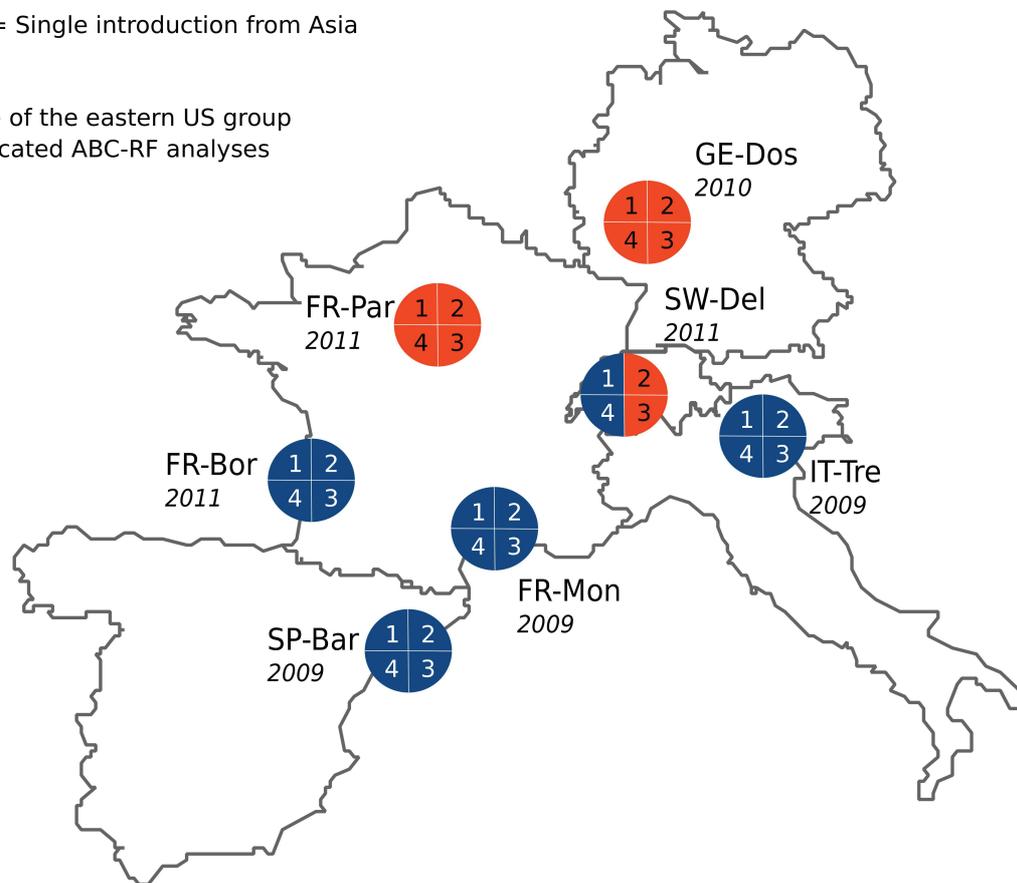


Fig. 2. Admixed or non-admixed origin of *D. sukikii* in Europe inferred for each sample sites.

Potential source populations include (among others) the Asian native range (represented by the Japanese sample site JP-Tok) and the invasive genetic group from eastern US represented by one of the four invasive sample sites collected in this area (fig. 1). Four replicate independent ABC-RF treatments corresponding to the analysis 3b (table 2) were hence carried out for each targeted European sample site using one of the four eastern US sample site. The treatments labeled 1, 2, 3 and 4 in the pies of the figure have been carried out with the sample sites US-NC, US-Wis, US-Gen and US-Col, respectively. A pie quarter in blue indicates that the best scenario corresponds to a single introduction event from Asia. A pie quarter in red indicates that the best scenario corresponds to an admixture event between Asia and eastern US. Dates in italic correspond to the dates of first record of the European sample sites.

of simulated datasets in the reference table for ABC-RF and ABC-LDA (i.e. 10,000 simulations per scenario for both types of analysis; table 2, table S3). This result was particularly marked for the complex analyses (i.e. those in which many introduction scenarios were compared and many summary statistics were computed). For instance, prior error rates were around 0.295 for ABC-RF versus 0.425 for ABC-LDA for the analyses 5a and 5b (21 compared scenarios and 424 summary statistics computed).

Regarding computational effort, we found that, for a given observed dataset, an ABC-RF treatment to select the best scenario and compute its posterior probability required a ca. 50 times shorter computational duration than when processing a ABC-LDA treatment with a traditional reference table of large size (i.e. 500,000 simulated datasets per scenario). Moreover, estimation of prior error rates with ABC-RF took only a few additional minutes of computation time for 10,000 pseudo-observed datasets whereas with ABC-LDA it lasted several hours to several days (depending on the analysis processed) for 500 pseudo-observed datasets.

The same invasion scenario had the highest probability for all eleven analyses carried out using ABC-RF and ABC-LDA with large number of simulated datasets (table 2; table S3). Posterior probabilities of the best scenario estimated using ABC-RF were not systematically higher (or lower) than those using ABC-LDA with large number of simulated datasets. In particular, there was hence no clear trend of overestimation of posterior probability when using ABC-LDA with large number of simulated datasets versus ABC-RF, a potential bias suggested by Pudlo *et al.* (2016). We found evidence of instability in the estimation of posterior probabilities and hence model selection when using ABC-LDA with a small numbers of simulated datasets in the reference table (i.e. 10,000 simulations per scenario as for ABC-RF). In this case, the scenario with the highest probability was different from that selected using either ABC-LDA with the large reference table or ABC-RF, in four and two of the eleven analyses when using the prior sets 1 and 2, respectively (results not shown).

Refining the origins of the primary introduction events from the native area

Additional ABC-RF analyses were run to determine which geographical area in the native range was the most likely origin of each of the three primary introductions (i.e. in Hawaii, western US, and Europe). The Hawaiian invasive population most likely originated from Japan, with a mean posterior probability of 0.959 (origin O1 in fig. 1). For Watsonville (western US, sample site US-Wat), the primary introduction from Asia most likely originated from southeast China with a probability of 0.860 (origin O2 in fig. 1). Finally, for Europe the primary introduction from Asia most likely originated from northeast China with a probability of 0.855 (origin O3 in fig. 1). Similar posterior probabilities were obtained using the prior set 1 (values presented above) and the prior set 2 (results not shown). For inferences regarding Europe, similar posterior probabilities were also obtained using various sets of sample sites from Europe and the eastern US (results not shown).

Estimation of parameter distributions for admixture rate and bottleneck severity

We used the most probable worldwide invasion scenario (fig. 1) to estimate the posterior distributions of parameters related to bottleneck severity associated to the foundation of new populations and genetic admixture rates between differentiated sources (i.e. A1 to A5 events in fig. 1).

The posterior distributions parameters relating to bottlenecks and admixture substantially differed from the priors, indicating that genetic data were informative for such parameters (table 3 and 4; and see tables S6-S7 for results using an alternative set of representative sample sites). Regarding bottleneck severity, we found that bottlenecks tended to be more severe for populations founded by individuals originating from a source located on a different continent (extra-continental origin) than for populations founded by individuals originating from a source located on the same continent (intra-continental origin; table 3 and fig. 1). The bottleneck severity for populations founded by individuals corresponding to a combination of the two types of sources (extra- and intra-continental origins) was either weak or moderate. The strongest bottleneck severity was found for the first invasive population in Hawaii, which showed the lowest level of genetic variation among all invasive populations. Bottleneck severity values were globally slightly stronger in European than in US populations, suggesting a smaller number of founding individuals and/or a slower demographic recovery in European populations.

Regarding admixture (table 4), we found that: (i) for the first recorded invasive population in western US (i.e. Watsonville; sample site US-Wat), the genetic contribution from China was substantially larger than that from Hawaii (median value of 0.759), (ii) for the western US population US-Sok, the secondary genetic contribution from Hawaii was small compared to that from Watsonville (median value of 0.237), and (iii) the sample sites from northern Europe (e.g. GE-Dos in Germany) contained a rather low proportion of genes originating from the eastern US (median value of 0.286). We found more balanced admixture rates for the populations from Brazil (BR-PA) and La Réunion (FR-Reu), but information on this parameter was less accurate in these cases as indicated by large 90% credibility intervals.

Model-posterior checking

When applying an ABC model-posterior checking analysis on the final invasion scenario detailed in fig. 1, we found that, when considering the prior set 1, only 54 of the 1141 summary statistics used as test quantities had low posterior predictive p-values (i.e. $0.002 < ppp\text{-values} < 5\%$). Moreover, the latter *ppp*-values values were not significant when applying the Benjamini and Hochberg 1995's method as an attempt to control the false discovery rate. Similar results were obtained

Introduction type	Sample site	Prior 1					Prior 2				
		mean	median	mode	q5%	q95%	mean	median	mode	q5%	q95%
Extra continental	US-Haw	0.517	0.500	0.495	0.326	0.755	0.490	0.484	0.477	0.382	0.615
	US-Wat	0.181	0.138	0.114	0.041	0.448	0.143	0.127	0.118	0.059	0.275
	IT-Tre	0.268	0.179	0.171	0.059	0.837	0.236	0.189	0.172	0.101	0.553
	BR-PA	0.158	0.126	0.104	0.020	0.407	0.208	0.193	0.172	0.067	0.399
	FR-Reu	0.259	0.215	0.186	0.051	0.626	0.245	0.214	0.190	0.069	0.534
Intra continental	US-SD	0.135	0.117	0.106	0.039	0.283	0.117	0.104	0.091	0.053	0.224
	US-NC	0.089	0.067	0.061	0.015	0.230	0.111	0.098	0.090	0.046	0.218
Extra + intra continental	US-Sok	0.116	0.099	0.088	0.027	0.250	0.119	0.106	0.099	0.052	0.229
	GE-Dos	0.189	0.177	0.155	0.061	0.356	0.205	0.189	0.171	0.088	0.372
Prior values		0.628	0.199	NA	0.021	1.904	0.517	0.500	NA	0.326	0.754

Table 3. Bottleneck severity in invasive populations of *D. sukukii*.

Note: Extra-continental introductions correspond to a long distance introduction from a source located apart from the continent of the focal population, intra-continental introduction corresponds to an introduction event from a source located on the same continent than the focal population, and Extra+Intra continental introduction corresponds to a combination of the two types of sources. Mean, median and mode estimates as well as bounds of 90% credibility intervals (q5% and q95%), are indicated for each bottleneck severity parameter. We roughly classified the estimated bottleneck severity values into three classes (represented here by the three shades of gray): weak (i.e. median value of bottleneck severity < 0.12, in light gray), moderate (0.12 < bottleneck severity < 0.22, in gray) and strong (i.e. bottleneck severity > 0.3 in dark gray). The set of sample sites used for the ABC estimations presented here include: (i) for the native area: Japan (JP-Tok+JP-Sap), South-East China (CN-Nin) and North-East China (CN-Lan+CN-Lia), and (ii) for the invaded range: US-Wat, US-Sok and US-SD for western US, US-NC for eastern US, IT-Tre for southern Europe, GE-Dos for northern Europe, BR-PA for South-America (Brazil) and FR-Reu for La Réunion island. Code names of the sample sites are the same as in fig. 1, and table S1 in which bottleneck severity classes are also given for each sample site. See table S6 for results on a different set of representative sample sites.

	Admixture event	Admixture rate (gene fraction from pop x)	mean	median	mode	q5%	q95%
Prior 1	A1 (US-Wat = US-Haw + CN-Nin)	r_{US-Wat} (China – CN-Nin)	0.759	0.761	0.774	0.659	0.854
	A2 (US-Sok = US-Haw + US-Wat)	r_{US-Sok} (USA – US-Wat)	0.237	0.230	0.227	0.092	0.400
	A3 (GE-Dos = IT-Tre + US-NC)	r_{GE-Dos} (USA – US-NC)	0.286	0.278	0.266	0.112	0.481
	A4 (BR-PA = US-NC + US-SD)	r_{BR-PA} (USA – US-SD)	0.467	0.461	0.442	0.187	0.754
	A5 (FR-Reu = IT-Tre + GE-Dos)	r_{FR-Reu} (Europe – GE-Dos)	0.388	0.380	0.426	0.132	0.670
Prior 2	A1 (US-Wat = US-Haw + CN-Nin)	r_{US-Wat} (China – CN-Nin)	0.761	0.762	0.766	0.684	0.833
	A2 (US-Sok = US-Haw + US-Wat)	r_{US-Sok} (USA – US-Wat)	0.224	0.219	0.199	0.114	0.350
	A3 (GE-Dos = IT-Tre + US-NC)	r_{GE-Dos} (USA – US-NC)	0.312	0.306	0.284	0.152	0.487
	A4 (BR-PA = US-NC + US-SD)	r_{BR-PA} (USA – US-SD)	0.422	0.418	0.429	0.247	0.613
	A5 (FR-Reu = IT-Tre + GE-Dos)	r_{FR-Reu} (Europe – GE-Dos)	0.370	0.368	0.356	0.178	0.569

Table 4. Posterior distributions of admixture rates for the five admixture events inferred for the final worldwide invasion scenario described in fig. 1.

Note: Admixture events are denoted as in figure 1. Each admixture rate parameter r points to the name of the admixed site and corresponds to the fraction of genes originating from the source site in parentheses ($1-r$ genes originate from the other source site). Mean, median and mode estimates as well as bounds of 90% credibility intervals (q5% and q95%) are indicated for each admixture parameter. Estimations assuming the prior set 1 and the prior set 2 (table S8) are provided. The set of representative sample sites used for the ABC estimations presented here is the same than in table 3. Code names of the population sites are the same as in fig. 1 and table S1. See table S7 for results on a different set of representative sample sites.

when assuming the prior set 2 and when considering different sets of representative sample sites in our analyses (results not shown). These results support that the final worldwide invasion scenario (with associated parameter posterior distributions) matches relatively well with the observed dataset. In agreement with this, the projections of the simulated datasets on the principal component axes from the final model-posterior combination were well grouped and centered on the target point corresponding to the observed dataset (fig. S4). Due to the modest size of the dataset (i.e. 25 microsatellite loci), however, only situations of major inadequacy of the model-posterior combination to the observed dataset are likely to be identified.

Discussion

In the present paper, we decipher the routes taken by *D. suzukii* in its invasion worldwide, evaluating evidence of bottlenecks and genetic admixture associated with different introductions, and we quantify the efficiency of ABC-RF relative to the more standard ABC-LDA method for choosing among introduction scenarios.

A complex worldwide invasion history

Prior to this study and that of Adrion *et al.* (2014), our understanding of the worldwide introduction pathways of *D. suzukii* was based on historical and observational data, which were incomplete and potentially misleading. Our results indicate three distinct introductions from the native range – to western Europe, the western side of North America, and to Hawaii, which accords well to findings of Adrion *et al.* (2014) from a different set of markers (X-linked sequence data). The almost simultaneous invasion of North America and Europe from Asia could be due to increased trade between these areas facilitating transport of this species between regions. Alternatively, or in addition, adaptation to human-altered habitats (in this case, changes in agricultural practices) within the native range of the species, could have promoted invasion to similarly altered habitats worldwide (i.e., anthropogenically induced adaptation to invade; Hufbauer *et al.* 2012). However, the two invaded continents were most likely colonized by flies from distinct Chinese geographic areas (fig. 1), requiring that adaptation to agriculture occurred concomitantly at several locations within the native range. This is possible, but not evolutionarily parsimonious, and further data would be required to test this explanation. Subsequent introductions from the three primary founder locations to other locations result in a complex history of invasion. Both primary and secondary introductions show signs of demographic bottlenecks, and several invasive populations have multiple sources and thus experienced genetic admixture between differentiated native or invasive populations.

The genetic relationships between worldwide *D. suzukii* populations define the genetic

variation they harbor, and may continue to shape further spread of alleles among populations. Gene flow among populations through continuous dispersal or more punctual admixture events permits the dissemination of allelic variants and new mutations, with dramatic consequences for evolutionary trajectories (Lenormand 2002). We found evidence for asymmetric genetic admixture from North American towards (northern) European populations. If this signal of admixture reflects recurrent and on-going dispersal events rather than a single past secondary introduction event, then variants that arise in North America would be more likely to spread to Europe than the reverse. Discriminating between ongoing and past dispersal events is tricky, however, especially for recent invasions (Benazzo *et al.* 2015). If the occurrence of recurrent gene flow is confirmed, then *D. suzukii* populations from Europe may not lag behind American ones in evolutionary terms. From an applied perspective, if resistance to control practices emerges in North America, special efforts to stop the importation of *D. suzukii* to Europe would help prevent the evolution of resistance in Europe, and thus reduce damage to crops (e.g. Caprio & Tabashnik 1992).

Advances in methods for discriminating among complex models

To our knowledge, the present study is the first to use the recently developed ABC-RF method (Pudlo *et al.* 2016) to test competing models characterized by different levels of complexity. Because ABC-RF reliably discriminates among models with reduced computational demand compared to standard ABC methods, it is particularly suitable for analyzing complex invasion scenarios involving multiple invasive populations and multiple potential sources, including replicating such analyses on different sample sets.

We found that, at least for some ABC-RF analyses, the posterior probability of the best scenario and the global statistical power to choose among alternative invasion scenarios (as measured by the prior error rate), were relatively low. For instance, in three of the eleven analyses (analyses 3b, 5a and 5b) the best scenario had relatively low probabilities (i.e., ranging from 0.500 to 0.630) and relatively high prior error rates (i.e., ranging from 0.300 to 0.400). Although replicate analyses carried out on different sample sets and prior sets pointed to the same best scenario for these three analyses, such low probability values and high prior error rates suggest a moderate level of confidence in the choice of model in these cases. Similarly, we found that the scenario choice leading to the conclusion of the presence versus absence of admixed genes from the eastern US in various European sample sites relied on probability values that were sometimes as low as 0.500. Moreover, in a minority of cases, the best scenario was different depending on the sample site considered as representative of the eastern US or the native area. The observed pattern of an uneven spatial distribution of the presence of genes of eastern US origin in Europe, following roughly a north to south gradient, should hence be taken cautiously. The analyses of additional European sample sites are needed to clarify this issue. The above inferential

uncertainties also most likely find their sources in the relatively small number of markers genotyped relative to the number, complexity, and similarity of some of the competing models. Such results indicate that there is room for improving the robustness of our inferences at least for a subset of our analyses.

From the perspective of methodological comparison, we carried out a subset of model choice analyses using both ABC-RF and the more standard ABC-LDA method (Estoup *et al.* 2012). We found that ABC-RF considerably out-performed ABC-LDA for selecting the “true” model when using similarly small reference tables for both type of analyses (e.g. 10,000 simulated datasets per scenario), especially when analyzing complex settings (e.g. when comparing more than ten scenarios using more than 100 summary statistics). The efficiency of ABC-LDA for choosing among introduction scenarios turned out to be better than predicted by Pudlo *et al.* (2016) when using large reference tables (e.g. 500,000 simulated datasets per scenario). ABC-RF is thus particularly appealing and competitive when the need for large reference tables renders more standard ABC methods computationally too costly. The much lower computation cost of ABC-RF was obvious. For a given observed dataset, selecting the best scenario and computing its posterior probability was ca. 50 times faster than when processing an ABC-LDA treatment with a traditional reference table of large size. Moreover, estimation of prior error rates with ABC-RF took only a few additional minutes of computation time whereas it lasted several hours to several days with ABC-LDA.

Using simulated datasets and toy examples for which the “true” posterior probabilities could be calculated, Pudlo *et al.* (2016) demonstrated that ABC-RF provided more reliable posterior probability of the best scenario than standard ABC methods when using the same (small) number of simulated datasets in the reference table (see fig. S3 in Pudlo *et al.* 2016). As true posterior probabilities are unknown and not computable for our real data and complex evolutionary models, we could not test this finding directly. We found, however, that the scenario with the highest probability was different in several analyses when using ABC-LDA with a small number of simulated datasets. This indicates substantial instability in the estimation of posterior probabilities when using ABC-LDA with small reference tables, especially when analyzing complex settings.

Bottlenecks and genetic admixture

We found strong evidence that bottlenecks and admixture both occurred during the course of the invasion of *D. sukukii*. A first sign of bottlenecks is a reduction in neutral genetic diversity. We found that all invasive *D. sukukii* populations were characterized by significantly lower genetic variation than native ones, with a loss of 6.4% (US-Col) to 23.3% (US-Haw) of heterozygosity and of 27.3% (US-Wat) to 54.2% (US-Haw) of allelic diversity, respectively (fig. S5). Adrion *et al.* (2014) find a similar significant loss of diversity in their dataset focused on gene sequences. Dlugosch & Parker

(2008) and Uller & Leimu (2011) reviewed studies of neutral genetic diversity in a large number of species of animals, plants, and fungi and compared nuclear molecular diversity within introduced and source populations. Overall they found that a loss of variation was the most frequent feature in invasive populations. However, reductions in genetic variation were on average modest (e.g., average loss of 18.7% and 15.5% of heterozygosity and allelic diversity, respectively; Dlugosch & Parker 2008). The reductions observed in our study in invasive *D. sukikii* populations were thus in the same range for heterozygosity but higher for allelic diversity (cf. average loss of 11.1% and 34.5% of heterozygosity and allelic diversity, respectively). A larger loss of allelic diversity than heterozygosity after a bottleneck is nevertheless expected by theory (Nei 1975).

The type of introduction pathway explains, at least partly, the severity of the reduction in genetic diversity. Bottlenecks tended to be less severe for populations founded by individuals from the same continent than for populations founded by individuals originating from a different continent. A simple explanation for such pattern is that, due to greater proximity, populations originating from the same continent are likely to be initially founded by a larger numbers of individuals and also to experience recurrent gene flow. Populations founded by individuals from both the same and different continents experienced weak or moderate reductions in diversity, suggesting that multiple introduction pathways restore diversity.

In agreement with previous studies on invasions on similarly large geographical scales (e.g. Lombaert *et al.* 2014 and reference therein), we found that admixture events were frequent in the worldwide invasion history of *D. sukikii*. We identified at least five admixture events for which sources were native and/or invasive populations (i.e. A1 to A5 events in fig. 1). The genetic contributions of the sources were unbalanced, except for the populations from Brazil and La Réunion; but genetic information regarding admixture was considerably less accurate in the latter cases, which involved weakly differentiated source populations.

We have provided here the information necessary to evaluate whether bottlenecks and admixture have influenced outcomes in this invasion. Quantitative genetics studies in the lab focusing on the key source, bottlenecked, and admixed populations could now examine the fitness and adaptive potential of those groups (e.g. Lavergne & Molofsky 2007; Facon *et al.* 2008; Turgeon *et al.* 2011).

Conclusions and perspectives

Understanding invasion pathways provides key information for understanding the role of evolutionary processes in biological invasions. For example, only does our research reveal sources of invasive populations for further relevant comparisons using quantitative genetics studies in the lab. Additionally, we found that the invasion of Europe by *D. sukikii* was distinct from that of North America, with limited and asymmetrical gene flow between these two main invaded areas. This

situation provides the opportunity to evaluate evolutionary trajectories in replicate into two separate temperate climate regions, similarly to research by Gilchrist *et al.* (2001, 2004) on the pace of clinal evolution in the invasive fruit fly *Drosophila subobscura*. There are also distinct invasion pathways for areas with warmer climates (i.e. Hawaii, La Réunion and Brazil) that will make interesting comparisons.

Retracing invasion routes and making inferences about demographic processes is only possible if there is adequate polymorphism within populations and significant genetic differentiation among them. As is evident here, microsatellite data can provide enough variation to reveal important pathways. We concur with Adrion *et al.* (2014), however, that genome-wide data (i.e. next generation sequencing approaches) will be tremendously powerful in further discriminating among complex invasion scenarios (see also Cristescu 2015, Estoup *et al.* 2016). Because of the reduced computational resources demanded by ABC-RF, this method will be particularly useful for analysis of massive single nucleotide polymorphism datasets (Pudlo *et al.* 2016). In addition to the selectively neutral demographic inference leading to the reconstruction of routes of invasion, population (and quantitative) next generation sequencing approaches are quite promising for studying the evolution of phenotypic traits in natural populations (Wray 2013; Gautier 2015). Specifically, they can be used to better understand the genetic architecture of traits underlying invasion success (Bock *et al.* 2015). *Drosophila suzukii* is a good species for such research given (i) its short generation time and viability in laboratory conditions, which facilitate experimental approaches to study quantitative traits of interest (Asplen *et al.* 2015) and (ii) the availability of annotated genome assemblies for this species (Chiu *et al.* 2013; Ometto *et al.* 2013), along with the huge amount of genomic resources available in its close relative species *D. melanogaster* (Groen & Whiteman 2016).

Materials and methods

Sampling and genotyping

Adult *D. suzukii* were sampled from the field at a total of 23 localities (hereafter termed sample sites) distributed throughout most of the native and invasive range of the species (table S1 and fig. 1). Samples were collected between 2013 and 2015 using baited traps and sweep nets, and stored in ethanol. Native Asian samples consisted of a total of six sample sites including four Chinese and two Japanese localities. Samples from the invasive range were collected in Hawaii (1 sample site), Continental US (7 sites), Europe (7 sites), Brazil (1 site) and in the French island of La Réunion (1 site). For each sample site, 15 to 44 adult flies were genotyped at 25 of the 28 microsatellite loci described in Fraimout *et al.* (2015), resulting in a total of 685 genotyped individuals. Three microsatellite loci from Fraimout *et al.* (2015) (i.e. loci DS31, DS42 and DS45) were not included

due to the presence of null alleles at frequencies > 10% in some sample sites (results not shown). DNA extraction and PCR amplification as well as allele scoring were performed following the methodology detailed in Fraimout *et al.* (2015).

General framework for defining focal population groups and sets of scenarios to be compared using ABC

One of the many challenges of ABC analyses is to optimally define sets of sample sites to be processed chronologically as potential sources and targets in a way that do not hamper the computational effort. As stated in Estoup & Guillemaud (2010) and Lombaert *et al.* (2014) the number and complexity of competing scenarios is proportional to the number of genetic groups to be accounted in ABC analyses. Following Lombaert *et al.* (2014), we used the results obtained from three genetic clustering methods along with historical and geographical information to define genetic groups that could include several sample sites (Appendix S1). We defined seven main genetic groups from our set of 23 sample sites: Asia, Hawaii, western US, eastern US, Europe, Brazil and La Réunion. To make the computations less intensive with respect to computer resources, ABC treatments were carried out using one representative sample site for each genetic group per replicate analysis. In other words, when a genetic group included more than two sample sites, we considered the two most differentiated sample sites (i.e. with the highest F_{ST} values in table S3) as representative of the group: CN-Shi and JP-Tok for the group Asia, all three sample sites for the genetically heterogeneous Western US group, US-NC and US-Wis for the group eastern US, GE-Dos and IT-Tre for the group Europe. Each of the sample sites retained (i.e. the representative sample sites) was included alternatively in the sample set considered for ABC treatments. This makes it possible to replicate analyses of scenario choice on different sample sets representing most of the genetic variation within and between genetic groups. All ABC analyses described in table 1 were thus replicated with all possible combinations of representative sample sites available for each genetic group. This allowed us to test the robustness of scenario choices with respect to the sample sets considered as representative of genetic groups.

Nested ABC analyses processed sequentially

Following Lombaert *et al.* (2014), we used historical information (i.e., dates of first observation of each invasive populations; table S1) to define eleven sets of competing introduction scenarios that were analyzed sequentially. The first set of compared scenarios considers the second oldest invasive population as target (the oldest one necessarily originating from the native range) and determines its introduction history. Step by step, subsequent analyses use the results obtained from the previous analyses, until the most recent invasive populations are considered. Each of the

eleven analyses included a set of compared introduction scenarios where the target invasive population (i.e. the focal population) can either directly derive by a single split (i.e. introduction event) from one of the historically compatible source populations, or be the result of an admixture event between one of all possible combinations of pairwise sources. For instance, in a model with one target and two possible sources, three scenarios could be formalized with the target population being derived from a single source population, from the other possible source, or from an admixture between the two sources (see fig. S6 for an illustration and table S2 for details).

The first set of scenario choice analyses 1a to 1d aimed at making inferences about the introduction pathways for the three sample sites in western US (table 1). We first evaluated the Asian or Hawaiian or admixed Asian+Hawaiian origin of each three western US sample sites. We then refined our modeling by formalizing seven competing scenarios describing the possible relationships among the three western US sample sites and their extra-continental sources (i.e., Asia and Hawaii; analysis 1d). Once the invasion history was resolved for the western US group, we sought to identify the population source(s) of the eastern US and European genetic groups, independently. Dates of first records in these two groups widely overlap and the possibility that the western US was a common source for both areas could not be ruled out. We therefore compared in analysis 2a a set of six scenarios to assess whether the eastern US group directly derived from Asia, Hawaii or western US or if it resulted from an admixture event between one pair of such sources. A similar set of six scenarios was used to assess the most likely origin of European populations, independently from the eastern US group (analysis 2b).

Owing to the results of analyses 2a and 2b and due to the overlap in first record dates for eastern US and Europe, we then performed two analyses to investigate the possibility of gene flows (i.e. admixture) between the eastern US and European groups, considering eastern US as target population with Europe as potential source and *vice versa* (analyses 3a and 3b, reciprocally). Analysis 4 aimed at clarifying the admixture in northern Europe sample sites identified in analysis 3b. We tested whether this admixture occurred between eastern US and a secondary introduction from Asia, or between individuals from eastern US and southern Europe. Finally, we investigated the origins of the most recent invasive populations in Brazil and La Réunion, separately, with all older groups and pairwise admixtures as potential sources (analyses 5a and 5b). Note that in agreement with the genetic clustering results (Appendix S1; fig. 1, S1 and S2) we did not test for the possibility of Brazil being the source of La Réunion, and reciprocally.

Historical, demographic and mutational parameters

For all scenarios formalized for ABC analyses, we modeled an introduction event as a divergence without subsequent gene-flow from the source population (or two source populations in case of admixture) at a time corresponding to the date of first record in the invaded area translated into a

number of generations (assuming 12 generation per year; Lin *et al.* 2014). The divergence event was immediately followed by a bottleneck period characterized by a lower effective population size, and a straight return to a stable (large) effective population size. We modeled our uncertainty regarding the identity of the source populations in the slightly structured native Asian area by including native "ghost populations" (i.e. unsampled populations) in our scenarios. Such unsampled native populations were modeled as a single split from an ancestral Asian population (without any change in effective population size) at a time defined by a loose flat prior, which includes zero at the lower bound (table S8). See Lombaert *et al.* (2014) and references therein for justifications of using unsampled populations when modeling invasion scenarios. A detailed illustration of the above modeling design is provided in fig. S6.

Prior distributions for historical, demographic and mutational parameters were defined taking into account the historical and demographic parameter values available from empirical studies on *D. suzukii* (Lin *et al.* 2014, Hauser 2011, Asplen *et al.* 2015) and microsatellite mutational data on *D. melanogaster* (Shug *et al.* 1998). We considered a first set of prior distributions (hereafter termed prior set 1), which was composed of rectangular (i.e. bounded uniforms) distributions (see table S8). To evaluate the robustness of our ABC inferences to prior choice, we also considered a second set of more peaked prior distributions (hereafter termed prior set 2; table S8).

Model choice using ABC-RF and ABC-LDA

For all ABC-RF and ABC-LDA analyses, we used the software DIYABC v.2.1.0 (Cornuet *et al.* 2014) to simulate datasets constituting the reference tables. A reference table includes a given number of datasets that have been simulated for different scenarios using parameter values drawn from prior distributions, each dataset being summarized with a pool of statistics.

Following Pudlo *et al.* (2016), ABC-RF treatments were processed on reference tables including 10,000 simulated datasets per scenario. Datasets were summarized using the whole set of summary statistics proposed by DIYABC (Cornuet *et al.* 2014) describing genetic variation per population (e.g. number of alleles), per pair (e.g. genetic distance) or per triplet (e.g. admixture rate) of populations, averaged over the 25 loci (see the DIYABC 2.1.0 user-manual p.16 for details about such statistics), plus the linear discriminant analysis (LDA) axes as additional summary statistics. The total number of summary statistics ranged from 39 to 424 depending on the analysis (table 2). When the number of scenarios in an analysis exceeded ten, we processed ABC-RF treatments on reference tables including 100,000 simulated datasets to avoid computer memory issues associated to the sub-bootstrapping procedure processed for reference tables with more than 100,000 simulated datasets; see the section *Practical recommendations regarding the implementation of the algorithms* in Pudlo *et al.* (2016). We checked that such number was

sufficient by evaluating the stability of prior error rates (i.e. the probability to choose a wrong model when drawing model index and parameter values into priors) and posterior probabilities estimations on 80,000, 90,000 and 100,000 simulated datasets. The number of trees in the constructed random forests was fixed to $n = 500$; see Pudlo *et al.* 2016, Breiman 2001 for justifications of considering a forest of 500 trees and fig. S7 for an illustration. For each ABC-RF analysis, we predicted the best scenario, estimated its posterior probabilities and prior error rates over 10 replicate runs of the same reference table. We used the `abcrf` R package (v1.1.0; Pudlo *et al.* 2016) to perform all ABC-RF analyses. In Appendix S2, we provide several scripts written in R programming language (R Development Core Team 2008) for computing random forest analyses in R with the `abcrf` package, when starting from simulated datasets generated with DIYABC v.2.1.0.

For sake of methodological comparisons, we also duplicated a subset of analyses using a more standard ABC method (Beaumont *et al.* 2002; Cornuet *et al.* 2008) improved by using the regression algorithms from Estoup *et al.* (2012) based on linear discriminant analysis (LDA) to avoid the curse of dimensionality and correlation among explanatory variables (i.e. multi-collinearity) during the regression step. This method is called ABC-LDA. Following previous analyses of this type (e.g. Lombaert *et al.* 2014), ABC-LDA treatments were processed on reference tables including 500,000 simulated datasets per scenario. As for ABC-RF, datasets were summarized using the whole set of summary statistics proposed by DIYABC (Cornuet *et al.* 2014). Posterior probabilities associated to each scenario were estimated by polychotomous logistic regression (Cornuet *et al.* 2008), modified following Estoup *et al.* (2012), on the 1% of the simulated datasets closest to the observed dataset. We also estimated for each analysis a prior error rate over 500 simulated datasets drawn into priors using reference tables including 500,000 simulated datasets per scenario. To provide a fair comparison of classification error with respect to the computational effort, prior error rates were also estimated using reference tables including the same (and hence lower) number of simulated datasets per scenario as for ABC-RF treatments.

Refining the origins of the primary introduction events from the native area

We refined our worldwide invasion scenario by assessing the most likely origins of the primary introductions from the Asian native range. To this aim, we carried out additional ABC-RF treatments to choose which geographical area in the native range was the most likely origin of each of the three primary introductions from Asia identified in the final invasion scenario, i.e. in Hawaii, western US (sample site US-Wat), and Europe (fig. 1). Capitalizing on the observed pattern of genetic differentiation among all Asian sample sites, we considered four possible geographical origins: O1 = Japan (represented by the sample sites JP-Sap+Jp-Tok that were pooled as they did not show any significant differentiation), O2 = South-East China (represented by the sample site CN-Nin), O3 = North-West China (represented by the sample sites CN-Lan+CN-Lia that were pooled as they

did not show any significant differentiation), and O4 = South-West China (represented by the sample site CN-Shi). In each of these three model choice analyses (cf. one analysis per primary introduction), we considered the previously inferred introduction histories and compared four scenarios corresponding to variation of those histories that only differed by the geographical origin of the primary Asian introduction (models O1, O2, O3 and O4). We used the ABC-RF method with 10,000 datasets per model simulated using the prior set 1 or 2, and with three replications of the RF process. For inferences regarding Europe, we carried out ABC-RF treatments on various sets of sample sites from Europe and eastern US to further evaluate the robustness of our results.

Parameter estimation

For ABC parameter estimation, we considered a set of 12 key sample sites representative of the inferred final invasion history (see fig. 1). This set of sample sites included the three native origins of primary introductions from the native range (O1, O2 and O3), the Hawaiian sample (US-Haw), the three western US samples (US-Wat, US-Sok and US-SD), the eastern US sample US-NC, two European samples (IT-Tre and GE-Dos), the Brazilian sample (BR-PA) and the sample from La Réunion (FR-Reu). To evaluate the robustness of our parameter estimations with respect to sample sites we carried out a second analysis replacing the sample sites US-NC, IT-Tre and GE-Dos by US-Wis, SP-Bar and FR-Par, respectively. We also replicated parameter estimations using the prior set 1 and the prior set 2.

ABC-RF was developed to tackle the inferential issue of model choice. So far no RF-based solution exists to estimate parameter distributions under a given model (Pudlo *et al.* 2016). Standard ABC algorithms for parameter estimation may suffer from the curse of dimensionality and correlation among explanatory variables (i.e. multi-co-linearity) during the regression step, and hence yield poor results when the number of statistics is much too large (reviewed in Blum *et al.* 2013). To avoid such potential problems as well as for the sake of simplicity and computation efficiency (i.e. statistical techniques for choosing summary statistics have not been implemented in DIYABC; Blum *et al.* 2013), we used a standard ABC methodological framework (Beaumont *et al.* 2002; Cornuet *et al.* 2008) applied to a subset of “expert-chosen” statistics proposed in DIYABC (e.g. Lombaert *et al.* 2014; Brouat *et al.* 2014) to estimate posterior parameter distributions under the final invasion scenario. We hence considered 95 summary statistics in total: the mean number of alleles and the mean genetic diversity (per locus and population sample), all pairwise F_{ST} 's, and some crude estimates of admixture rates based on the five population triplets for which admixture events have been inferred (i.e. A1 to A5 events described in fig. 1 corresponding to five AML statistics; Choisy *et al.* 2004, see the DIYABC 2.1.0 user-manual p16 for details about such statistics). Using parameter values drawn from the prior sets 1 or 2, we produced a reference table containing ten million simulated datasets. Following Beaumont *et al.* (2002), we then used a local

linear regression to estimate the parameter posterior distributions. We took the 10,000 (1 %) simulated datasets closest to our observed dataset for the regression, after applying a logit transformation to parameter values. We found that considering instead the 1 % closest simulated datasets and/or a larger sets of summary statistics had only weak effects on the estimated parameter distributions, at least for the subset of parameters we were interested in (i.e., admixture rates and bottleneck severity) (results not shown).

We focused our investigations on the parameters related to two types of demographic events that are considered to be potentially important drivers of invasion success: severity of the bottleneck associated with the foundation of new populations and genetic admixture (e.g. Estoup *et al.* 2016). Bottleneck severity was defined as the ratio between the duration of the bottleneck and the number of effective individuals during this period (e.g. Pascual *et al.* 2007). Following the final invasion scenario detailed in fig. 1, we defined three categories of introduction situations: bottleneck severity for populations founded only by individuals originating from a different continent (e.g. sample sites US-Haw, US-Wat, IT-Tre, BR-PA and FR-Reu), bottleneck severity for populations founded only by individuals originating from the same continent (e.g. sample sites US-SD and US-NC) and bottleneck severity for populations founded by a combination of the two types of sources (e.g. sample sites US-Sok and GE-Dos). Genetic admixture is evident when a population is inferred to have more than one distinct source (A1 to A5 events in fig. 1).

Model-posterior checking

We used the ABC model-posterior checking method implemented in DIYABC (Cornuet *et al.* 2010) to evaluate how well the final worldwide invasion scenario with associated parameter posterior distributions matches with the observed dataset. This method was largely inspired by statistical frameworks detailed in Gelman *et al.* (2003) and Cook *et al.* (2006). The principle is as follows: if a scenario-posterior combination fits the observed data correctly, then data simulated under this scenario with parameters drawn from associated posterior distributions should be close to the observed data. The lack of fit of the model to the data can be measured by determining the frequency at which test quantities measured on the observed dataset are extreme with respect to the distributions of the same test quantities computed from the simulated datasets (i.e. the posterior predictive distributions). In practice, the test quantities are chosen among the large set of ABC summary statistics proposed in the program DIYABC. For each test quantities (q), a lack of fit of the observed data with respect to the posterior predictive distribution can be measured by the cumulative distribution function values defined as $\text{Prob}(q_{\text{simulated}} < q_{\text{observed}})$. Tail-area probability can be easily computed for each test quantity as $\text{Prob}(q_{\text{simulated}} < q_{\text{observed}})$ and $1.0 - \text{Prob}(q_{\text{simulated}} < q_{\text{observed}})$ for $\text{Prob}(q_{\text{simulated}} < q_{\text{observed}}) \leq 0.5$ and > 0.5 , respectively. Such tail-area probabilities, also named posterior predictive p-values (i.e. *ppp*-values; Meng 1994), represent the probabilities that the

replicated data (simulated ABC summary statistics) could be more extreme than the observed data (observed ABC summary statistics). In practice, *ppp*-values can be interpreted as a guideline for tracking model-posterior misfit: a few *ppp*-values around 0.05 are not necessarily a big deal whereas the presence of *ppp*-values around 0.001 is pretty suspect. Hence, too many observed summary statistics falling in the tails of distributions, especially if some of those *ppp*-values are pretty low, cast serious doubts on the adequacy of the model-posterior combination to the observed dataset. Finally, because *ppp*-values are computed for a number of (often non-independent) test statistics, a method such as that of Benjamini and Hochberg (1995) can be used to control the false discovery rate (e.g. Cornuet *et al.* 2010).

We carried out the ABC model-posterior checking analysis on our observed microsatellite dataset as follows. From 10^7 datasets simulated under the final invasion scenario detailed in fig. 1 and using the same subset of 95 statistics previously retained for parameter estimation, we obtained a posterior sample of 10^4 values from the posterior distributions of parameters, as described in the previous section *Parameter estimation*. We then simulated 10^4 datasets with parameter values drawn with replacement from this posterior sample. As underlined in many text books in statistics (e.g. Gelman *et al.* 2003), it is advised against performing model checking using information that have already been used for training (i.e. model fitting; see also Cornuet *et al.* 2010 for illustrations on simulated datasets). Optimally, model-posterior checking should be based on test quantities that do not correspond to the summary statistics which have been used for obtaining the parameters posterior distributions. This is naturally possible with DIYABC as the package propose a large choice of summary statistics. Our set of test statistics therefore included the 1141 remaining single, pairwise and three-population summary statistics available in DIYABC that were not used for previous ABC parameter estimation.

Acknowledgements

AF was supported by state funding by the Agence Nationale de la Recherche, through the LabEx ANR-10-LABX-0003-BCDiv, of the program "Investissements d'avenir" (ANR-11-IDEX-0004-02). AE and RH were supported by the Languedoc Roussillon region (France) through the European Union program FEFER FSE IEJ 2014-2020 (project CEPADROL), the INRA scientific department SPE (AAP-SPE 2016), and the French Agropolis Fondation (Labex Agro-Montpellier) through the AAP "International Mobility" (CfP 2015-02). GL acknowledges support from Cornell University's New York State Agricultural Experiment Station federal formula funds, project # 2012-13-195. XC, MK, IM and JZ were funded by the European Union's 7th Framework Programme under grant agreement No. 613678 (DROPSA). MD acknowledges support from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). MP was supported by project CTM2013-48163. The

authors acknowledge Julio Pedraza and the UMS 2700 OMSI for computational support. Molecular data used in this work were produced through the genotyping and sequencing facilities of ISEM (Institut des Sciences de l'Evolution-Montpellier) and Labex CEMEB (Centre Méditerranéen Environnement Biodiversité). The authors thank Masanori Toda for his help in *D. suzukii* sampling.

References

- Adrion JR, Kousathanas A, Pascual M, Burrack HJ, Haddad NM, Bergland AO, Machado H, Sackton TB, Schlenke TA, Watada M, Wegmann D, Singh ND. 2014. *Drosophila suzukii*: the genetic footprint of a recent, worldwide invasion. *Mol Biol Evol.* 31: 3148–3163.
- Alcala N, Jensen J, Telenti A, Vuilleumier S. 2015. The genomic signature of population reconnection following isolation: from theory to HIV. *G3.* 6:107-120.
- Asplen MK, Anfora G, Biondi A, Choi DS, Chu D, Daane KM, Gibert P, Gutierrez AP, Hoelmer KA, Hutchison WD, Isaacs R, Jiang ZL, Kárpáti Z, Kimura MT, Pascual M, Philips CR, Plantamp C, Ponti L, Véték G, Vogt H, Walton VM, Yu Y, Zappalà L, Desneux N. 2015. Invasion biology of spotted wing *Drosophila* (*Drosophila suzukii*): a global perspective and future priorities. *J Pest Sci.* 88: 469–494.
- Atallah J, Teixeira L, Salazar R, Zaragoza G, Kopp A. 2014. The making of a pest: the evolution of a fruit-penetrating ovipositor in *Drosophila suzukii* and related species. *Proc R Soc B.* 281: 20132840–20132840.
- Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics.* 162: 2025–2035.
- Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annu. Rev Ecol Evol Syst.* 41: 379-406.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B.* 57: 289-300.
- 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: 2609-2625.
- Blum M, Nunes M, Prangle D, Sisson S. 2013. A comparative review of dimension reduction methods in approximate Bayesian computation. *Stat Sci.* 28:189-208.
- Bock DG, Caseys C, Cousens RD, Hahn MA, Heredia SM, Hübner S, Turner KG, Whitney KD, Rieseberg LH. 2015. What we still don't know about invasion genetics. *Mol Ecol.* 24: 2277-2297.
- Boissin E, Hurley B, Wingfield MJ, Vasaitis R, Stenlid J, Davis C, de Groot P, Ahumada R, Carnegie A, Goldarazena A, Klasmer P, Wermelinger B, Slippers B. 2012. Retracing the routes of introduction of invasive species: the case of the *Sirex noctilio* woodwasp. *Mol Ecol.* 21: 5728–5744.

- Breiman L. 2001. Random forests. *Machine Learning* 45: 5 – 32.
- Benazzo A, Ghirotto S, Torres Vilaca ST, Hoban S. 2015. Using ABC and microsatellite data to detect multiple introductions of invasive species from a single source. *Heredity* 115: 262–272.
- Brouat C, Tollenaere C, Estoup A, Loiseau A, Sommer S, Soanandrasana R, Rahalison L, Rajerison M, Piry S, Goodman SM, Duplantier JM. 2014. Invasion genetics of a human commensal rodent: the black rat *Rattus rattus* in Madagascar. *Mol Ecol.* 23(16): 4153-4167.
- Caprio M, Tabashnik BE. 1992. Gene flow accelerates local adaptation among finite populations: simulating the evolution of insecticide resistance. *J Econ Entomol.*85: 611–620.
- Calabria G, Máca J, Bächli G, Serra L, Pascual M. 2012. First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe. *J Appl Entomol.* 136 : 139–147.
- Chiu JC, Jiang XT, Zhao L, Hamm CA, Cridland JM, Saelao P, Hamby KA, Lee EK, Kwok RS, Zhang G, Zalom FG. 2013. Genome of *Drosophila suzukii*, the spotted wing drosophila. *G3.* 3: 2257 – 2271.
- Choisy M, Franck P, Cornuet JM. 2004. Estimating admixture proportions with microsatellites: comparison of methods based on simulated data. *Mol Ecol.* 13: 955 – 968.
- Cook S, Gelman A, Rubin DB. 2006. Validation of software for Bayesian models using posterior quantiles. *J Comput Graph Stat.* 15: 675–692.
- Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, Guillemaud T, Estoup A. 2008. Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics.* 24: 2713–2719.
- Cornuet JM, Ravigné V, Estoup A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *Bmc Bioinformatics.* 11: 401.
- Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A. 2014. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics.* 30, 1187–1189.
- Cristescu ME. 2015. Genetic reconstructions of invasion history. *Mol Ecol.* 24: 2212 – 2225.
- Deprá M, Poppe JL, Schmitz HJ, De Toni DC, Valente VLS. 2014. The first records of the invasive pest *Drosophila suzukii* in the South American continent. *J Pest Sci.* 87: 379–383.
- Dlugosch KM, Parker IM. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol.* 17:431–49.
- Dlugosch KM, Anderson SR, Braasch J, Alice Cang F, Gillette H. 2015. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Mol Ecol.* 24:2095—111.
- Edmonds CA, Lillie AS, Cavalli-Sforza LL. 2004. Mutations arising in the wave front of an expanding population. *Proc Natl Acad Sci USA.* 101, 975–979.

- Ellstrand NC, Schierenbeck KA. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc Nat Acad Sci USA*. 97: 7043-7050.
- Estoup A, Guillemaud T. 2010. Reconstructing routes of invasion using genetic data: why, how and so what?. *Mol Ecol*. 19, 4113–4130.
- Estoup A, Lombaert E, Marin JM, Guillemaud T, Pudlo P, Robert CP, Cornuet JM. 2012. Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Mol Ecol Res*.12(5) : 846-855.
- Estoup A, Ravigné V, Hufbauer RA, Vitalis R, Gautier M, Facon B. 2016. Is there a genetic paradox of biological invasion? *Annu Rev Ecol Evol Syst*. 47: 51-72.
- Facon B, Genton B, Shykoff J, Jarne P, Estoup A, David P. 2006. A general eco-evolutionary framework for understanding bioinvasions. *Trends Ecol Evol*. 21:130-35.
- Facon B, Pointier JP, Jarne P, Sarda V, David P. 2008. High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Curr Biol*. 18:363 – 67.
- Fraimout A, Loiseau A, Price DK, Xuereb A, Martin JF, Vitalis R, Fellous S, Debat V, Estoup A. 2015. New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest. *Eur J Entomol*. 112(4): 855.
- Gautier M. 2015. Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*. 201:1555-79.
- Gelman A, Carlin JB, Stern HS, Rubin DB. 2003. *Bayesian Data Analysis*. London: Chapman & Hall/CRC.
- Gilchrist GW, Huey RB, Serra L. 2001. Rapid evolution of wing size clines in *Drosophila subobscura*. *Genetica* 112, 273–286.
- Gilchrist GW, Huey RB, Balanyà J, Pascual M, Serra L, Noor M. 2004. A time series of evolution in action: a latitudinal cline in wing size in South American *Drosophila subobscura*. *Evolution* 58, 768–780.
- Groen SC, Whiteman NK. 2016. Using *Drosophila* to study the evolution of herbivory and diet specialization. *Curr Opin Insect Sci*. 14, 66-72.
- Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM, Estoup A. 2010. Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity* 104: 88–99.
- Hauser M. 2011. A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Manag Sci*. 67: 1352–1357.
- Hufbauer RA, Facon B, Ravigné V, Turgeon J, Foucaud J, Lee CE, Rey O, Estoup A. 2012. Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evol. Appl*. 5: 89-101.

- Kaneshiro KY. 1983. *Drosophila* (Sophophora) *suzukii* (Matsumura). *Proc Hawaiian Entomol Soc.* 24:179.
- Karsten M, Jansen van Vuuren B, Addison P, Terblanche JS. 2015. Deconstructing intercontinental invasion pathway hypotheses of the Mediterranean fruit fly (*Ceratitidis capitata*) using a Bayesian inference approach: are port interceptions and quarantine protocols successfully preventing new invasions? *Divers. Distrib.* 21: 813–825.
- Keller SR, Taylor DR. 2010. Genomic admixture increases fitness during a biological invasion: Admixture increases fitness during invasion. *J Evol Biol.* 23: 1720–1731.
- Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB. 2004. Genetic variation increases during biological invasion by a Cuban lizard. *Nature.* 431:177–81.
- Lavergne S, Molofsky J. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc Natl Acad Sci USA.* 104: 3883–3888.
- Lawson Handley L, Estoup A, Evans D, Thomas C, Lombaert E, Facon B, Aebi A, Roy H. 2011. Ecological genetics of invasive alien species. *BioControl.* 56:409-428.
- Lee C. 2002. Evolutionary genetics of invasive species. *Trends Ecol Evol.* 17:386-391.
- Lee JC, Bruck DJ, Curry H, Edwards D, Haviland DR, Van Steenwyk RA, Yorgey BM. 2011. The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest Manag Sci.* 67: 1358–1367.
- Lenormand T. 2002. Gene flow and the limits of natural selection. *Trends Ecol Evol* 17:183–189.
- Lin QC, Zhai YF, Zhang AS, Men XY, Zhang XY, Zalom FG, Zhou CG, Yu Y. 2014. Comparative Developmental Times and Laboratory Life tables for *Drosophila suzukii* and *Drosophila melanogaster* (Diptera: Drosophilidae). *Fla Entomol.* 97 : 1434–1442.
- Lombaert E, Guillemaud T, Lundgren J, Koch R, Facon B, Grez A, Loomans A, Malausa T, Nedved O, Rhule E, Staverlokk A, Steenberg T, Estoup A. 2014. Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Mol Ecol.* 23: 5979–5997.
- Matsumura S. 1931: 6000 illustrated insects of Japan-empire (in Japanese). Toko shoin, Tokyo, 1497 pp.
- Meng XL. 1994. Posterior predictive p-values. *Annals Stat.* 22: 1142-1160.
- Nei M, Maruyama T, Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution.* 29: 1–10.
- Nei M. 1987. Molecular Evolutionary Genetics. *New York: Columbia University Press.*
- Ometto L, Cestaro A, Ramasamy S, Grassi A, Revadi S, Siozios S, Moretto M, Fontana P, Varotto C, Pisani D, Dekker T, Wrobel N, Viola R, Pertot I, Cavalieri D, Blaxter M, Anfora G, Rota-Stabelli O. 2013. Linking genomics and ecology to investigate the complex evolution of an invasive drosophila pest. *Genome Biol. Evol.* 5, 745–757.
- Pascual M, MP Chapuis MP, Mestres F, Balanyà J, Huey RB, GW Gilchrist GW, Serra L, Estoup A.

2007. Introduction history of *Drosophila subobscura* in the New World: a microsatellite based survey using ABC methods. *Mol Ecol.* 16: 3069-3083.
- Peischl S, Excoffier L. 2015. Expansion load: recessive mutations and the role of standing genetic variation. *Mol. Ecol.* 24: 2084–2094.
- Pudlo P, Marin JM, Estoup A, Cornuet JM, Gautier M, Robert CP. 2016. Reliable ABC model choice via random forests. *Bioinformatics.* 32: 859–866.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Robert CP, Cornuet JM, Marin JM, Pillai NS. 2011. Lack of confidence in approximate Bayesian computation model choice. *Proc Natl Acad Sci USA.* 108: 15112-15117.
- Rius M, Darling JA. 2014. How important is intraspecific genetic admixture to the success of colonising populations? *Trends Ecol Evol.* 29: 233–242.
- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE. 2001. The population biology of invasive species. *Annu Rev Ecol Syst.* 305-332.
- Schug MD, Hutter CM, Noor MA, Aquadro CF. 1998. Mutation and evolution of microsatellites in *Drosophila melanogaster*. in: Mutation and Evolution. Springer, pp. 359–367.
- Simberloff D. 2013. Biological invasions: Much progress plus several controversies. *Contrib Sci.* 9: 7-16.
- Turgeon J, Tayeh A, Facon B, Lombaert E, De Clercq, Berkvens N, Lungren JG, Estoup A. 2011. Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *J Evol Biol.* 24: 1044-1052.
- Uller T, Leimu R. 2011. Founder events predict changes in genetic diversity during human-mediated range expansions. *Glob Change Biol.* 17: 3478–85.
- Verdu P, Austerlitz F, Estoup A, Vitalis R, Georges M, Théry S, Froment A, Le Bomin S, Gessain A, Hombert JM, Van der Veen L, Quintana-Murci L, Bahuchet S, Heyer E. 2009. Origins and genetic diversity of pygmy hunter-gatherers from Western Central Africa. *Curr Biol.* 19 : 312–318.
- Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM. 2009. Something in the way you move: dispersal pathways affect invasion success. *Trends Ecol Evol.* 24: 136–144.
- Wray G. 2013. Genomics and the evolution of phenotypic Traits. *Annu Rev Ecol Evol Syst.* 44:51-72.
- Yoshida T, Goka K, Ishihama F, Ishihara M, Kudo S. 2007. Biological invasion as a natural experiment of the evolutionary processes: introduction of the special feature. *Ecol Res.* 22 : 849–854.

Zenni RD, Lamy JB, Lamarque LJ, Porté AJ. 2014. Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: implications for tree invasion biology. *Biol Inv.* 16: 635–644.

Chapitre 2 : Génétique quantitative, évolution de la matrice G et divergence phénotypique.

Article 3.

Frainout, A., Navarro, N., Chantepie, S., Kimura, M.T., Teplitsky, C. & Debat, V.
Quantitative genetics of *Drosophila suzukii* wing shape: early phases of divergence during a worldwide invasion. *in prep.*

Quantitative genetics of *Drosophila suzukii* wing shape: early phases of divergence during a worldwide invasion

Antoine Fraimout, Nicolas Navarro, Stéphane Chantepie, Masayoshi Watada, Masahiro T. Kimura, Celine Teplitsky* and Vincent Debat*

Abstract

How phenotypic evolution is driven by evolutionary mechanisms of selection and drift is a long standing question in evolutionary biology. Specifically, how divergence is constrained by the genetic correlations summarized in the **G** matrix and how **G** itself evolves throughout period of drift and selection needs more empirical work. One general issue in studies of phenotypic evolution in extent populations is the lack of an appropriate ancestral population to compare patterns of divergence to. Biological invasions represent a particularly convenient situation to address this issue as they provide both ancestral and derived populations available for measurement of quantitative traits. Moreover, invasions often involve complex demo-genetic events including bottlenecks and admixtures, as well as putative selective pressures and genetics drift acting throughout the introduction and spread in the new range. Here we study the phenotypic divergence and the evolution of the **G** matrix estimated from wing shape data in the highly invasive pest *Drosophila suzukii*. The history of this global invasion was recently characterized by population genetics methods providing a particularly favorable situation to investigate the evolution of genetic architecture. We compared the **G** matrices of ancestral populations to derived populations in two major invaded ranges in Europe and USA, and finely quantify phenotypic divergence in wing shape using geometric morphometrics. Our results suggest that neutral evolution has most probably shaped some aspect of the eigenstructure of **G** as well as phenotypic divergence along the direction of the ancestral **G** matrix. However, we argue that stabilizing selection on wing shape may have maintained stability of **G** throughout the invasion process.

Key words: *Drosophila suzukii*, G-matrix, geometric morphometrics, biological invasion, quantitative genetics.

Introduction

Aside from their harmful effects on ecosystems and human societies, biological invasions represent promising models for the study of rapid evolution as they typically involve environmental change, new selective regimes and demographic shuffling (Keller & Taylor 2008). The extent to which the colonization of a new environment can trigger evolutionary responses in introduced populations has been a central topic of discussion in the past decade of invasion genetics studies (Prentis *et al.* 2008, Eroukhmanoff 2009, Bacigalupe *et al.* 2013). Evolutionary potential of invasive populations mostly depends on the available genetic variation, as phenotypic evolution requires heritable variation for selection to act upon. However the amount and nature of this genetic variation is itself highly dependent on processes acting early on in the invasion (e.g. depletion of genetic variation from founder effect, strong selective pressure or genetic drift). Understanding how quantitative genetic variation and genetic architecture evolve in the early phases of an invasion is a requisite to assess how they may affect the response to selection and potentially facilitate the success of the invasion. Although often discussed theoretically (e.g. Lee 2002; Bacigalupe 2009; Eroukhmanoff 2009), this question requires more empirical investigation (Bock *et al.* 2015, Dlugosch *et al.* 2015).

Most phenotypic traits are not independent within organisms and thus do not evolve independently. At the genetic level, these relationships are summarized in **G**, the genetic (co)variance matrix. A predicted response to selection integrating the effect of these genetic covariances can be obtained from the multivariate breeder's equation (Lande 1979):

$$R = \mathbf{G}\beta$$

where the response to selection R equals the product of **G** and the directional selection gradient β . Depending on the direction of selection, the genetic correlations among traits can facilitate or constrain response to selection over one generation and ultimately the rate of evolution (Agrawal & Stinchcombe 2009, Fischer *et al.* 2016). Particularly, an alignment between the major axis of **G** (i.e. g_{\max}) and the selection gradient (i.e. high genetic variance in the combination of selected traits) will induce a stronger response to selection towards the fitness optimum and an easier adaptive divergence in that direction (Steppan *et al.* 2002, Arnold *et al.* 2008). In a situation where g_{\max} is not aligned with selection, the evolutionary response would nevertheless be biased towards g_{\max} due to genetic correlations, thereby constraining evolution along these lines of least evolutionary resistance (Schluter 1996). In the context of evolutionary diversification, it is thus expected that the **G** matrix of an ancestral population would constrain derived populations to diverge along these lines. However, how **G** will impact evolutionary trajectories and our ability to predict them will depend on its stability.

The stability of **G** has been a subject of discussion among evolutionary biologists (Roff 2000, Steppan *et al.* 2002, Bégin & Roff 2003, Arnold *et al.* 2008), and biological invasions provide an exceptional context to address this issue empirically (Eroukhmanoff 2009, Bacigalupe 2011,

Franks *et al.* 2012). Evolutionary mechanisms such as selection, migration and genetic drift have been shown to influence the stability and structure of **G**, both theoretically (e.g. Jones *et al.* 2003, Guillaume & Whitlock 2007), and empirically under natural and laboratory conditions (e.g. Philips *et al.* 2001, Cano *et al.* 2004, Eroukhmanoff and Svenson 2011). All of these mechanisms are likely to be involved in the process of biological invasion at different stages (Lee 2002, Prentis *et al.* 2008, Rius & Darling 2014). At the early stage of invasion, depletion of genetic variation from random sampling of individuals in the native range and founder effect (i.e. demographic bottleneck) is expected. Strong selective pressures are also expected during these early stages of invasion as a result of the mismatch between native and invasive environments (Yoshida *et al.* 2007). During later stages of the invasion, populations at the margin of the distribution (i.e. expanding populations) are also likely to experience changes in their allele frequencies due to repeated founder events and genetic drift (Oliveri 2009) or repeated selection events across heterogeneous environments (Parker *et al.* 2003, Kilkenny & Galloway 2012). Evolution during biological invasions have been shown to affect the structure of **G** in several cases (e.g. Calsbeek *et al.* 2011, Bacigalupe 2013) but how the combined effects of neutral, selective and demographic mechanisms are expected to shape the structure of **G** during the course of an invasion remains unclear from empirical data.

Here we focus on the rapid worldwide invasion of *Drosophila suzukii* (Matsumura 1931) to investigate the potential effects of demo-genetic events associated with the early steps of the invasion process on the structure and stability of **G**. This Asian species is a highly invasive pest which has recently invaded Europe and North America simultaneously (see Asplen *et al.* 2015 for review). Using a population genetics approach, we have recently deciphered its invasion routes, allowing us to unambiguously identify populations historical relationships (Framout *et al.* 2015, 2016; see material and methods). We thus use natural populations of known history, and contrast samples from the native range and the European and North-American invasive ranges. For each of these populations, we estimate **G** matrices of wing shape derived from geometric morphometric analysis. The evolution of *Drosophila* wing shape has been studied at both the micro- and macro-evolutionary levels. Houle *et al.* (2003) have reported a clear phylogenetic signal in wing shape across species, that however cannot fully account for the observed divergence, suggesting a role for natural selection on shape evolution. Overall, wing shape nevertheless appears relatively stable across *drosophilids* (Houle *et al.* 2003, Hansen & Houle 2004, 2008). This stability does not result from absolute genetic constraints (i.e. a lack of genetic variation), as artificial selection studies in the lab have demonstrated that wing shape, at least in *D. melanogaster*, is highly evolvable (Weber 1992, Mezey & Houle 2005). These results suggest that stabilizing selection rather than absolute genetic constraints (i.e. a lack of variation) might be responsible for the overall stability found in the family, either directly acting on wing shape or indirectly via pleiotropic effects on correlated traits (e.g. Hansen & Houle 2004, 2008). This is somewhat paradoxical as within species, rapid evolution

of wing size and shape has been documented in several *Drosophilid* species, with marked latitudinal clines having repeatedly evolved in nature (Azevedo *et al.* 1998, Huey *et al.* 2000, Gilchrist *et al.* 2001, 2004). Overall the adaptive value of wing shape variation remains controversial. Therefore, and because of its implication in flight and courtship in *Drosophila*, investigating the evolution of wing shape at the early stages of *D. sukii's* invasion is particularly appealing.

Using an isofemale line breeding design (Hoffmann and Parsons, 1988, David *et al.* 2005, Berger *et al.* 2013) and a Bayesian statistical framework (Hadfield 2010), we study six natural populations of *D. sukii* and ask whether they retained signs of their invasion history in their genetic architecture. We address the following questions: i) Can we observe structural differences in **G** between native and invasive species? ii) Can neutral patterns of variation identified in *D. sukii* invasion history explain differences in **G** among native and invasive populations? iii) Can we observe signature of selection from the correlation between divergences and the major axes of the ancestral **G** matrix?

Specifically, at least three hypotheses can be tested in this context. The null hypothesis (H_0), would be that due to the very short time scale at which our study takes place (i.e. less than ten years of invasion), no divergence in mean phenotype and/or its variation would be detectable. Absence of divergence could also be due to high similarities between native and invasive habitats for the species, thus implying no adaptive challenges upon invasion (Facon *et al.* 2006, Hufbauer *et al.* 2012, Bacigalupe *et al.* 2013). Alternatively (H_1) some divergence might be detected among populations, but resulting from neutral evolution. In this situation, the known patterns of neutral genetic variation from *D. sukii's* invasion history (Frainout *et al.* 2016) suggests several predictions: 1) Regarding phenotypic means, we would expect divergence proportional to the molecular distance, the highest divergence separating populations exhibiting the highest F_{ST} values. 2) In terms of effects on the **G** matrix, one would mostly expect quantitative effects on the amount of variation (i.e. volume of the matrix) altered by bottleneck and genetic drift. For instance, drift can affect many aspects of the **G** matrix such as shape and orientation (Roff 2000, Philips *et al.* 2001, Whitlock *et al.* 2002). Here because of founders effect, we mostly expect a reduction in the volume of **G**, reflecting lower additive genetic variance in small populations. Additionally, we expect that such variation in volume would be proportional to the amount of neutral genetic variation present in the populations. For example, admixed populations would present higher volumes. However neutral genetic variation is generally a poor predictor of quantitative variation (Reed & Frankham 2001, Sgro *et al.* 2011), and the validity of this prediction is thus questionable. Finally (H_2), selection may also be at play. Although detailed, consensual predictions on the effect of selection on **G** are not available (e.g. Roff 2000, Jones *et al.* 2004, Chapuis *et al.* 2008, Doroszuck *et al.* 2008, Arnold 2008) one expected effect of directional selection on **G** is an alignment with the gradient of selection through correlated selection (Philips & McGuigan 2006,

Arnold *et al* 2008; Eroukhmanoff 2009). Two additional hypotheses can thus be tested in relation to selective effects in the invasion of *D. suzukii*. If the phenotypic divergence among populations is driven by selection then: (1) the ancestral Asian **G** might have constrained the divergence among populations by biasing it along its main directions (Schluter 1996). One would thus expect a correlation between the divergence matrix and the ancestral **G** matrix. (2) Selection might have induced an alignment of **G** along the selection gradient. If divergence from the ancestral population is indeed adaptive, one should thus expect **G** matrix of each derived invasive population to be aligned with the corresponding direction of divergence and thus possess more additive genetic variance along this direction.

Material and methods

Samples, lines and data acquisition

Drosophila suzukii adults were sampled using banana bait traps and net swiping at six localities representative of their global distribution range in Japan (native range), USA and France (invasive range) (Table 1). For each range, two populations were sampled, representative of a 'core' population and a 'margin' population corresponding to a North/South latitudinal gradient, as well as representative of the invasion history in USA and France. For the Japanese flies, a population from Sapporo (Hokkaido island, North) and one from Tokyo (Honshu island, South) were sampled. In both localities, traps were placed in gardens of the university campus. In the invasive ranges, the northern population was sampled in Dayton (Oregon) and the southern one in Watsonville (California) for the US samples. Sampling habitats corresponded to vineyards in Dayton, and strawberry fields in Watsonville. French samples were collected in Paris (North) in the botanical garden of the Natural History Museum and in Montpellier (South) close to vineyards. An average of 30 isofemale lines (Hoffmann & Parsons 1988, David *et al.* 2005; Berger *et al.* 2013) were established for each locality by performing single matings in separated rearing vials and expanding F1 offspring in a subsequent series of replicate vials. Rearing medium consisted of corn-starch and yeast with added anti-biotic (hydroxy-4 benzoate). Lines were randomly positioned in incubators and kept for five generations at $22 \pm 1^\circ\text{C}$. For each line from each population, 5 to 10 males *D. suzukii* were then randomly chosen for phenotyping. In total, 1323 individual wings distributed across 177 isofemale lines were phenotyped (Table 1). Images of the wings were acquired with a Leica DFC 420 digital camera mounted on a Leica Z6 APO microscope. Fifteen landmarks were defined on the dorsal face of the right wing (Fig. 1) and digitized using an ImageJ plugin written by Christian Klingenberg (personal communication).

Morphometric and statistical analyses

Procrustes generalized least squares superimposition was used to extract shape information from

Name of the population	Historical status	Type of population	Geographical coordinates	n° of isofemale lines	n° of individuals phenotyped
Tokyo	Native	core	35° 41' 22.155" N 139° 41' 30.143" E	31	192
Sapporo		margin	43° 3' 43.545" N 141° 21' 15.754" E	24	192
Watsonville	Invasive	core	36° 54' 36.832" N 121° 45' 24.82" E	26	191
Dayton		margin	45° 13' 14.422" N 123° 4' 34.368" E	38	313
Montpellier		core	43° 36' 38.768" N 3° 52' 36.177" E	33	226
Paris		margin	48° 51' 23.81" N 2° 21' 7.998" E	25	209

Table 1. Population of *D. sukuzii* used in this study with historical, geographical and data informations.

The range of distribution of each population is indicated in the column 'Historical status'. Differences among population within countries are specified in the 'Type of population' column where a 'core' population corresponds to a southern latitude population and/or a population of first records, and a 'margin' population corresponds to a northern latitude population and/or a dispersing population from the origin point of the invasion. Informations about the data are presented in the two last columns of the table.



Figure 1. Fifteen landmarks used to quantify wing shape of *D. sukuzii*.

Homologous landmarks are placed at the veins intersection and superimposed using the Procrustes superimposition method.

the landmarks coordinates (GLS, e.g. Dryden & Mardia 1998). Shape data have been used in several quantitative genetics studies (Klingenberg & Leamy 2001, Monteiro 2003, Hansen & Houle 2004, Mezey & Houle 2005, Myers *et al.* 2006, Adams 2011, Leinonen *et al.* 2011; Martinez-Abadias *et al.* 2012; see Klingenberg, 2003, 2010 ; and Polly 2008 for discussions) and are suitable for the estimation of **G**. A PCA was first applied to the set of 30 coordinates and the non-null 26 components were conserved as input data. Centroid size was used as a size variable in the subsequent analyses and compared between all pairs of populations.

As divergence in wing shape appeared to be weak on PCA plots we further performed a between-group principal component analysis (bgPCA) to visualize the overall variation in wing shape among populations (e.g. Mitteroecker & Bookstein 2011). We then estimated the statistical significance of the shape divergence among populations by applying a multivariate analysis of variance (MANOVA) with population as a main effect, combined with permutational pairwise MANOVAs, using the RVAideMemoire R package (v.0.9-60, Hervé 2016).

To investigate a possible effect of allometry (shape divergence correlated with size), we applied a multivariate regression of shape on size. This procedure was legitimate as allometry was found to be very low within populations (not shown), suggesting that no bias due to pooling populations with different allometries was expected in the global regression (e.g. Klingenberg & Monteiro 2005; Perrard *et al.* 2012).

G matrix estimations

We estimated the six **G** matrices corresponding to the six sampled populations under a Bayesian framework using the R package MCMCglmm (version 2.22.1; Hadfield 2010). For each population, we separately estimated the **G** matrix from a multivariate model with wing shape as a response variable and isofemale line as a random effect. Thus, in this design, traits (co)variances among lines represent broad sense genetic (co)variances. We also estimated the pooled within-group genetic covariance matrix from both Japanese populations datasets by running a model with an added fixed effect for population, and used this matrix as a proxy for the ancestral **G** matrix. Wing shape data corresponded for each population to the 11 first (out of the 26) non-null principal components which explained 92.4% of the variance. To assess prior sensitivity, we ran the models using several sets of priors. We used both informative and non-informative priors and the posterior distributions of (co-)variances did not depend on the prior (results not shown). We retained a prior set with variance estimates following a 'broken-stick' model of decrease of variance, appropriate for the use of PC scores representing decreasing proportion of explained variance (Jolliffe 2002), with null covariances and a degree of belief $\nu = n + 0.002$ (where n is the number of traits, Houle & Meyer 2015). We ran each model for 1 200 000 iterations with a burnin period of 200 000 and a

thinning rate of 1000. The convergence of models was checked visually from plots of MCMC chains and posterior distributions as well as autocorrelation level (<0.05 in all models, results not shown).

Matrices comparisons

For each pair of **G** matrices among and within continents, we calculated the pairwise difference using a variety of metrics describing the size, shape, orientation and similarity between matrices (Robinson & Beckerman 2013, Teplitsky *et al.* 2014). The Bayesian framework of these analyses allowed for tests of significance for each metric, as combinations of MCMC posterior values allow to calculate confidence intervals.

Size of the matrices (i.e. volumes) were calculated as the sum of their eigenvalues. We estimated the shape of the matrix by its eccentricity, (i.e. the ratio of the first, to the second eigenvalue; Jones *et al.* 2003). We also estimated the difference in the proportion of variance along \mathbf{g}_{\max} for each pairwise comparisons of **G**. Finally differences in orientation were estimated by performing the vector comparison algorithm implemented in the MorphoJ software (Klingenberg 2011) to compare angles between pairs of \mathbf{g}_{\max} . This test assesses statistical significance by comparing the observed angle to a distribution of angles between random vectors of the same dimensionality (see Klingenberg & Marugan-Lobon 2012). More specifically, we compared angles between the first eigenvector of the ancestral **G** matrix (computed as the pooled within-group **G** matrix of the Sapporo and Tokyo samples) and all four derived **G** matrices from invasive populations. We also compared the first eigenvectors of the ancestral **G** matrix to those of the total divergence matrix **D**, computed as the covariance matrix based on the populations mean shapes.

Ovaskainen's d

Ovaskainen *et al.* (2008) developed a measure of distance between the probability distributions of two matrices called 'd', which can be used to assess the statistical support for a difference in probability distributions.

$$d = \left(\frac{1}{2} \int_{R^d} \frac{[f(x) - g(x)]^2}{f(x) + g(x)} dx \right)^{\frac{1}{2}}$$

Here, $f(x)$ and $g(x)$ are the density probability distributions of two genetic variance-covariance matrices on R^d , where x represents a random sample drawn in each distribution. From a practical perspective, to integrate this function, we drew numerous random samples x in the distributions ($n=10^6$) and approximated 'd' using the mean of all draws.

The statistical support for a difference in probability distributions based on the distance d can be

easily calculated by drawing two random samples (A, B) from the posterior distribution of the G-matrix in population j (G_{P_j}) and in population k (G_{P_k}) using the following equation:

$$\Psi_d(G_{P_j}, G_{P_k}) = \left[d(G_{P_j}^A, G_{P_j}^B) + d(G_{P_k}^A, G_{P_k}^B) \right] - \left[d(G_{P_j}^A, G_{P_k}^A) + d(G_{P_j}^B, G_{P_k}^B) \right]$$

The first part of the right hand side the equation represents the distance between the matrices within a population, and the second part represents the distance between populations. It is thus a test of whether the differences between populations are larger than the differences within population. The statistical support for the difference between two matrices can be evaluated by using all the point estimates of their posterior distributions as random samples and computing the 95% credible interval of Ψ_d (Ovaskainen *et al.* 2008).

Principal coordinate analyses

Principal coordinates analyses (PCOA; Mardia *et al.* 1979) were applied to the set of **G** matrices. This approach is convenient as it allows to graphically displays the similarities and differences among a set of matrices at once. This analysis, also referred to as non-metric multidimensional scaling, has been applied to a wide range of morphometric data and to diverse types of covariance matrices (Debat *et al.* 2006, 2009, Jamniczky & Hallgrímsson 2009, Mitteroecker & Bookstein 2009, Breno *et al.* 2013). PCOA is related to principal components analysis but uses a matrix of distance between each pairs of matrices as the input data. We used various metrics to measure distance among matrices, including one minus their correlation (*r* index) and the Riemannian metric proposed by Mitteroecker & Bookstein (2009). The principal coordinates are axes that successively account for the maximum amount of the information contained in the distance matrix. Such an ordination allows one to visualize the relationship among covariance matrices.

Projection along divergence axes

An expected effect of directional selection on G is an alignment of the matrix with the selection gradient (e.g. Arnold *et al.* 2008, Eroukhmanoff 2009). If the divergences are triggered by selection, one would thus expect each derived population to exhibit a higher amount of variation along the direction of its divergence from the common ancestral population. To test this hypothesis we projected each MCMC matrix samples onto the divergence axis defined by the contrast between the native ancestral population and each invasive population

$$G_{proj} = (\bar{inv} - \bar{anc})G(\bar{inv} - \bar{anc})^t = dGd^t$$

where **d** is the contrast vector between the ancestral (anc) and invasive (inv) population shape

means, t stands for the transpose. For the ancestral mean shape, we used the pooled average between the two Japanese populations. We computed the proportion of the total genetic variance expressed along this direction as:

$$V_g^{proj} = \frac{dGd^t}{\sum V_g}$$

Results

Wing shape variation

Although the divergence among populations was weak, the bgPCA shows a clear signal of geographic differentiation for wing shape, opposing the pairs of populations according to their geographic origin (i.e. Japan vs France, vs USA; Fig. 2). This shape divergence was confirmed by the significant population effect in the MANOVA (Wilk's lambda = 0.169, df = 5, $p < 0.001$). Pairwise comparisons among populations indicated significant differences between all pairs of populations.

It is noteworthy that the two French populations appear further apart from the Japanese and US populations on the first bgPC, suggesting a stronger phenotypic divergence.

Allometry was weak (about 3% of shape variation), and mostly triggered by the smaller size of the North American flies. The patterns of divergence between populations described above (with the French populations being the most different) was not affected when using the residuals of a multivariate regression of size on shape, suggesting this divergence is mostly independent of size.

Matrix comparisons

Size, shape and orientation

Results of the pairwise comparisons of the three metrics of **G** matrix structure can be found in Table 2. We found significant differences in the size of the **G** matrix among three pairs of populations: Watsonville (southern US) and Paris (northern French), Watsonville and Montpellier (southern French) and Watsonville and Dayton (northern US). The volume of the **G** matrices of both French populations were smaller than **G** from the Watsonville population but did not show differences among themselves. Matrices of the French populations were also marginally (not significantly) smaller than matrices of Japanese populations. The volume of the Watsonville **G** matrix was also significantly larger than the volume of the **G** matrix from the Dayton population. Overall, except for the Watsonville population, **G** matrices of the invasive populations showed a trend for a smaller volume than the matrices from the native range. No significant difference in eccentricity between any pairs of populations was detected. Finally, we compared the amount of variance distributed along *gmax*. This metric comparison did not reveal any differences among all

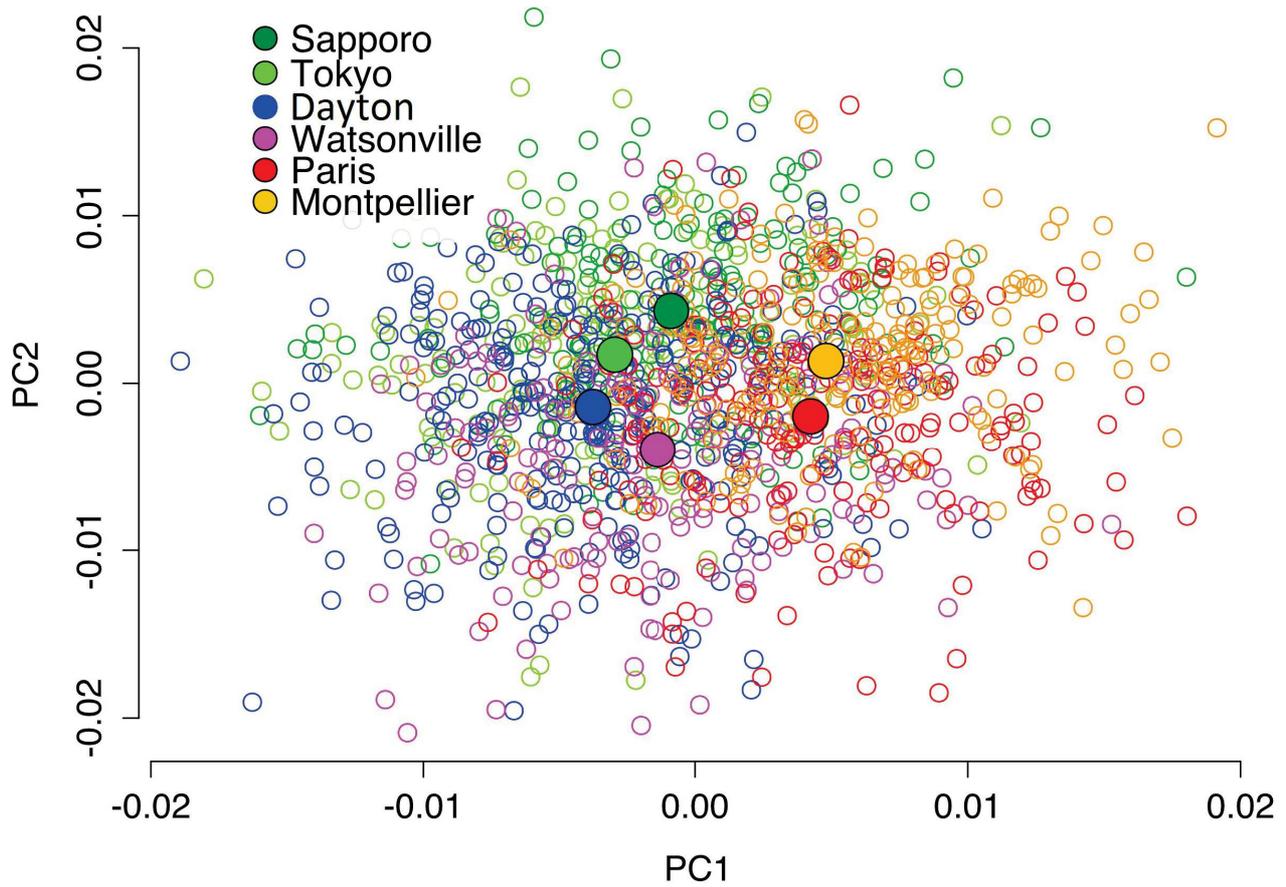


Figure 2. Plot of the between-group PCA.

Mean shape values for each populations were projected back onto the PCA space and are represented by the filled colored circle. Open circles represent individuals.

volume

Population	Tokyo	Sapporo	Dayton	Watsonville	Paris	Montpellier
Tokyo		0 [-0.28:0.08]	0 [-0.1:0.08]	-0.04 [-0.39:0.05]	0 [-0.08:0.09]	0 [-0.09:0.07]
Sapporo	1.21 [0.81:1.79]		-0.04 [-0.32:0.07]	0 [-0.15:0.09]	0 [-0.29:0.06]	0 [-0.26:0.07]
Dayton	0.88 [0.61:1.24]	0.74 [0.47:1.01]		-0.18 [-0.52:0]	0 [-0.11:0.07]	0 [-0.09:0.09]
Watsonville	1.44 [0.93:2.07]	1.15 [0.67:1.66]	1.53 [1.04:2.27]		-0.12 [-0.44:0.05]	-0.09 [-0.4:0.05]
Paris	0.88 [0.63:1.39]	0.76 [0.49:1.1]	1.01 [0.65:1.51]	0.7 [0.4:0.99]		0 [-0.09:0.09]
Montpellier	1 [0.66:1.37]	0.71 [0.49:1.12]	1.07 [0.74:1.51]	0.67 [0.44:0.99]	1.02 [0.65:1.46]	

eccentricity

Population	Tokyo	Sapporo	Dayton	Watsonville	Paris	Montpellier
Tokyo		0 [-0.38:0.31]	0 [-0.28:0.32]	0 [-0.45:0.34]	0 [-0.34:0.32]	0 [-0.51:0.44]
Sapporo	0.85 [0.43:2.06]		0 [-0.36:0.32]	0.01 [-0.5:0.47]	0 [-0.38:0.42]	0 [-0.56:0.4]
Dayton	1.04 [0.41:1.78]	0.84 [0.4:1.83]		0 [-0.43:0.37]	0 [-0.38:0.31]	0 [-0.48:0.41]
Watsonville	1.1 [0.45:2.23]	1.08 [0.42:2.22]	0.98 [0.49:2.29]		0 [-0.47:0.38]	0 [-0.43:0.54]
Paris	0.81 [0.43:2.02]	0.86 [0.38:1.95]	0.9 [0.44:2.08]	0.72 [0.31:1.78]		0 [-0.58:0.34]
Montpellier	1.09 [0.51:2.53]	0.89 [0.39:2.39]	1.01 [0.52:2.4]	0.86 [0.35:2.26]	1.13 [0.45:2.28]	

gmax

Population	Tokyo	Sapporo	Dayton	Watsonville	Paris	Montpellier
Tokyo		0 [-0.17:0.11]	0 [-0.11:0.09]	0 [-0.17:0.11]	0 [-0.12:0.1]	0 [-0.19:0.09]
Sapporo	0.97 [0.68:1.7]		0 [-0.12:0.11]	0 [-0.14:0.12]	0 [-0.14:0.1]	0 [-0.13:0.1]
Dayton	1.02 [0.69:1.54]	0.88 [0.6:1.39]		0 [-0.13:0.11]	0 [-0.09:0.12]	0 [-0.14:0.11]
Watsonville	1.19 [0.71:1.67]	0.97 [0.6:1.5]	0.94 [0.67:1.55]		0 [-0.13:0.1]	0 [-0.13:0.11]
Paris	1 [0.67:1.57]	0.86 [0.58:1.43]	0.91 [0.58:1.46]	0.79 [0.58:1.43]		0 [-0.15:0.11]
Montpellier	1.12 [0.75:1.71]	1.03 [0.65:1.61]	1.16 [0.73:1.62]	1.06 [0.67:1.58]	1.12 [0.67:1.66]	

Table 2. Pairwise comparisons for three metrics of G matrix shape.

On the upper diagonal is the difference between the two characteristics (volume, eccentricity or amount of genetic variance along gmax), so the test is whether the confidence interval encompass one. On the lower diagonal is the ratio between these variables (rows/column), so the test is whether the confidence interval includes 1. Significant differences are indicated by bold text.

pairwise matrix comparisons. On average, there was 35% of variance on \mathbf{g}_{\max} (Table 3). For angular comparisons, all matrices but Watsonville's \mathbf{G} present highly correlated eigenvectors (the three first only were compared, Table 3). For Watsonville, the first two eigenvectors are clearly inverted, suggesting that the main directions of \mathbf{G} are also shared by this population. Similarly, the total divergence matrix \mathbf{D} and the ancestral \mathbf{G} matrix (pooled within-group Japanese matrix) had highly correlated first two eigenvectors (Table 3).

Matrices distances and similarities

Measure of distance between probabilities of distribution (computed as Ovaskainen's d) of all pairwise comparisons of matrices did not reveal any dissimilarity among populations. Graphical representation of matrices distances using the Principal coordinates analysis nevertheless revealed a clear geographical signal using the r metric (Fig. 3). The first principal coordinate axis opposes matrices according to their continental origins. As observed for phenotypic data (Fig. 2), French populations were further apart from the Japanese and the US populations on this first axis, suggesting a stronger divergence. Furthermore, Japanese populations occupied a central position in each PCOA plots suggesting an intermediate distance from the two other sets of matrices. The second and third principal coordinate axes corresponded to within countries variation with no particular latitudinal consistency among samples. It is worth noticing that confidence intervals (as plotted by concentric colored lines in Fig. 2) did overlap along these two secondary axes, suggesting that these differences are weak.

Projection along divergence axes

We projected the proportion of total genetic variance of each \mathbf{G} matrix from each invasive population onto the axes of divergence between the ancestral mean shape and each invasive mean shapes (Fig. 4). For each of the four divergence axes, no significant difference was detected between the proportion of genetic variance expressed by the different \mathbf{G} matrices along the axis. We could however observe a trend for a higher proportion of variance expressed for \mathbf{G} matrices projected on their respective divergence axis.

Discussion

We conducted a quantitative genetics analysis of wing shape data in the highly invasive pest *D. suzukii*, to identify signs of phenotypic and genetic divergence in the early stages of this recent invasion. Knowing the demographic history of this recent invasion, we were able to formulate and test hypotheses regarding the stability and evolution of \mathbf{G} under this particular context. In that regard, it is worth stressing that although the most likely genetic source populations of Europe and US invasions were identified in continental Asia (i.e. China, Fraimout *et al.* 2016), F statistics and

a)	Montpellier		
Ancestral G matrix	PC1	PC2	PC3
PC1	0,00610	0,40696	<.00001
PC2	0,93588	0,00002	0,81726
PC3	0,00001	0,32838	0,12117

b)	Paris		
Ancestral G matrix	PC1	PC2	PC3
PC1	<.00001	0,84283	0,96492
PC2	0,72747	<.00001	0,70800
PC3	0,66525	0,18734	<.00001

c)	Watsonville		
Ancestral G matrix	PC1	PC2	PC3
PC1	0,13087	0,00152	0,04901
PC2	<.00001	0,00559	0,26689
PC3	0,28578	0,54850	<.00001

d)	Dayton		
Ancestral G matrix	PC1	PC2	PC3
PC1	<.00001	0,51559	0,35499
PC2	0,26131	0,00003	0,00182
PC3	0,73980	0,00014	0,00117

e)	Montpellier		
Paris	PC1	PC2	PC3
PC1	0,00047	0,37728	0,00127
PC2	0,37076	<.00001	0,00986
PC3	0,00052	0,59006	0,69636

f)	Watsonville		
Dayton	PC1	PC2	PC3
PC1	0,38405	0,02333	0,00003
PC2	0,00002	0,06356	0,72568
PC3	0,07422	<.00001	0,07931

g)	D matrix		
Ancestral G matrix	PC1	PC2	PC3
PC1	0,03493	0,89172	0,00283
PC2	0,37801	0,00173	0,17760
PC3	0,23025	0,00002	0,11672

Table 3. Vector comparisons.

Vectors comparison were performed between the ancestral **G** matrix and each invasive **G** matrix (a – d) and within each invaded range in France (e) and USA (f). Orientation of the **D** matrix and the ancestral **G** matrix was also performed (g). Bold text indicates significant correlation among axes.

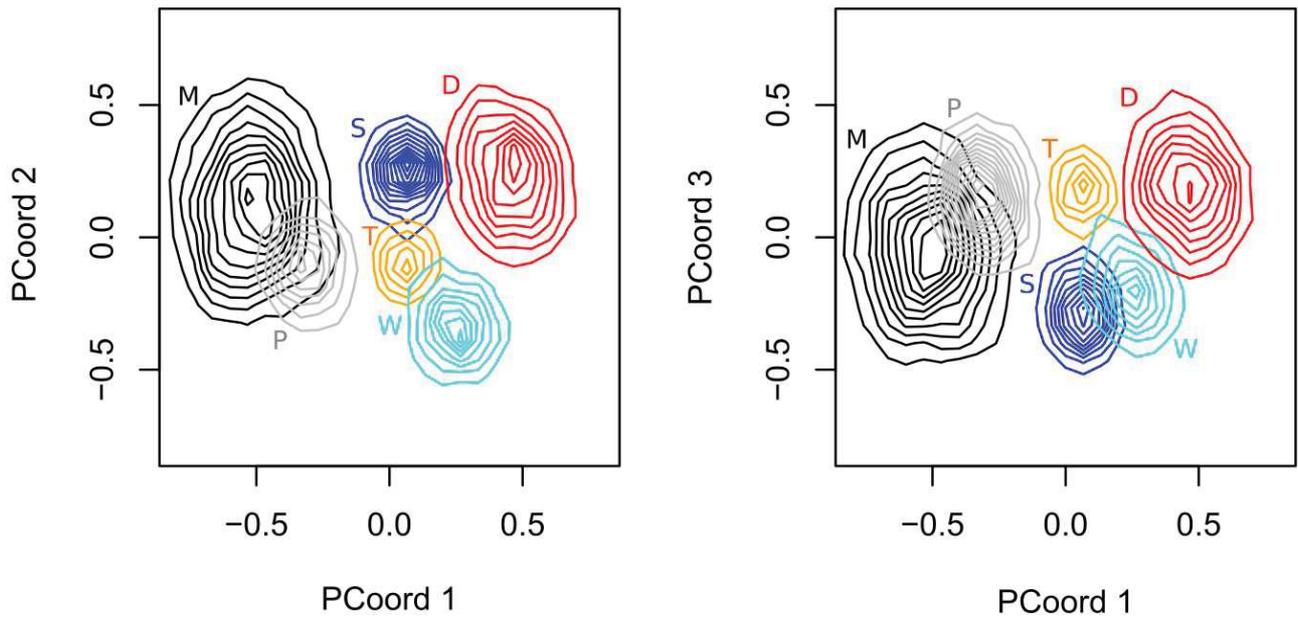


Figure 3. Graphical representation the PCOA.

Each matrix is represented by a color and the first initial of the population's locality (M, red: Montpellier; P, gray: Paris; S, blue: Sapporo; T, orange: Tokyo; D, red: Dayton; W, light-blue: Watsonville). Concentric circles represent the confidence interval obtained from 1000 MCMC samples.

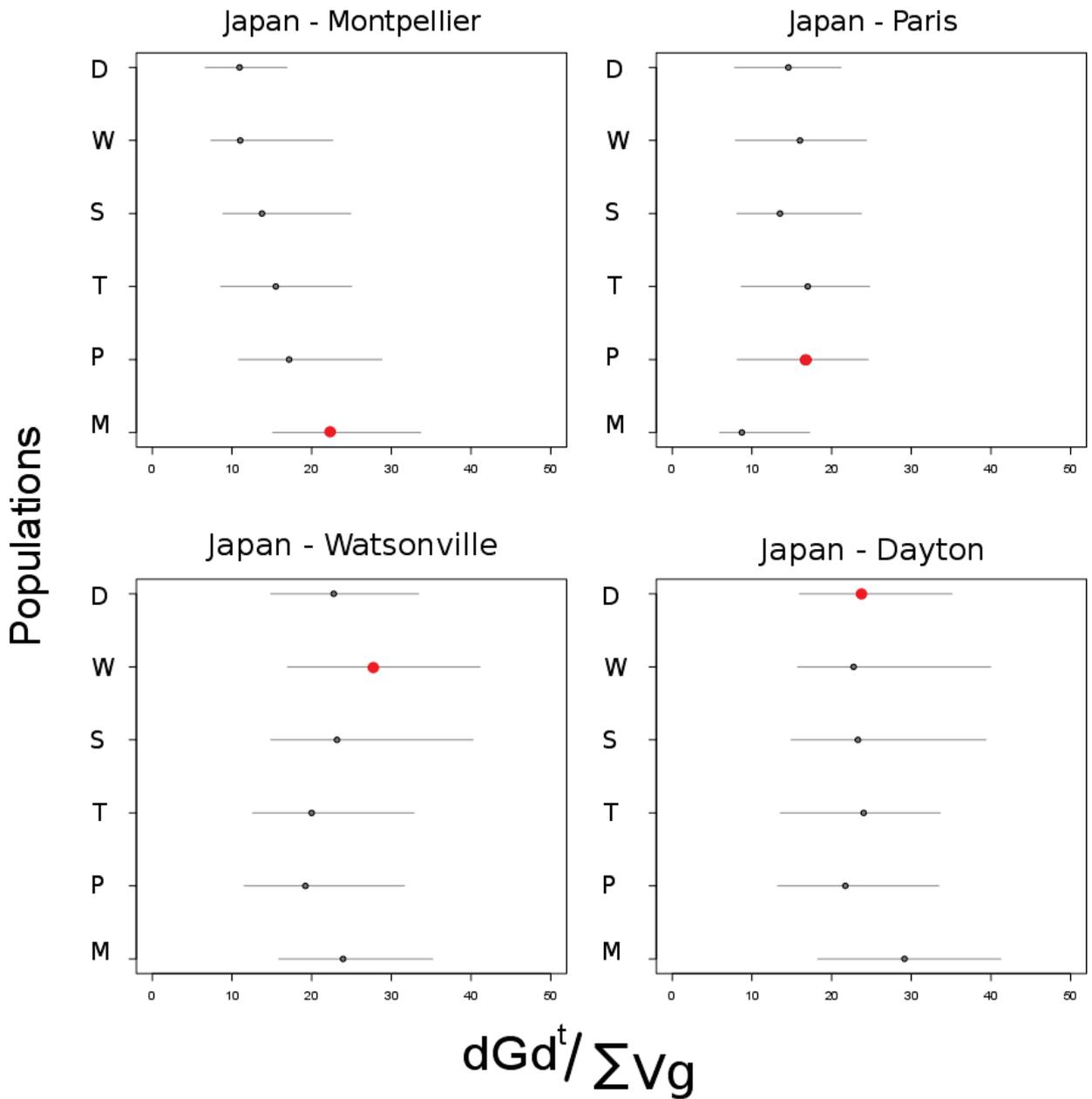


Figure 4. Matrices projection on divergence axes.

G matrices were projected along the four axes of divergence between the ancestral **G** matrix (i.e. the pooled within-group covariance matrix for Japanese populations) and all derived invasive populations: Axis 1 = Montpellier (M); Axis 2 = Paris (P), Axis 3 = Watsonville (W) and Axis 4 = Dayton (D).

Bayesian clustering methods reveal weak differentiation among these sources and our Japanese population samples, thereby suggesting that the use of Japanese populations as ancestral ones is an acceptable approximation.

*Divergence of **G** Matrices*

Under a quantitative genetics framework, we investigated if divergence in wing shape was reflected by a change in the genetic (co-)variance matrix **G**. Most studies investigating the evolution of **G** in natural populations usually perform comparisons between long diverged populations (e.g. Cano *et al.* 2004, Bacigalupe *et al.* 2013) and the actual pace at which **G** may evolve in natural conditions remains unclear. Evaluating the support for predictions regarding the stability of **G** its expected evolution and its effects on population divergence requires thus further empirical studies (McGuigan 2006, Revell 2007). As stated by Jones *et al.* (2003) “[...] *the current state of analytical theory on G-matrix evolution is that existing theory cannot guarantee stability of the G-matrix (Turelli 1988), but also does not guarantee instability.*” Following this assertion, computer based simulation studies on **G** stability (Jones *et al.* 2003, 2004, Revell 2007, Arnold *et al.* 2008) have identified specific parameters promoting the stability of **G** structure and orientation (e.g. correlative selection, population size, pleiotropic mutation). However, in an empirical context, these parameters remain largely unknown and one should thus use *a priori* informations on the evolutionary history of the focal population to consider if **G** is expected to be stable or not. In the case of *D. suzukii*, a relative instability of **G** could be expected from repeated bottlenecks, admixture, drift and the overall complexity of the invasion history (Fraitout *et al.* 2016). Such an instability in **G** would thus render difficult the interpretation of patterns of differences between **G** matrices. Especially, if stochastic processes of drift are dominating, derived **G** matrices could diverge in all directions from the ancestral **G** (Arnold *et al.* 2001). Overall, our results however suggest that most aspects of the eigenstructure of **G** (shape and orientation) have remained mostly stable throughout this complex demographic context, a somewhat surprising result. The small amount of time for divergence might be responsible for this low divergence. However, other studies have suggested that **G** might evolve remarkably quickly (e.g. Careau *et al.* 2015), suggesting that other factors might be at play. For example, it is conceivable that stabilizing selection on wing shape would contribute to constraining its variation and explain the stability of the **G** matrices despite destabilizing effects of demographic events.

We did find that the total amount of genetic variance of **G** (its volume) varied among populations, which relate to theoretical expectations regarding neutral evolution. Reduction in **G** volume is expected under certain scenarios such as small population size and genetic drift (Roff 2000, Jones *et al.* 2003, 2004). According to the results of Fraitout *et al.* (2016), all invasive populations of *D. suzukii* experienced loss of genetic diversity and various intensity of bottlenecks upon introduction.

Volume comparisons were consistent with the neutral predictions of lower volumes for matrices in invasive populations (although not significant) and higher volume of **G** for the most admixed population (in particular, the Watsonville population is highly admixed).

Divergence in populations means

All pairs of populations showed significant differences in wing shape and we were able to identify a geographically structured signal in the data. Is this divergence neutral and the effect of stochastic demographic sampling and genetic drift, or is it adaptive? The similarity between the ancestral **G** matrix and the divergence matrix is congruent with an hypothesis of pure drift.

However, comparing the alignment of each of the divergence axes to the major axis of their corresponding **G** matrix (e.g. projection of Paris **G** matrix onto the axis of divergence between Japan and Paris, Fig. 4), we observed a trend suggesting such an alignment – although not statistically significant. As selection is supposed to induce an alignment of **G** with the selection gradient, this result could be indicative of an ongoing adaptive divergence in its early stages.

However, we are lacking any evidence of the very nature of the selection that would be at play.

Geographical variation of wing morphology in *Drosophila* has been described in several naturally occurring species (e.g. Haas & Tolley 1998, Hoffmann & Shirriffs 2002) and is expected to occur as well in a rapid colonization context (Huey *et al.* 2000). Intraspecific variation in *Drosophila* wing morphology between geographic regions is often discussed as an adaptation to local environments, particularly when areas of distribution span various latitudes (e.g. James *et al.* 1997, Huey *et al.* 2000, Gilchrist *et al.* 2004). Parallel clines of wing morphology evolution in particular have been described in different *Drosophila* species and related to an adaptation to climatic conditions (Gilchrist *et al.* 2000). Despite the fact that northern and southern populations (i.e. core and margin) showed differences in wing shape, these differences were not consistent with latitudinal distributions. As the invasion of *D. sukikii* was initiated less than a decade ago in USA and Europe (Hauser 2011, Calabria *et al.* 2012) it is possible that adaptive geographical clines did not have enough time to evolve in this species. As a comparison, latitudinal clines in *Drosophila subobscura* after colonization of the new world, were reported after two decades although it could not be identified ten years after the start of the invasion (Huey *et al.* 2000, Gilchrist *et al.* 2004).

Conclusion

We observed an overall stability in **G** throughout the invasion of *D. sukikii* with the exception of matrices volume. Although this structural difference between matrices most likely resulted from neutral evolution (i.e. as measured by quantitative amount of neutral genetic variation, Fraimout *et al.* 2016), we argue that stabilizing parameters such as selection could explain the high stability observed among all of our **G** matrices comparisons. A more stringent test for deciphering the effects of selection and drift should involve direct comparisons of molecular and quantitative

variance such as Qst / Fst comparisons (e.g. Chapuis *et al.* 2008, Chenoweth & Blows 2008).

Acknowledgments

AF was supported by state funding by the Agence Nationale de la Recherche, through the LabEx ANR-10-LABX-0003-BCDiv, of the program “Investissements d’avenir” (ANR-11-IDEX-0004-02). The authors acknowledge K. Tamura, M. Toda, Peter Shearer, Todd Schlenke and Artyom Kopp for help in flies sampling. We thank Marion Guillaume for helping with the stock flies maintenance and Frédérique Peronnet for providing flies rearing medium.

References

- Adams, D.C., Rohlf, F.J., and Slice, D.E. (2013). A field comes of age: geometric morphometrics in the 21st century. *Hystrix* 24, 7–14.
- Agrawal, A.F., and Stinchcombe, J.R. (2009). How much do genetic covariances alter the rate of adaptation? *Proceedings of the Royal Society B: Biological Sciences* 276, 1183–1191.
- Arnold, S.J., Bürger, R., Hohenlohe, P.A., Ajie, B.C., and Jones, A.G. (2008). Understanding the evolution and stability of the G-matrix. *Evolution* 62, 2451–2461.
- Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., *et al.* (2015). Invasion biology of spotted wing *Drosophila* (*Drosophila suzukii*): a global perspective and future priorities. *Journal of Pest Science* 88, 469–494.
- Azevedo, R.B.R., James, A.C., McCabe, J., and Partridge, L. (1998). Latitudinal Variation of Wing: Thorax Size Ratio and Wing-Aspect Ratio in *Drosophila melanogaster*. *Evolution* 52, 1353.
- Bacigalupe, L.D. (2009). Biological invasions and phenotypic evolution: a quantitative genetic perspective. *Biological Invasions* 11, 2243–2250.
- Bacigalupe, L.D., Barrientos, K., Beckerman, A.P., Carter, M.J., Figueroa, C.C., Foster, S.P., Moore, A.J., Silva, A.X., and Nespolo, R.F. (2013). Can invasions occur without change? A comparison of **G**-matrices and selection in the peach-potato aphid, *Myzus persicae*. *Ecology and Evolution* 3, 5109–5118.
- Bégin, M., and Roff, D.A. (2003). The constancy of the G matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: a case study in crickets. *Evolution* 57, 1107.
- Berger, D., Postma, E., Blanckenhorn, W.U., and Walters, R.J. (2013). Quantitative genetic divergence and standing genetic (co)variance in thermal reaction norms along latitude: genetics of thermal reaction norm differentiation. *Evolution* 67, 2385–2399.
- Björklund, M. (2004). CONSTANCY OF THE G MATRIX IN ECOLOGICAL TIME. *Evolution* 58, 1157.
- Bock, D.G., Caseys, C., Cousens, R.D., Hahn, M.A., Heredia, S.M., Hübner, S., Turner, K.G.,

- Whitney, K.D., and Rieseberg, L.H. (2015). What we still don't know about invasion genetics. *Molecular Ecology* 24, 2277–2297.
- Breno, M., Bots, J., and Van Dongen, S. (2013). Heritabilities of Directional Asymmetry in the Fore- and Hindlimbs of Rabbit Fetuses. *PLoS ONE* 8, e76358.
- Burrack, H.J., Fernandez, G.E., Spivey, T., and Kraus, D.A. (2013). Variation in selection and utilization of host crops in the field and laboratory by *Drosophila suzukii* Matsumura (Diptera: Drosophilidae), an invasive frugivore: Selection and utilization of host crops by the invasive frugivore *D. suzukii*. *Pest Management Science* 69, 1173–1180.
- Calabria, G., Máca, J., Bächli, G., Serra, L., and Pascual, M. (2012). First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe: First record of *Drosophila suzukii* in Europe. *Journal of Applied Entomology* 136, 139–147.
- Calsbeek, B., Lavergne, S., Patel, M., and Molofsky, J. (2011). Comparing the genetic architecture and potential response to selection of invasive and native populations of reed canary grass: Genetic architecture of invasive populations. *Evolutionary Applications* 4, 726–735.
- Cano, J.M., Laurila, A., Palo, J., and Merilä, J. (2004). Population differentiation in G matrix structure due to natural selection in *Rana temporaria*. *Evolution* 58, 2013–2020.
- Careau, V., Wolak, M.E., Carter, P.A., and Garland, T. (2015). Evolution of the additive genetic variance–covariance matrix under continuous directional selection on a complex behavioural phenotype. *Proceedings of the Royal Society B: Biological Sciences* 282, 20151119.
- Chapuis, E., Martin, G., and Goudet, J. (2008). Effects of Selection and Drift on G Matrix Evolution in a Heterogeneous Environment: A Multivariate Qst-Fst Test With the Freshwater Snail *Galba truncatula*. *Genetics* 180, 2151–2161.
- Chenoweth, S.F., and Blows, M.W. (2008). Q_{ST} meets the G matrix: the dimensionality of adaptive divergence in multiple correlated quantitative traits. *Evolution* 62, 1437–1449.
- David, J.R., Gibert, P., Legout, H., Pétavy, G., Capy, P., and Moreteau, B. (2005). Isofemale lines in *Drosophila*: an empirical approach to quantitative trait analysis in natural populations. *Heredity* 94, 3–12.
- Debat, V., Milton, C.C., Rutherford, S., Klingenberg, C.P., and Hoffmann, A.A. (2006). Hsp90 and the quantitative variation of wing shape in *Drosophila melanogaster*. *Evolution* 60, 2529–2538.
- Debat, V., Debelle, A., and Dworkin, I. (2009). Plasticity, canalization, and developmental stability of the *Drosophila* wing: joint effects of mutations and developmental temperature. *Evolution* 63, 2864–2876.
- Dlugosch, K.M., Anderson, S.R., Braasch, J., Cang, F.A., and Gillette, H.D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology* 24, 2095–2111.

- Doroszuk, A., Wojewodzcic, M.W., Gort, G., and Kammenga, J.E. (2008). Rapid Divergence of Genetic Variance-Covariance Matrix within a Natural Population. *The American Naturalist* 171, 291– 304.
- Eroukhmanoff, F. (2009). Just How Much is the G-matrix Actually Constraining Adaptation? *Evolutionary Biology* 36, 323–326.
- Eroukhmanoff, F., and Svensson, E.I. (2011). Evolution and stability of the G-matrix during the colonization of a novel environment: G-matrix evolution. *Journal of Evolutionary Biology* 24, 1363–1373.
- Facon, B., Genton, B., Shykoff, J., Jarne, P., Estoup, A., and David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* 21, 130–135.
- Fischer, E.K., Ghalambor, C.K., and Hoke, K.L. (2016). Plasticity and evolution in correlated suites of traits. *Journal of Evolutionary Biology* 29, 991–1002.
- Fraimout, A., Loiseau, A., Price, D.K., Xuereb, A., Martin, J.-F., Vitalis, R., Fellous, S., Debat, V., and Estoup, A. (2015). New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest. *European Journal of Entomology*.
- Franks, S.J., Wheeler, G.S., and Goodnight, C. (2012). Genetic variation and evolution of secondary compounds in native and introduced populations of the invasive plant *Melaleuca quinquenervia*: evolution in an invasive plant. *Evolution* 66, 1398–1412.
- Gilchrist, A.S., Azevedo, R.B.R., Partridge, L., and O'higgins, P. (2000). Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. *Evolution & Development* 2, 114–124.
- Gilchrist, G.W., Huey, R.B., Balanyà, J., Pascual, M., and Serra, L. (2004). A time series of evolution in action: a latitudinal cline in wing size in South-American *Drosophila subobscura*. *Evolution* 58, 768.
- Guillaume, F., & Whitlock, M. C. (2007). Effects of migration on the genetic covariance matrix. *Evolution*, 61(10), 2398-2409.
- Haas, H.L., and Tolley, K.A. (1998). Geographic variation of wing morphology in three Eurasian populations of the fruit fly, *Drosophila lummei*. *Journal of Zoology* 245, 197–203.
- Hadfield, J.D., and others (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* 33, 1–22.
- Hansen, T.F., and Houle, D. (2004). Evolvability, stabilizing selection, and the problem of stasis. In Pigliucci & Preston. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford University Press (2004)
- Hansen, T.F., and Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* 21, 1201–1219.

- Hauser, M. (2011). A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Management Science* 67, 1352–1357.
- Hoffmann, A.A., and Parsons, P.A. (1988). The analysis of quantitative variation in natural populations with isofemale strains. *Genetique, Selection, Evolution* 20, 1.
- Hoffmann, A.A., and Shirriffs, J. (2002). Geographic variation for wing shape in *Drosophila serrata*. *Evolution* 56, 1068–1073.
- Houle, D., Mezey, J., Galpern, P., and Carter, A. (2003). Automated measurement of *Drosophila* wings. *BMC Evolutionary Biology* 3, 1.
- Huey, R.B., Gilchrist, G.W., Carlson, M.L., Berrigan, D., and Serra, L. (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287, 308–309.
- Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O., and Estoup, A. (2012). Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions: Anthropogenically induced adaptation to invade. *Evolutionary Applications* 5, 89–101.
- Jamniczky, H.A., and Hallgrímsson, B. (2009). A comparison of covariance structure in wild and laboratory murid crania. *Evolution* 63, 1540–1556.
- Jolliffe, I. (2002). *Principal component analysis*. John Wiley & Sons, Ltd.
- Jones, A.G., Arnold, S.J., and Bürger, R. (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* 57, 1747–1760.
- Jones, A.G., Arnold, S.J., and Burger, R. (2004). Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* 58, 1639.
- Keller, S.R., and Taylor, D.R. (2008). History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* 11, 852–866.
- Kilkenny, F.F., and Galloway, L.F. (2013). Adaptive divergence at the margin of an invaded range; adaptive divergence in margin populations. *Evolution* 67, 722–731.
- Klingenberg, C.P. (2010). Evolution and development of shape: integrating quantitative approaches. *Nature Reviews Genetics*.
- Klingenberg, C.P. (2011). MorphoJ: an integrated software package for geometric morphometrics: Computer program note. *Molecular Ecology Resources* 11, 353–357.
- Klingenberg, C.P., and Leamy, L.J. (2001). Quantitative genetics of geometric shape in the mouse mandible. *Evolution* 55, 2342–2352.
- Klingenberg, C.P., Mebus, K., and Auffray, J.-C. (2003). Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evolution & Development* 5, 522–531.

- Klingenberg, C., and Monteiro, L. (2005). Distances and Directions in Multidimensional Shape Spaces: Implications for Morphometric Applications. *Systematic Biology* 54, 678–688.
- Klingenberg, C.P., and Marugan-Lobon, J. (2013). Evolutionary Covariation in Geometric Morphometric Data: Analyzing Integration, Modularity, and Allometry in a Phylogenetic Context. *Systematic Biology* 62, 591–610.
- Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 402–416.
- Lee, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17, 386–391.
- Lee, J.C., Bruck, D.J., Curry, H., Edwards, D., Haviland, D.R., Van Steenwyk, R.A., and Yorgey, B.M. (2011). The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest Management Science* 67, 1358–1367.
- Leinonen, T., Cano, J.M., and Merilä, J. (2011). Genetics of body shape and armour variation in threespine sticklebacks: Genetics of threespine stickleback body shape. *Journal of Evolutionary Biology* 24, 206–218.
- Martínez-Abadías, N., Esparza, M., Sjøvold, T., González-José, R., Santos, M., Hernández, M., and Klingenberg, C.P. (2012). Pervasive genetic integration directs the evolution of human skull shape: genetic integration in the human skull. *Evolution* 66, 1010–1023.
- Matsumura S. 1931: 6000 illustrated insects of Japan-empire (in Japanese). Toko shoin, Tokyo, 1497 pp.
- Mezey, J.G., and Houle, D. (2005). The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* 59, 1027–1038.
- Mitteroecker, P., and Bookstein, F. (2011). Linear Discrimination, Ordination, and the Visualization of Selection Gradients in Modern Morphometrics. *Evolutionary Biology* 38, 100–114.
- Monteiro, L.R., Duarte, L.C., and dos Reis, S.F. (2003). Environmental correlates of geographical variation in skull and mandible shape of the punar rat *Thrichomys apereoides* (Rodentia: Echimyidae). *Journal of Zoology* 261, 47–57.
- Olivieri, I. (2009). Alternative mechanisms of range expansion are associated with different changes of evolutionary potential. *Trends in Ecology & Evolution* 24, 289–292.
- Ovaskainen, O., Cano, J.M., and Merila, J. (2008). A Bayesian framework for comparative quantitative genetics. *Proceedings of the Royal Society B: Biological Sciences* 275, 669–678.
- Parker, I.M., Rodriguez, J., and Loik, M.E. (2003). An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology* 17, 59–72.
- Phillips, P.C., Whitlock, M.C., and Fowler, K. (2001). Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* 158, 1137–1145.

- Phillips, P.C., and K.L. McGuigan. 2006. Evolution of genetic variance-covariance structure. In C. Fox and J. Wolf (eds.), *Evolutionary Genetics: Concepts and Case Studies*. Oxford University Press. Oxford, England.
- Polly, P.D. (2008). Developmental Dynamics and G-Matrices: Can Morphometric Spaces be Used to Model Phenotypic Evolution? *Evolutionary Biology* 35, 83–96.
- Prentis, P.J., Wilson, J.R.U., Dormontt, E.E., Richardson, D.M., and Lowe, A.J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science* 13, 288–294.
- Puentes, A., Granath, G., and Ågren, J. (2016). Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*: G matrix structure in populations of *A. lyrata*. *Evolution* 70, 2370–2386.
- Reed, D.H., and Frankham, R. (2001). How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55, 1095.
- Revell, L.J. (2007). The G matrix under fluctuating correlational mutation and selection. *Evolution* 61, 1857–1872.
- Rius, M., and Darling, J.A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution* 29, 233–242.
- Robinson, M.R., and Beckerman, A.P. (2013). Quantifying multivariate plasticity: genetic variation in resource acquisition drives plasticity in resource allocation to components of life history. *Ecology Letters* 16, 281–290.
- Roff, D. (2000). The evolution of the G matrix: selection or drift? *Heredity* 84, 135–142.
- Schluter, D. (1996). Adaptive Radiation Along Genetic Lines of Least Resistance. *Evolution* 50, 1766.
- Sgrò, C.M., Lowe, A.J., and Hoffmann, A.A. (2011). Building evolutionary resilience for conserving biodiversity under climate change: Conserving biodiversity under climate change. *Evolutionary Applications* 4, 326–337.
- Steppan, S.J., Phillips, P.C., and Houle, D. (2002). Comparative quantitative genetics: evolution of the G matrix. *Trends in Ecology & Evolution* 17, 320–327.
- Teplitsky, C., Robinson, M.R., and Merilä, J. (2014). Evolutionary potential and constraints in wild populations. *Quantitative Genetics in the Wild* 190–208.
- Weber, K. E. (1992). How small are the smallest selectable domains of form?. *Genetics*, 130(2), 345-353.
- Whitlock, M.C., Phillips, P.C., and Fowler, K. (2002). Persistence of changes in the genetic covariance matrix after a bottleneck. *Evolution* 56, 1968.
- Yoshida, T., Goka, K., Ishihama, F., Ishihara, M., and Kudo, S. (2007). Biological invasion as a natural experiment of the evolutionary processes: introduction of the special feature. *Ecological Research* 22, 849–854.

Chapitre 3 : Plasticité phénotypique et performance de vol chez *Drosophila suzukii*.

Article 4.

Fraimout, A., Jacquemart, P., Aponte, D.J., Descamps, T., Herrel, A., Cornette, R. & Debat, V. Phenotypic plasticity of *Drosophila suzukii* wing to developmental temperature: implications for flight performance. *in prep.*

Phenotypic plasticity of *Drosophila suzukii* wing to developmental temperature: implications for flight performance

Antoine Fraimout, Pauline Jacquemart, David J. Aponte, Thierry Descamps, Anthony Herrel, Raphaël Cornette & Vincent Debat.

Abstract

Phenotypic plasticity has been proposed as a facilitating mechanism in the success of biological invasions. If plastic traits are advantageous in the new range, selection can act on this plastic variation and induce post-invasion increase of phenotypic plasticity. In order to test the hypothesis of increased plasticity in introduced populations, particular attention should be paid to the relationship between the focal plastic trait, the environmental stimulus and the performance of the trait. *Drosophila* wing is a particularly relevant candidate for experimental studies on phenotypic plasticity. Wing morphology is known for its plastic variation under different experimental temperatures, but this plastic response has rarely been investigated in a functional context of flight performance. Here, we investigate the effect of temperature on wing morphology and flight performance of the invasive pest *Drosophila suzukii*. This highly invasive *drosophilid* species has rapidly dispersed across continent through human activities, but is also expected to actively disperse in the invasive range. By finely quantifying wing morphology and flight trajectories of individuals raised under different temperature treatments, we test whether invasive populations of *D. suzukii* show higher phenotypic plasticity than native one. Moreover, we investigate the effect of the plasticity in wing morphology on flight performance. Our results suggest that (1) no difference in phenotypic plasticity can be observed between invasive and native populations and (2) temperature has a strong effect on wing morphology and flight performance (in particular speed and acceleration), indicative of a possible adaptive value of wing plasticity.

Key words: *Drosophila suzukii*, phenotypic plasticity, flight performance, aspect-ratio, geometric morphometrics, wing shape

Introduction

Phenotypic plasticity – the capacity of a genotype to produce different phenotypes in different environments – is a ubiquitous property of organisms in response to environmental stimuli (West-Eberhard 1989, 2003). Its role in facilitating adaptation to new conditions and ultimately, evolution, has been proposed in the literature (e.g. Pigliucci 2005, Friedrich & Meyer 2016) and plasticity is in many cases thought to be adaptive (e.g. Via *et al.* 1995, Ghalambor *et al.* 2007). The adaptive nature of phenotypic plasticity relies on the assumption that plastic individuals may express advantageous phenotypes (i.e. phenotypes close to a fitness optimum, Schlichting 1986, Price *et al.* 2003) in response to an environmental change, thus increasing their fitness. By inducing novel phenotypic variation in the population upon which natural selection can act, phenotypic plasticity has the potential to set Darwinian evolution in motion without prior alteration of the genetic makeup of the plastic trait (Levis & Pfennig 2016). Increased survival of plastic individuals in changing environmental conditions, or during the colonization of new habitats have thus been suggested (Ghalambor *et al.* 2007). The latter situation has brought phenotypic plasticity among the hypotheses which could explain the success of biological invasions (e.g. Richards *et al.* 2006, Davidson *et al.* 2011). Invasive species are among the most successful colonialists in their ability to survive, reproduce and disperse in new environments (Lee 2002). The underlying biological mechanisms associated to such success are however still debated and in need for empirical studies (Bock *et al.* 2015, Estoup *et al.* 2016). One consensual expectation however is that the adaptive potential of invasive species (i.e. their ability to respond to selection) can be enhanced by genetic variance (Lavergne & Molofsky 2007) and/or phenotypic plasticity (Yu-Peng *et al.* 2007). Plasticity may for instance promote phenotypic adjustment to new conditions after introduction in cases where genetic variation in the direction of selection is lacking due to demographic bottlenecks or genetic drift.

A common question in invasion biology is whether the remarkable success of invasive species could be explained by higher phenotypic plasticity compared to non-invasive species (Richards *et al.* 2006, Davidson *et al.* 2011). This macro-evolutionary approach is however limited by how the phylogeny may impede comparisons of traits across species (Funk 2008). It has been proposed that such comparisons between native and invasive ranges would benefit from the use of closely related species (Muth & Pigliucci 2006). One way to circumvent the issue of phylogenetic independence is therefore to directly compare invasive populations within a species to their native counterparts (Kaufman & Smouse 2001). This micro-evolutionary approach furthermore allows to address the question of underlying genetic variation for plastic traits (i.e. G x E interactions; Ghalambor *et al.* 2007) from population and quantitative genetics point of view. Within-species population differences in phenotypic plasticity have been also suggested in biological invasions, with invasive populations expected to be more plastic and perform better in their new range (Parker *et al.* 2013). However the relationship between plasticity and the performance of a

population is fairly difficult to assess experimentally as it requires several assumptions (reviewed in Forsman 2015). For instance, performance in relation to the investigate trait(s) needs to be rigorously defined. Specifically, as plasticity may produce phenotypic variation in non-adaptive directions (Ghalambor *et al.* 2007) the significance of the focal trait and its function for the organism need to be assessed *a priori* (Forsman 2015). Similarly, the environmental stimulus used to induce plasticity should reflect ecological significance for the species.

Among the factors which can trigger plastic responses, and be experimentally manipulated, temperature has been shown to affect both morphological and life-history traits under laboratory conditions in a variety of biological models (e.g. Rhen & Lang 1995, David *et al.* 1997, Atkin *et al.* 2006, Smith *et al.* 2013). Especially for ectotherm organisms such as insects, temperature is a relevant ecological parameter as it is highly influential on life-cycle and geographic distribution in the wild (Deutsch *et al.* 2008, Régnière *et al.* 2012). In *Drosophila* extensive laboratory work has shown that developmental temperature – the temperature experienced by the individual from egg to adult – has various and strong effects on *Drosophila* phenotypes including both morphological (e.g. thorax length, pigmentation, Barker & Krebs 1995, Gibert *et al.* 2000), and life history traits (e.g. fecundity, longevity, Nunney & Cheung 1997). Wing morphology in particular is well known to be plastic to temperature (e.g. David *et al.* 1994, Partridge *et al.* 1994, Imasheva *et al.* 2000; Debat *et al.* 2003, Gilchrist *et al.* 2004; Bublly & Loeschcke, 2005, David *et al.* 2006, Debat *et al.* 2009, Pitchers *et al.* 2013, Torquato *et al.* 2014, Przybylska *et al.* 2016). *Drosophila* wing shape has been shown to be a phylogenetically conserved trait but also highly evolvable in the laboratory (Weber 1992, Houle *et al.* 2003) suggesting strong stabilizing selection on this trait (Hansen & Houle 2004, 2008). However, geographic variation for wing morphology in wild *Drosophila* populations also suggests adaptive evolution to latitudinal and altitudinal (i.e. thus environmental) clines (e.g. Bitner-Mathe & Klaczko 1999, Huey *et al.* 2000, Gilchrist *et al.* 2001, 2004). Whether wing size and shape plasticity is adaptive or not is a contentious issue. Clinal divergence for thermal plasticity in wing morphology has been documented and this plastic clines are in the same direction as the genetic clines (e.g. Petavy *et al.* 1997; Morin *et al.* 1999; Pitchers *et al.* 2013). Wing function in flight should allow for specific test of the relationship between phenotypic plasticity, wing morphology and flight performance.

Flight performance in *Drosophila* has been studied under several experimental context (e.g. Marden *et al.* 1997, Card & Dickinson 2008) and is expected to be influenced by temperature (Lehman 1999, Dillon & Frazier 2006, Frazier *et al.* 2008). Specifically, several aerodynamic properties of the wing have been proposed to affect flight performance under different different thermal conditions (Gilchrist & Huey 2004). The wing aspect-ratio for instance (i.e. the ratio of wing length to width, Azevedo *et al.* 1998) is a common measure of flight performance in aerodynamic studies of wing structures (Dudley 2002) and has been described in several *Drosophila* species (e.g. Bublly & Loeschke 2005). From these studies, some general expectations can be formulated

regarding the association between wing morphology and flight. Ellington (1984) suggests for example that elongated wing with conserved wing area should yield greater aerodynamic forces and overall better flight performances. Azevedo *et al.* (1998) showed in *D. melanogaster* that higher wing aspect-ratio was advantageous for flight in low temperatures (to compensate for low wing-beat frequencies in the cold), a result that was later on attributed by Frazier *et al.* (2008) to adaptive developmental plasticity. The association between phenotypic plasticity and flight performance can thus be empirically tested in *Drosophila* species, and can allow testing for hypotheses regarding adaptive plasticity and biological invasions.

Here, we investigate the effect of developmental temperature on the wing morphology of the invasive pest *Drosophila suzukii* and test if changes in wing size and shape affects flight performance. The spotted-wing *Drosophila suzukii* is an Asian fruit fly species from the *melanogaster* group (Kopp & True 2002). The species is a particularly successful invader which has colonized more than 20 countries across Europe, North- and South-America (Hauser 2011, Calabria *et al.* 2010, Depra *et al.* 2014), from its native Asian range in less than a decade (see Asplen *et al.* 2015 for detailed review). Although rapid trans-continental migration is most probably human mediated (i.e. global trade of fruits, to which *D. suzukii* is associated; Lee *et al.* 2011), the species also shown rapid spread within invaded countries (Fraitout *et al.* 2016), and release/recapture experiments of marked individuals suggest that *D. suzukii* can actively disperse through flight (Anfora, pers. comm.).

Using geometric morphometrics and semi-automated tracking of flights obtained from high-speed video cameras, we respectively analyze wing shape and flight performance of *D. suzukii* individuals reared at different temperatures. We test the hypothesis that temperature-induced morphological variation in the wing will affect flight performance. Furthermore, by using populations derived from the native and invasive ranges, we test whether invasive populations exhibit higher phenotypic plasticity for wing morphology than native populations.

Material & Methods

Study populations and rearing

Drosophila suzukii adults were sampled using banana bait traps and net swiping at three localities representative of their global distribution range in Japan (native range), USA and France (invasive range). For each country, a northern population was sampled, to grant comparisons between thermally comparable conditions. The Japanese population was sampled in Sapporo, the US population was sampled in Dayton (Oregon) and the French population was sampled in Paris. Flies stocks were established by ten isofemale lines per locality by performing single matings in separated rearing vials and expanding F1 offspring in a subsequent series of replicate vials (Hoffmann & Parsons 1988, David *et al.* 2005). Stock populations were kept at $22 \pm 1^\circ\text{C}$ on rearing

medium consisting of corn-starch and yeast with added anti-biotic (hydroxy-4 benzoate). Prior to the experiment, flies from each basal population were transferred to oviposit in two batches of 20 new vials (i.e. 10 vials per batch) for 24 hours. For each population, after oviposition was ensured (i.e. presence of eggs in the medium), parent flies were removed and the two batches were separately placed in two incubators at $16 \pm 1^\circ\text{C}$ and $28 \pm 1^\circ\text{C}$ and one batch kept at stock temperature $22 \pm 1^\circ\text{C}$. Thus, each population was represented by 10 isofemale lines reared at three experimental temperatures resulting in 30 lines per geographic population. Lines were randomly positioned in incubators and kept at the experimental temperatures until two days after emergence to ensure sexing and optimal flight ability (Dillon & Frazier 2006). For each line from each population, 1 male *D. sukuzii* was then randomly chosen for flight analysis resulting in a total of 89 individual filmed and measured.

Flight acquisition and analysis

Flies were released at one side of a circular and transparent chamber (Fig. S1) covered on one side by white paper to maximize contrast and facilitate subsequent digitization. Each fly was released individually and allowed to fly freely in the chamber. Room temperature was not controlled but measured at $21 \pm 1^\circ\text{C}$. Unresponsive flies were stimulated by touching them briefly with a fine paint brush. We considered a successful flight to be at least 10 seconds without interruption, and three successful flights were recorded for each individual. Flies were then euthanized and preserved in 70% ethanol. Flights were recorded using three synchronized high-speed cameras (Prosilica GE680, Allied Vision Technologies GmbH, Stadroda, Germany) set at 200 frames per second. Two cameras were set up in dorsal view, one camera was set up in oblique frontal view. Thus, the fly was visible in at least two of the three views during the whole sequence recorded. Cameras were calibrated and scaled using a DLT routine (Hartley & Sturm, 1995) based on the digitization of a black-and-white checkerboard composed of ten one by one cm squares. The position of the fly was digitized and the screen-coordinates were calculated on the three synchronized views using a custom Matlab routine (Loco 2.8). Next, flight paths were smoothed with a low-pass quintic spline (Fig. S2) and the following flight parameters were calculated using a custom-written R script: sinuosity (defined as the ratio of the distance between the initial and final positions over the actual flight path length), peak flight speed, acceleration and deceleration, smallest and biggest angular turns and their associated peak angular speed, acceleration and deceleration. The highest values out of the three flights were retained for each parameter with the exception of the smallest turn angle for which we retained the minimal value. To ensure comparability, absolute values of the flight parameters were log-transformed prior analysis.

Geometric morphometrics

Wings were mounted on slides in a mixture of ethanol and glycerin. Cover slips were sealed using

nail-polish and maintained with small weights to flatten the wing as much as possible. Images of the wings were then acquired with a Leica DFC 420 digital camera mounted on a Leica Z6 APO microscope. Fifteen landmarks were defined on the dorsal face of the right wing (see Fig 1) and digitized using an ImageJ plugin written by Chris Klingenberg (personal communication). Procrustes generalized least squares superimposition was used to extract shape information from the landmark coordinates (GLS, e.g. Dryden & Mardia 1998). A PCA was first applied to the set of 30 coordinates and the non-null 26 components were conserved for the shape analyses. Centroid size was used as a size variable in the subsequent analyses and compared between all pairs of populations. We calculated the wing aspect-ratio (AR) as the ratio between wing length and wing width (Azevedo *et al.* 1998), defined as the distances computed between pairs of landmarks LM13 and LM3 (i.e. wing length) and between LM12 and LM14 (i.e. wing width), using the coordinates prior superimposition.

Statistical analyses

We first performed PCA and Linear Discriminant Analyses (LDA) on both shape and flight data sets to visually investigate variation, using temperature and population as grouping factors in the LDA. We then tested the effect of developmental temperature and of geographic origin on wing morphology and flight performance. For wing morphology, we performed uni- and multivariate analyses of variance (ANOVA and MANOVA) on wing size (i.e. centroid size) and wing shape (i.e. non-null PC scores) respectively, with population and temperature as main effects. The effect of temperature and population on flight parameters was tested by performing a MANOVA on the combined flight parameters (i.e. total flight performance) and subsequent ANOVAs for each variable independently. We then estimated the degree of covariation between flight performance and wing shape data using the Escoufier's RV coefficient (Escoufier 1973, Klingenberg 2009) and a two-blocks Partial Least Square regression (Rohlf and Corti 2000). To visualize the shape changes associates with particular components of flight, multivariate regressions of each of these parameters on shape data were applied (e.g. Monteiro 1999).

Results

Effect of temperature on wing morphology

As expected temperature had a strong effect on wing morphology (Fig. 2). A negative relationship between wing centroid size and temperature was observed for all populations (Fig. 2a) with flies raised at cooler temperature having larger wings. There was no difference in wing size among populations within-temperature. The amplitude of the size plastic response was higher for the US population compared to the Japanese population (i.e. an interaction between temperature and population was detected in the ANOVA, $p = 0.006$; Fig. 2a). There was a strong effect of temperature on wing aspect-ratio (ANOVA, $p < 0.001$; Fig. 2b), AR being positively correlated with

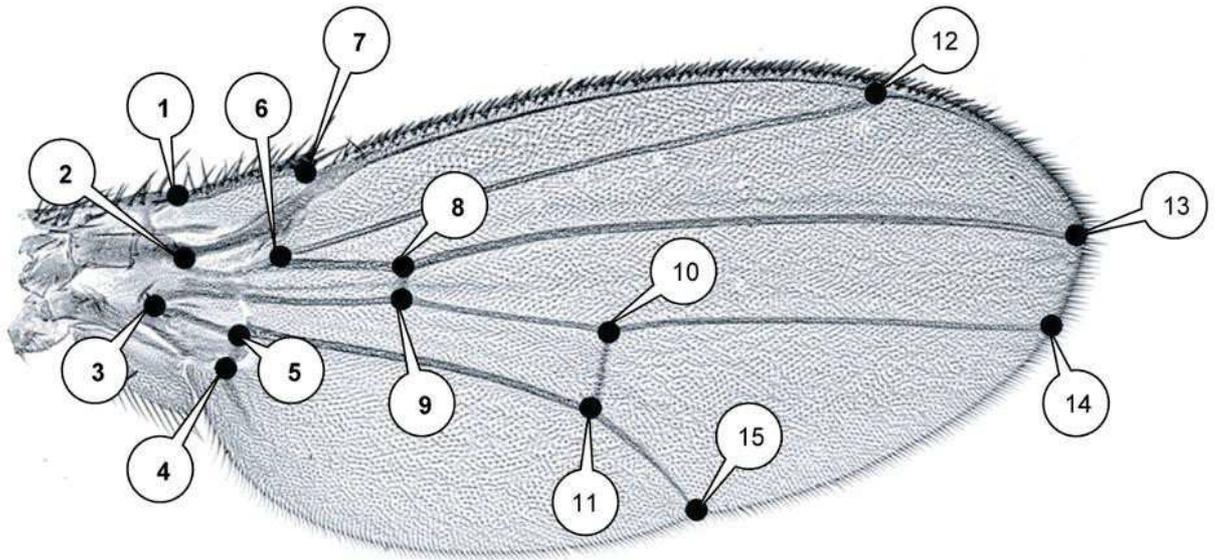


Figure 1. Positions of the fifteen landmarks used to quantify wing shape (from Debat *et al.* 2003).

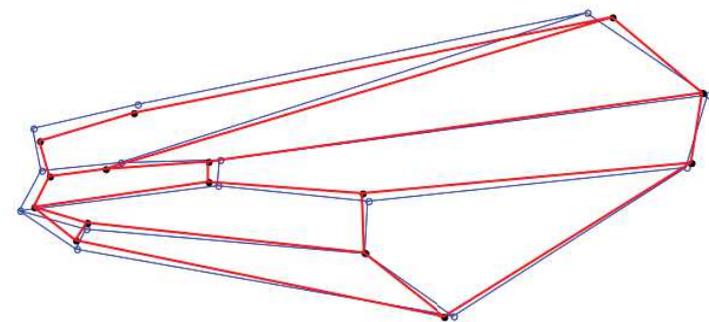
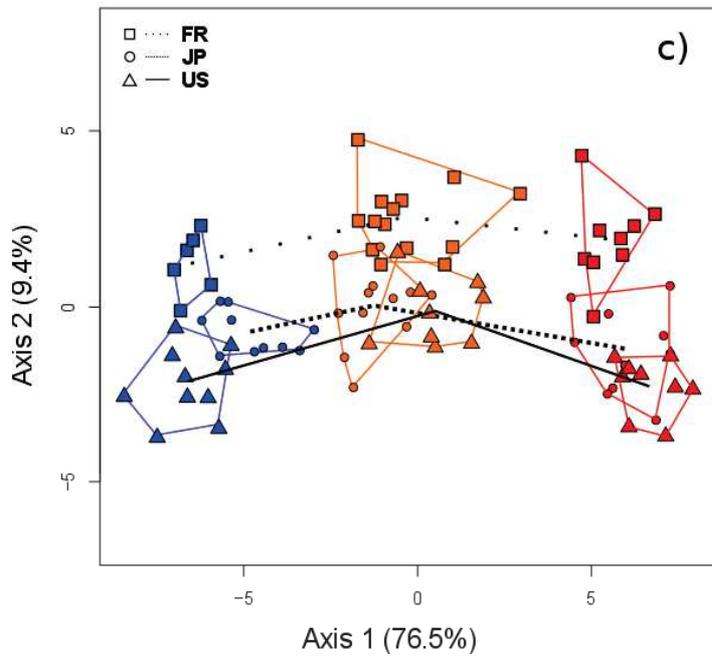
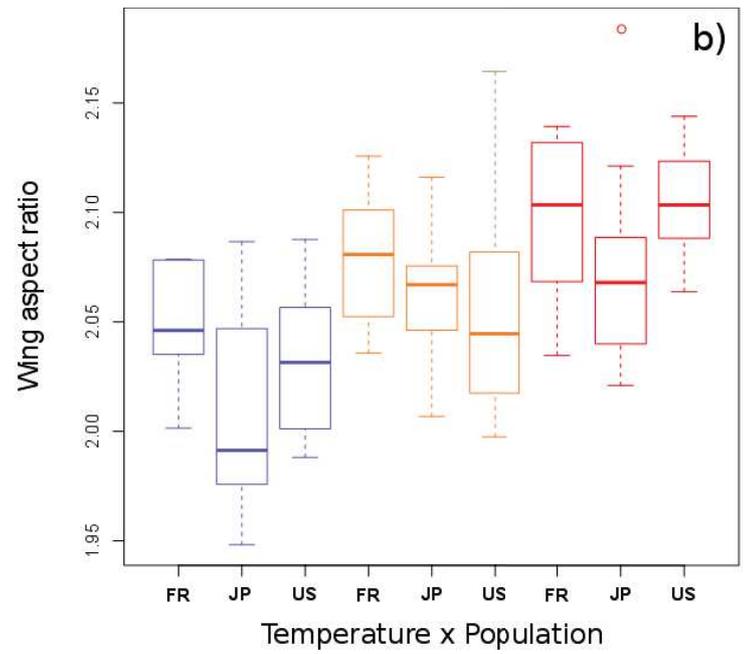
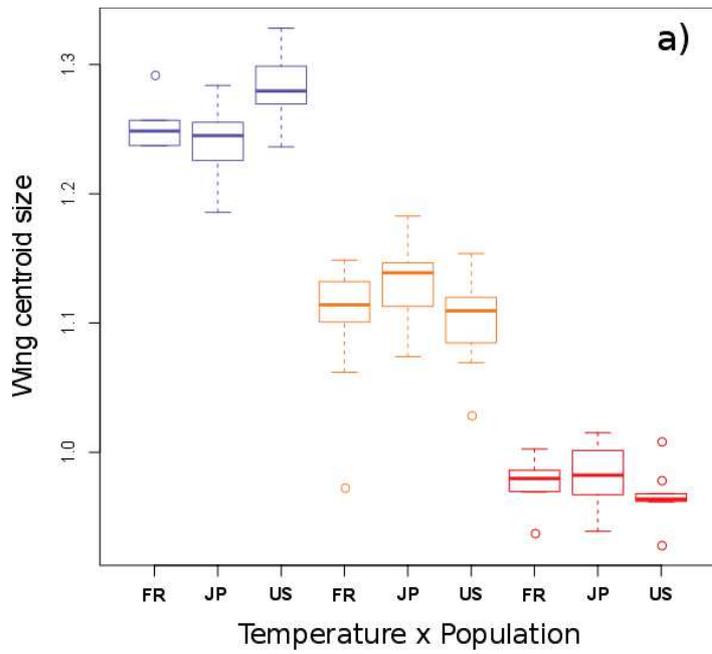


Figure 2. Effect of temperature on different aspects of *D. sukuzii* wing morphology.

The effect of developmental temperature on all three populations (FR: Montpellier, France; JP: Sapporo, Japan; US: Dayton, USA) on a) Wing centroid size; b) Wing aspect ratio and c) wing shape as represented by the Linear Discriminant Analysis (LDA). d) wing shape as represented by the first PC. Colours represent the three experimental temperatures: blue = 16°C, orange = 22°C, red = 28°C. Lines represent the shape reaction norms on the first discriminant plan.

temperature: relatively more elongated wings were produced at higher temperatures. There was no difference in the plastic response among populations for AR (i.e. no interaction between temperature and population effect).

The two main factors (population and temperature) had each an effect on wing shape (Fig. 2c, Table S1). The lack of interaction between temperature and population effects indicates that no difference in plasticity for wing shape between populations was detectable. Individuals from all three populations were thus pooled together for further shape/flight analyses. As wing shape can be influenced by size (i.e. allometry, e.g. Debat *et al.* 2003, Bolstad *et al.* 2015) we performed a MANOVA on the residuals from a regression between shape and size (i.e. the non-allometric component of shape). This procedure was valid as no allometry was detected within temperature, suggesting a lack of confounding effects of between- and within-temperature allometries. There was no effect of temperature on shape residuals after extraction of the allometric component (Table S2) indicating that the previous observed effect of temperature on wing shape was mostly allometric.

Effect of temperature on flight performance

We first investigated the correlations across flight parameters. Unsurprisingly, speed and acceleration/deceleration were strongly correlated. The same was found for angular parameters, but the two sets of parameters were only mildly correlated. Sinuosity and the smallest angular change were not correlated with any other parameter.

We investigated the interaction between temperature and population effects on flight. The standard MANOVA indicated an interaction between both effects ($p = 0.012$, Table S3). However, when performing a permutational MANOVA, the interaction between the two effects was lost ($p = 0.804$, Table S4). Considering this as an indication of a weak interaction at best, we decided to pool all individuals from all three populations for further analyses. Figure 3 shows the result from the PCA on the combined flight parameters. Although differentiation was rather weak, temperature had a significant effect on the total flight performance (MANOVA, $p = 0.013$ and permutational MANOVA, $p = 0.003$, Table S4). We then investigated which individual flight variable was affected by temperature. Temperature had a significant effect on velocity, acceleration, angular velocity and angular acceleration. More precisely, significant differences were observed for flies raised at 16°C compared to the two other experimental populations. Flies reared in colder temperature were capable of higher acceleration and faster flights (Fig. 4).

Association between shape and flight

Running multivariate regression of individual flight parameters on shape data, we found that velocity and acceleration had the strongest effect. The shape change associated with increased speed is shown on figure 4. Flies reared at the cooler temperature (i.e. fastest flies) had larger

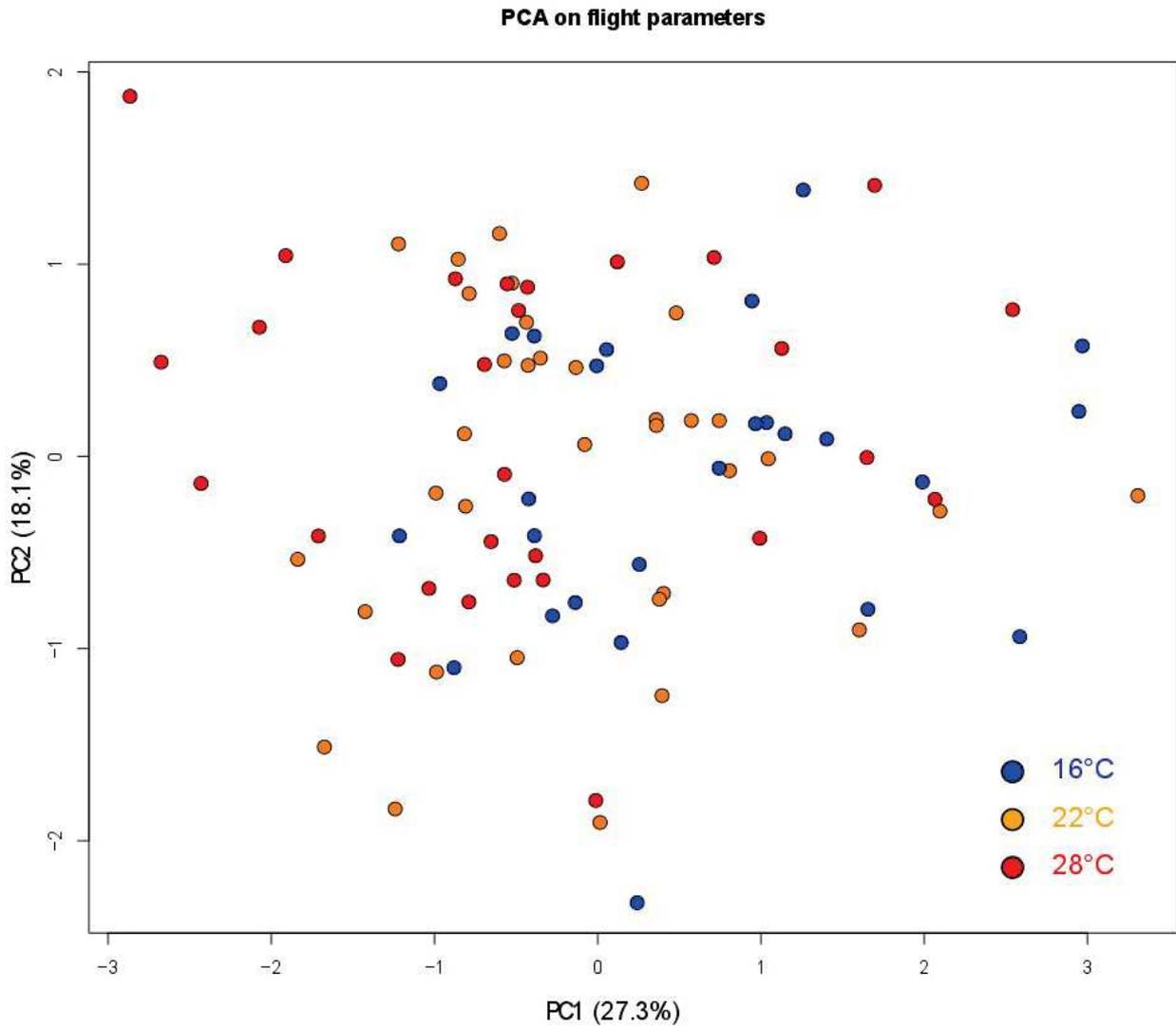


Figure 3. PCA on flight parameters and developmental temperature.

Filled circles represent individuals from all populations. Colors represent the three experimental temperatures: blue = 16°C, orange = 22°C, red = 28°C.

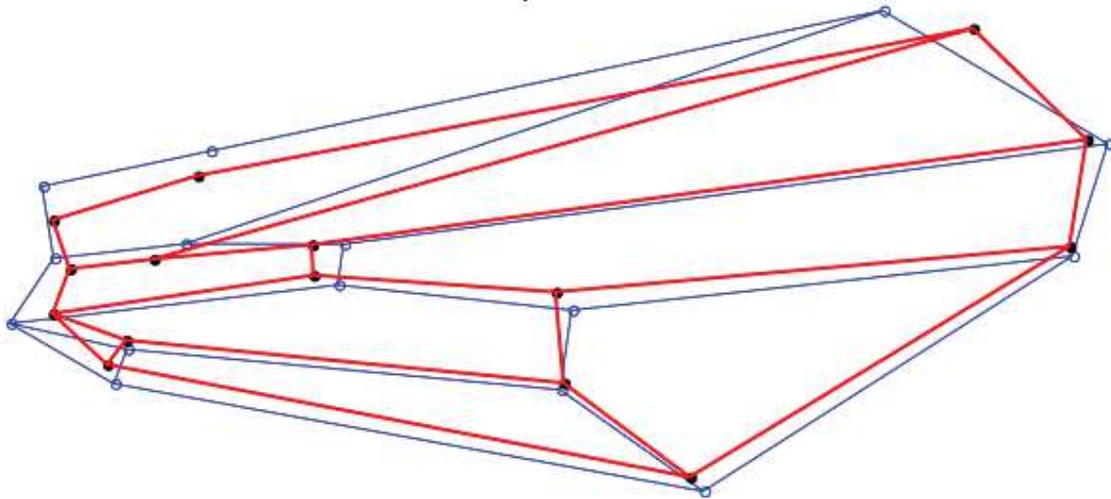
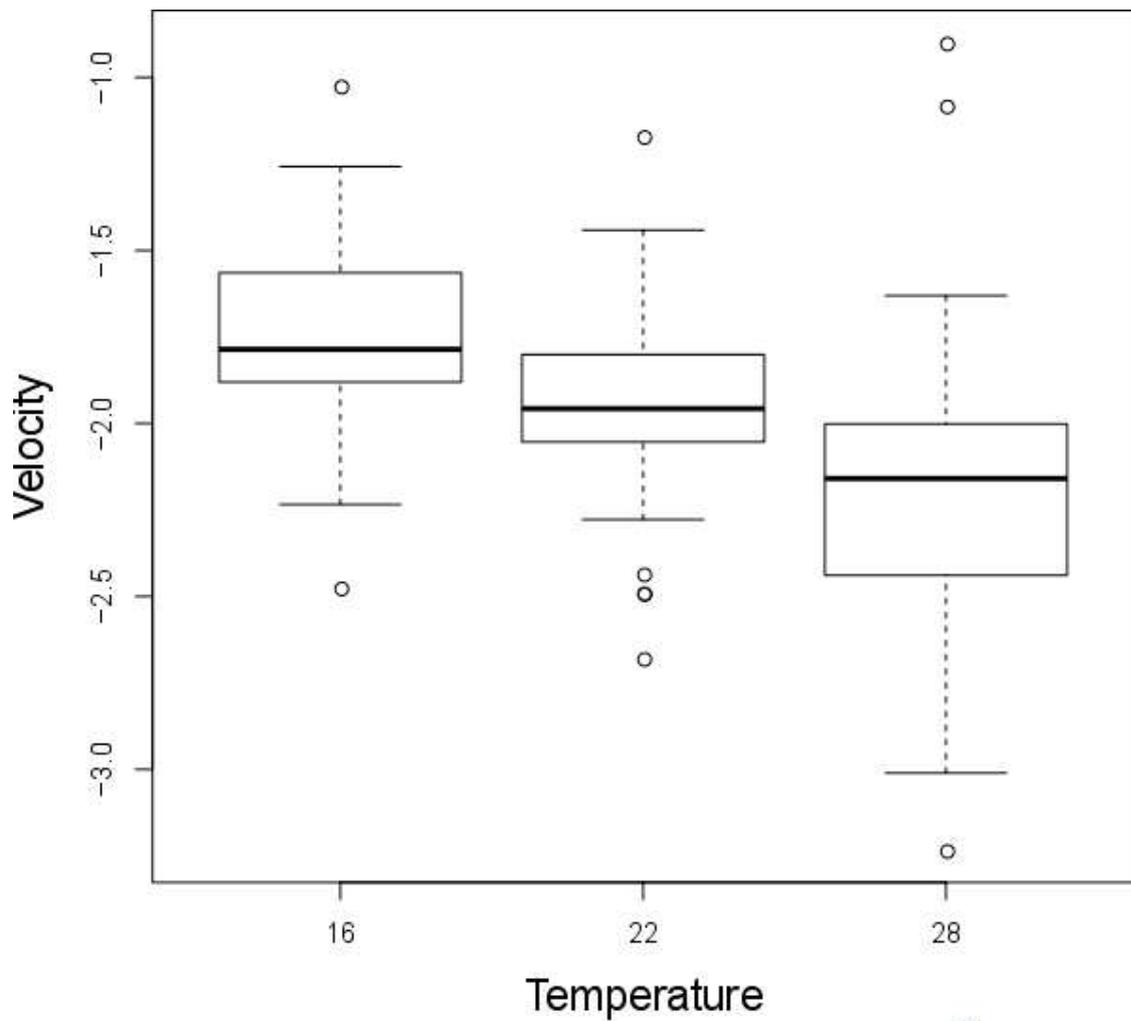


Figure 4. Temperature effect on flight velocity and associated shape change.

Boxplot represent the differences in velocity between the 3 temperatures. Under each boxplot is the associated wing shape as described by the first principal component of variation.

wings (i.e. in blue on Fig. 4). Shape change was less pronounced for acceleration (Fig. 4) We investigated the covariation of wing shape and flight. The RV test revealed a significant covariation between shape and flight ($p = 0.03$). The PLS analysis confirmed this result, showing that the covariation between shape and flight is driven by temperature (Fig. 5). The flight component of the PLS analysis (PLS block 2) is clearly driven by speed and acceleration. Unsurprisingly, the shape change associated with the first PLS axis is extremely similar to that found for the association with velocity, the fastest flies presenting the most rounded shape and being found at the lowest temperature.

Discussion

We studied the effect of developmental temperature on *D. sukukii* wing morphology and how this phenotypic plasticity may affect flight performance. As the focal trait is a crucial aspect of phenotypic plasticity experiments (Forsman 2015) we chose to focus on wing morphology, as this trait is highly plastic in *Drosophila* and involved in locomotory function (David *et al.* 1997).

As expected we found a strong effect of temperature on wing size and shape (Fig. 1). We found no overall difference in wing shape and flight plasticity between native and invasive populations of *D. sukukii*. This result does not support the hypothesis of a role for wing plasticity to temperature in the success of the invasion. Post-invasion evolution of increased plasticity (Bossdorf *et al.* 2005) is indeed a hypothesis commonly proposed to explain higher levels of phenotypic plasticity in introduced populations. In that case, only the individuals with the highest plastic response from the propagule pool will colonize the new habitat. As these individuals are continuously selected, the mean reaction norm should become steeper in the invasive range compared to the native populations. *D. sukukii* populations used in this study were derived from similar latitudes within their respective countries (i.e. northern populations), with annual average temperatures in all three localities being relatively similar (from meteorological databases, results not shown). Relative similarity between ranges could explain the absence of plasticity differences among populations.

Our results suggest that developmental temperature affects flight performance. More specifically, flies reared at the colder experimental temperature (i.e. 16°C) showed higher flight velocity and acceleration. Increased flight performance in cold conditions have been previously reported for *Drosophila melanogaster* (Frazier *et al.* 2008). In that case for example, longer wings (a morphological parameter expected to generate higher aerodynamic forces; Ellington 1984 in Frazier *et al.* 2008) in cold-reared flies were associated to increase lift performance in cold air. However, results from Frazier *et al.* (2008) were obtained from controlled flight temperature and might not be extrapolated to our results, as our flies – cold-, medium- of warm-reared – all flew at an intermediate temperature.

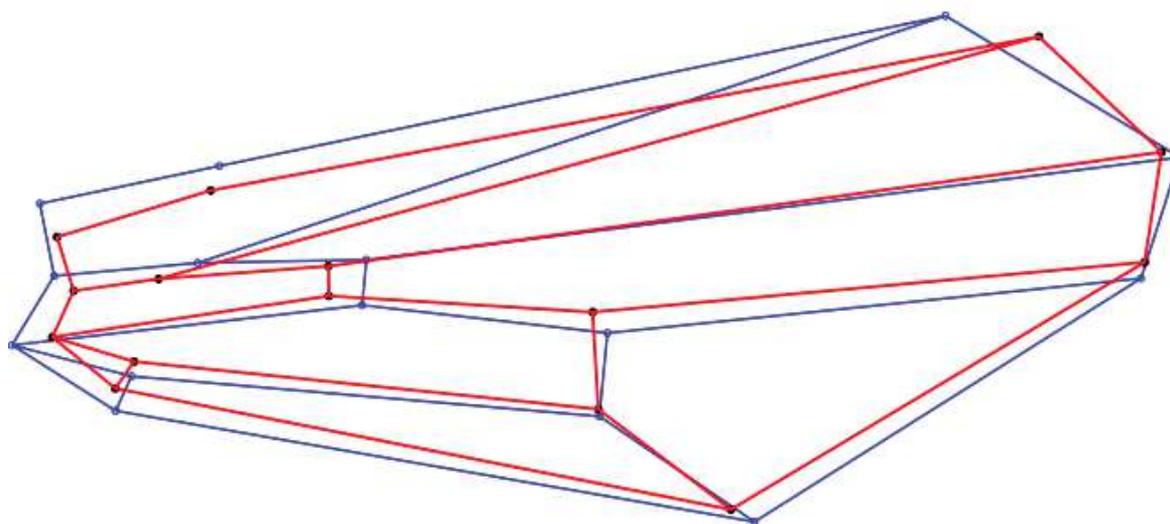
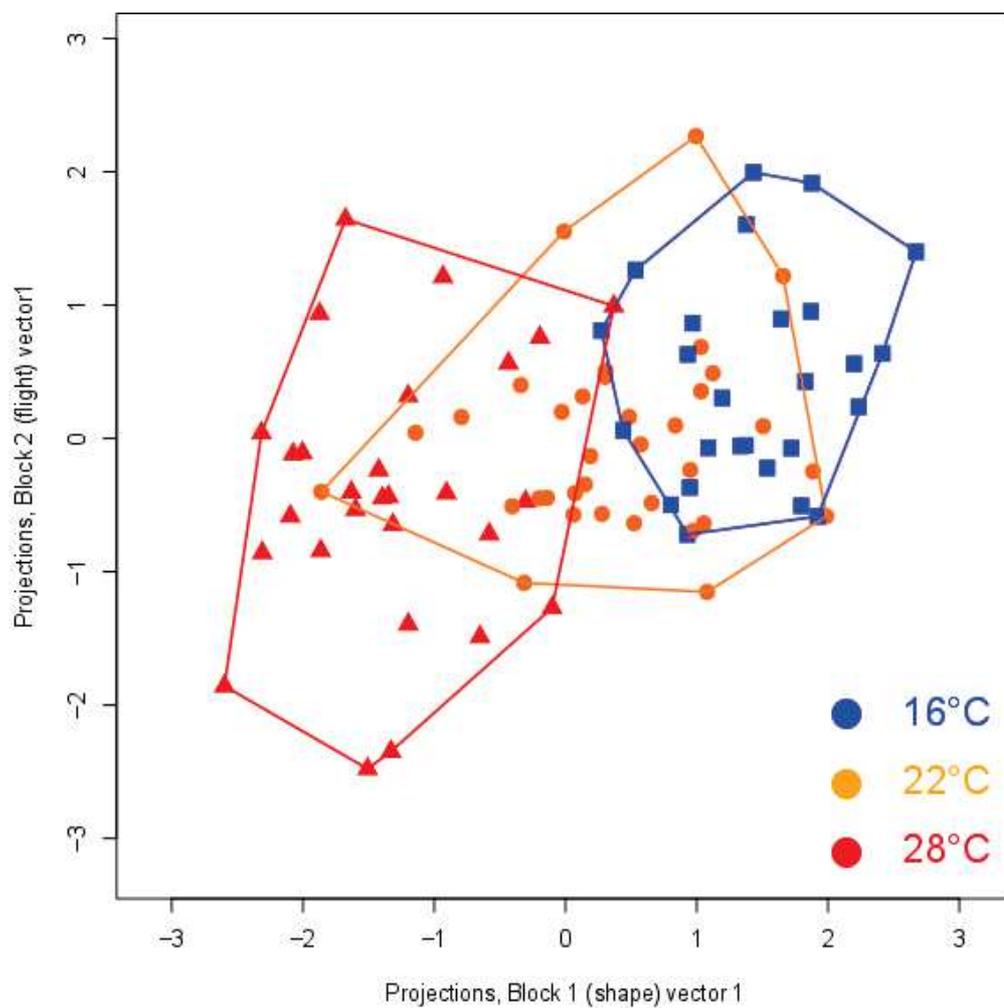


Figure 5. Plots of the PLS analysis and associated shape change on the PLS block 1. Colors represent experimental temperatures: blue = 16°C, orange = 22°C and red = 28°C.

Additionally, the shape change we detected is strikingly different from the one they describe. The positive relationship we found between the wing-aspect ratio and developmental temperature is indeed reversed from that reported in most previous studies of *D. melanogaster* (e.g. Azevedo *et al* 1998; Przybylska *et al.* 2016; but see Debat *et al* 2009). Wing-aspect ratio (AR) is used to describe wing planform and is related to aerodynamic efficiency (Dudley 2002). Higher AR, that is, longer and narrower wings, improves gliding flight whereas smaller AR are associated with increased flight maneuverability (e.g. Dudley, 2002; Chazot *et al* 2016). In *Drosophila melanogaster*, higher AR has been proposed to increase flight performance at cold temperature, where elongated wings might increase the mechanical efficiency of wing beat, compensating the often observed reduction in wing-beat frequency (Azevedo *et al.* 1998; Frazier *et al* 2008). The association of low AR and increased linear and angular speed and acceleration we detected in *D. suzukii* at colder temperature is thus unexpected. It is however consistent with the increased maneuverability predicted for rounder and shorter wings (e.g. Dudley 2002). Positive relationship between AR and temperature has nevertheless been reported in other *Drosophila* species (e.g. *D. buzzatii*: Bublly & Loeschke 2005), others cases are reviewed in Przybylska *et al.* (2016). Whether species ecology or the thermal conditions typically encountered might influence wing shape plasticity and flight performance is unknown. This observation would deserve more experimental analysis, especially in conditions of controlled air temperature during flights.

Another question deserves additional investigation: whether the effect of temperature on flight is actually induced by the plasticity of wing size and shape is difficult to assess. Particularly, we do not have any informations regarding neuromuscular activity variation potentially associated with developmental temperature, a parameter that is crucial for insect flight (Ellington 1984, Dudley 2002). It is thus conceivable that the detected flight difference might result from other sources than differences in wing morphology.

Finally, one could argue that flight performance could be better assessed through orientated flight behavioral experiment involving food acquisition or mating, rather than by quantifying free flight. However, preliminary tests of flights in the presence of fresh fruits in the flight chamber did not reveal any particular behavioral response (results not shown).

Adaptive plasticity

An explicit test for the adaptive nature of wing shape plasticity was beyond the reach of our experimental setup. Indeed, theory predicts that plasticity would be adaptive if the plastic trait is advantageous in the environment by which it is induced (or at least correlated with it) (Yeh & Price 2004). Flies reared in the cold should thus fly more efficiently in cold air than flies reared in warm conditions, and vice versa. Here, we show that lower developmental temperature increase flight speed and acceleration in intermediate air temperature. However our results do not disqualify the adaptive hypothesis. We indeed detected a covariation between wing shape and flight

performance. Under a null hypothesis of non-adaptive plasticity, no effects of shape variation on flight performance would be expected. Our results therefore encourage further investigation on the adaptive value of wing shape plasticity in relation to flight performance in *D. sukuzii*, with additional care on the flight environment.

Finally, it is worth trying to put our experiments back in an ecological context. *D. sukuzii* is considered to be a rather temperate species limited in its physiology by high temperatures and drought level (Ometto *et al.* 2013, Tochen *et al.* 2014). Here our lowest temperature of $16 \pm 1^\circ\text{C}$ might not fall in the colder end of temperature tolerance for *D. sukuzii*, especially for populations derived from northern latitude. For instance Dalton *et al.* (2011) investigated the seasonality of *D. sukuzii* in Marion, Oregon (ca. 60km from the sampling point of our US population). Their results indicate that *D. sukuzii* activity, as described by trap counts, was higher in late September when the daily mean temperature was between 15 and 20°C . Thus, rather than testing cold vs warm treatments around an optimal physiological temperature of 22°C , we might have actually exposed cold adapted *D. sukuzii* to a range of somewhat unusually high temperatures. This could explain why our populations – all sampled from northern latitudes – showed increased performance at 16°C compared to warmer temperatures.

This result would be particularly interesting as locomotor performance is usually inversely proportional to temperature in *Drosophila* (Lehmann 1999, Gibert *et al.* 2001, Dillon & Frazier 2006, but see Frazier *et al.* 2008). Additional work regarding the effect of temperature on flight should investigate the performance of populations derived from the southern range of *D. sukuzii* in USA, Europe and Japan.

Acknowledgments

AF was supported by state funding by the Agence Nationale de la Recherche, through the LabEx ANR-10-LABX-0003-BCDiv, of the program “Investissements d’avenir” (ANR-11-IDEX-0004-02). The authors acknowledge K. Tamura, M. Toda, P. Shearer for help in flies sampling. We thank Marion Guillaume for helping with the stock flies maintenance and Frédérique Peronnet for providing flies rearing medium.

References

- Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., *et al.* (2015). Invasion biology of spotted wing *Drosophila* (*Drosophila sukuzii*): a global perspective and future priorities. *Journal of Pest Science* 88, 469–494.
- Atkin, O.K. (2005). Phenotypic plasticity and growth temperature: understanding interspecific variability. *Journal of Experimental Botany* 57, 267–281.

- Azevedo, R.B.R., James, A.C., McCabe, J., and Partridge, L. (1998). Latitudinal Variation of Wing: Thorax Size Ratio and Wing-Aspect Ratio in *Drosophila melanogaster*. *Evolution* 52, 1353.
- Beldade, P., Mateus, A.R.A., and Keller, R.A. (2011). Evolution and molecular mechanisms of adaptive developmental plasticity: ADAPTIVE DEVELOPMENTAL PLASTICITY. *Molecular Ecology* 20, 1347–1363.
- Bitner-Mathé, B.C., and Klaczko, L.B. (1999). Plasticity of *Drosophila melanogaster* wing morphology: effects of sex, temperature and density. *Genetica* 105, 203–210.
- Bock, D.G., Caseys, C., Cousens, R.D., Hahn, M.A., Heredia, S.M., Hübner, S., Turner, K.G., Whitney, K.D., and Rieseberg, L.H. (2015). What we still don't know about invasion genetics. *Molecular Ecology* 24, 2277–2297.
- Bolstad, G.H., Cassara, J.A., Márquez, E., Hansen, T.F., van der Linde, K., Houle, D., and Pélabon, C. (2015). Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences* 112, 13284–13289.
- Bossdorf, O., Auge, H., Lafuma, L., Rogers, W.E., Siemann, E., and Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144, 1–11.
- Bubliy, O.A., and Loeschcke, V. (2005). Variation of life-history and morphometrical traits in *Drosophila buzzatii* and *Drosophila simulans* collected along an altitudinal gradient from a Canary island. *Biological Journal of the Linnean Society* 84, 119–136.
- Calabria, G., Máca, J., Bächli, G., Serra, L., and Pascual, M. (2012). First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe: First record of *Drosophila suzukii* in Europe. *Journal of Applied Entomology* 136, 139–147.
- Card, G., and Dickinson, M. (2008). Performance trade-offs in the flight initiation of *Drosophila*. *Journal of Experimental Biology* 211, 341–353.
- Chazot, N., Panara, S., Zilbermann, N., Blandin, P., Le Poul, Y., Cornette, R., Elias, M., and Debat, V. (2016). Morpho morphometrics: Shared ancestry and selection drive the evolution of wing size and shape in *Morpho* butterflies. *Evolution* 70, 181–194.
- Dalton, D.T., Walton, V.M., Shearer, P.W., Walsh, D.B., Caprile, J., and Isaacs, R. (2011). Laboratory survival of *Drosophila suzukii* under simulated winter conditions of the Pacific Northwest and seasonal field trapping in five primary regions of small and stone fruit production in the United States. *Pest Management Science* 67, 1368–1374.
- David, J.R., Moreteau, B., Gauthier, J.P., Petavy, G., Stockel, A., and Imasheva, A.G. (1994). Reaction norms of size characters in relation to growth temperature in *Drosophila melanogaster*: an isofemale lines analysis. *Genetics Selection Evolution* 26, 1.
- David, J. R., Gibert, P., Gravot, E., Petavy, G., Morin, J. P., Karan, D., & Moreteau, B. (1997). Phenotypic plasticity and developmental temperature in *Drosophila*: analysis and

- significance of reaction norms of morphometrical traits. *Journal of Thermal Biology*, 22(6), 441–451.
- David, J.R., Gibert, P., Legout, H., Pétavy, G., Capy, P., and Moreteau, B. (2005). Isofemale lines in *Drosophila*: an empirical approach to quantitative trait analysis in natural populations. *Heredity* 94, 3–12.
- Davidson, A.M., Jennions, M., and Nicotra, A.B. (2011a). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis: Invasive species have higher phenotypic plasticity. *Ecology Letters* 14, 419–431.
- Davidson, A.M., Jennions, M., and Nicotra, A.B. (2011b). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis: Invasive species have higher phenotypic plasticity. *Ecology Letters* 14, 419–431.
- Debat, V., Béugin, M., Legout, H., and David, J.R. (2003). Allometric and nonallometric components of *Drosophila* wing shape respond differently to developmental temperature. *Evolution* 57, 2773–2784.
- Debat, V., Debelle, A., and Dworkin, I. (2009). Plasticity, canalization and developmental stability of the wing: joint effects of mutations and developmental temperature. *Evolution* 63, 2864–2876.
- Deprá, M., Poppe, J.L., Schmitz, H.J., De Toni, D.C., and Valente, V.L.S. (2014). The first records of the invasive pest *Drosophila suzukii* in the South American continent. *Journal of Pest Science* 87, 379–383.
- Deutsch, C.A., Tewksbury, J.J., Huey, R.B., Sheldon, K.S., Ghalambor, C.K., Haak, D.C., and Martin, P.R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences* 105, 6668–6672.
- Dillon, M.E. Frazier, M.R. (2006). *Drosophila melanogaster* locomotion in cold thin air. *Journal of Experimental Biology* 209, 364–371.
- Ellington, C.P. (1984). The aerodynamics of hovering insect flight. I. The quasi-steady analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 305, 1–15.
- Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., and Facon, B. (2016). Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics* 47.
- Forsman, A. (2015). Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity* 115, 276–284.
- Frazier, M.R., Harrison, J.F., Kirkton, S.D., and Roberts, S.P. (2008). Cold rearing improves cold-flight performance in *Drosophila* via changes in wing morphology. *Journal of Experimental Biology* 211, 2116–2122.
- Friedrich, S.R., and Meyer, A. (2016). How plasticity, genetic assimilation and cryptic genetic

- variation may contribute to adaptive radiations. *Molecular Ecology*.
- Funk, J.L. (2008). Differences in plasticity between invasive and native plants from a low resource environment. *Journal of Ecology* 96, 1162–1173.
- Geng, Y.-P., Pan, X.-Y., Xu, C.-Y., Zhang, W.-J., Li, B., Chen, J.-K., Lu, B.-R., and Song, Z.-P. (2007). Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biological Invasions* 9, 245–256.
- Ghalambor, C.K., McKAY, J.K., Carroll, S.P., and Reznick, D.N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21, 394–407.
- Gibert, P., Moreteau, B., & David, J. R. (2000). Developmental constraints on an adaptive plasticity: reaction norms of pigmentation in adult segments of *Drosophila melanogaster*. *Evolution & development*, 2(5), 249-260.
- Gibert, P., Huey, R.B., and Gilchrist, G.W. (2001). Locomotor performance of *Drosophila melanogaster*: interactions among developmental and adult temperatures, age, and geography. *Evolution* 55, 205–209.
- Gilchrist, A.S., Azevedo, R.B.R., Partridge, L., and O'higgins, P. (2000). Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. *Evolution & Development* 2, 114–124.
- Gilchrist, G.W., Huey, R.B., and Serra, L. (2001). Rapid evolution of wing size clines in *Drosophila subobscura*. *Genetica* 112, 273–286.
- Gilchrist, G.W., Huey, R.B., Balanyà, J., Pascual, M., Serra, L., and Noor, M. (2004). A time series of evolution in action: a latitudinal cline in wing size in South American *Drosophila subobscura*. *Evolution* 58, 768–780.
- Hansen, T.F., and Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* 21, 1201–1219.
- Hansen, T.F., and Houle, D. (2004). Evolvability, stabilizing selection, and the problem of stasis. In Pigliucci & Preston. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford University Press (2004)
- Hauser, M. (2011). A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Management Science* 67, 1352–1357.
- Hoffmann, A.A., and Parsons, P.A. (1988). The analysis of quantitative variation in natural populations with isofemale strains. *Genetique, Selection, Evolution* 20, 1.
- Houle, D., Mezey, J., Galpern, P., and Carter, A. (2003). Automated measurement of *Drosophila* wings. *BMC Evolutionary Biology* 3, 1.
- Huey, R.B., Gilchrist, G.W., Carlson, M.L., Berrigan, D., and Serra, L. (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287, 308–309.

- Imasheva, A.G., Moreteau, B., and David, J.R. (2000). Growth temperature and genetic variability of wing dimensions in *Drosophila*: opposite trends in two sibling species. *Genetical Research* 76, 237–247.
- Kaufman, S.R., and Smouse, P.E. (2001). Comparing indigenous and introduced populations of *Melaleuca quinquenervia* (Cav.) Blake: response of seedlings to water and pH levels. *Oecologia* 127, 487–494.
- Klingenberg, C.P. (2009). Morphometric integration and modularity in configurations of landmarks: tools for evaluating a priori hypotheses. *Evolution & Development* 11, 405–421.
- Lavergne, S., and Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences* 104, 3883–3888.
- Lee, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17, 386–391.
- Lee, J.C., Bruck, D.J., Curry, H., Edwards, D., Haviland, D.R., Van Steenwyk, R.A., and Yorgey, B.M. (2011). The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest Management Science* 67, 1358–1367.
- Lehmann, F. O. (1999). Ambient temperature affects free-flight performance in the fruit fly *Drosophila melanogaster*. *Journal of Comparative Physiology B*, 169(3), 165-171.
- Levis, N.A., and Pfennig, D.W. (2016). Evaluating “Plasticity-First” Evolution in Nature: Key Criteria and Empirical Approaches. *Trends in Ecology & Evolution* 31, 563–574.
- Marden, J.H., Wolf, M.R., and Weber, K.E. (1997). Aerial performance of *Drosophila melanogaster* from populations selected for upwind flight ability. *Journal of Experimental Biology* 200, 2747–2755.
- Morin, Moreteau, Pétavy, and David (1999). Divergence of reaction norms of size characters between tropical and temperate populations of *Drosophila melanogaster* and *D. simulans*. *Journal of Evolutionary Biology* 12, 329–339.
- Muth, N.Z., and Pigliucci, M. (2006). Traits of invasives reconsidered: phenotypic comparisons of introduced invasive and introduced noninvasive plant species within two closely related clades. *American Journal of Botany* 93, 188–196.
- Nunney, L., and Cheung, W. (1997). The Effect of Temperature on Body Size and Fecundity in Female *Drosophila melanogaster*: Evidence for Adaptive Plasticity. *Evolution* 51, 1529.
- Ometto, L., Cestaro, A., Ramasamy, S., Grassi, A., Revadi, S., Siozios, S., Moretto, M., Fontana, P., Varotto, C., Pisani, D., *et al.* (2013). Linking Genomics and Ecology to Investigate the Complex Evolution of an Invasive *Drosophila* Pest. *Genome Biology and Evolution* 5, 745–757.
- Partridge, L., Barrie, B., Fowler, K., and French, V. (1994). Evolution and Development of Body Size and Cell Size in *Drosophila melanogaster* in Response to Temperature. *Evolution* 48,

1269.

- Pigliucci, M. (2005). Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* 20, 481–486.
- Pitchers, W.R., Brooks, R., Jennions, M.D., Tregenza, T., Dworkin, I., and Hunt, J. (2013). Limited plasticity in the phenotypic variance-covariance matrix for male advertisement calls in the black field cricket, *Teleogryllus commodus*. *Journal of Evolutionary Biology* 26, 1060–1078.
- Price, T.D., Qvarnström, A., and Irwin, D.E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 1433–1440.
- Przybylska, M.S., Brito, F.A. de, and Tidon, R. (2016). Ecological insights from assessments of phenotypic plasticity in a Neotropical species of *Drosophila*. *Journal of Thermal Biology* 62, 7–14.
- Régnière, J., Powell, J., Bentz, B., and Nealis, V. (2012). Effects of temperature on development, survival and reproduction of insects: Experimental design, data analysis and modeling. *Journal of Insect Physiology* 58, 634–647.
- Rhen, T., and Lang, J.W. (1995). Phenotypic plasticity for growth in the common snapping turtle: effects of incubation temperature, clutch, and their interaction. *American Naturalist* 726–747.
- Schlichting, C.D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics* 667–693.
- Smith, S., Bernatchez, L., and Beheregaray, L.B. (2013). RNA-seq analysis reveals extensive transcriptional plasticity to temperature stress in a freshwater fish species. *BMC Genomics* 14, 1.
- Tochen, S., Dalton, D.T., Wiman, N., Hamm, C., Shearer, P.W., and Walton, V.M. (2014). Temperature-Related Development and Population Parameters for *Drosophila suzukii* (Diptera: Drosophilidae) on Cherry and Blueberry. *Environmental Entomology* 43, 501–510.
- Torquato, L.S., Mattos, D., Matta, B.P., and Bitner-Mathé, B.C. (2014). Cellular basis of morphological variation and temperature-related plasticity in *Drosophila melanogaster* strains with divergent wing shapes. *Genetica* 142, 495–505.
- Weber, K. E. (1992). How small are the smallest selectable domains of form?. *Genetics*, 130(2), 345-353.
- West-Eberhard, M.J. (1989). Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology and Systematics* 249–278.
- Yeh, P.J., and Price, T.D. (2004). Adaptive Phenotypic Plasticity and the Successful Colonization of a Novel Environment. *The American Naturalist* 164, 531–542.

Chapitre 4 : Modélisation de la niche climatique de *Drosophila suzukii*

Article 5.

Frainout, A., Monnet, A.C. Global invasion of the spotted-wing *Drosophila suzukii*: new insights from climatic niche modeling. *in prep.*

Global invasion of the spotted-wing *Drosophila suzukii*: new insights from climatic niche modeling

Antoine Fraimout & Anne-Christine Monnet

Abstract

Biological invasions are one of the main threats to the loss of biodiversity worldwide and understanding the mechanisms allowing exogenous species to colonize their new environment is therefore critical to the management of invasive populations. One particular aspect of invasion biology is to define the environmental conditions within which invasive species can thrive (i.e. their ecological niche) to get insights on the potential role of adaptation in successful invasions, as well as to predict future invasion. Here, we use multi-correlative species distribution modeling to investigate the worldwide invasion of the spotted-wing drosophila, *Drosophila suzukii*. By modeling the climatic envelope of *D. suzukii* from occurrence data in its native and invasive range, we tested if a shift in the realized niche has occurred during invasion. Furthermore, we used recent populations genetics work on the invasion history of the species to test whether invasive populations have preferentially invaded areas with more similar climatic conditions than in their precise area of origins. Although accounting for the invasion history when calibrating models did not increased presence prediction in the invasive range, our model successfully predicted the distribution of *D. suzukii* in its native and invasive ranges. Overall our result suggest that *D. suzukii* has similar ecological requirements across its native and invasive range with the exception of western region of USA where a shift in the realized niche seems to have occurred. Finally we provide the first global niche modeling approach for the spread of *D. suzukii*, and point towards several yet uncolonized countries which might be at risk of infestation.

Key words: ecological niche modeling, exotic species, invasion routes, niche conservatism, niche overlap

Introduction

Coping with environmental variations can be challenging to many organisms, and the extent to which species can respond to changes of their ecological niche highly depends on the magnitude of departure from the physiological optimum. Invasive alien species represent one of the most common studied cases of adaptation to new environmental conditions as they provide researchers with natural experiments of response to a changing environment (Yoshida *et al.* 2007). In the light of evolutionary theory, invasive species are predicted to face fast adaptive genetic changes in response to the new selective pressures emerging from the mismatch between their native environment and the new colonized niche (Lee 2002; Sakai *et al.* 2001). However consensual underlying evolutionary mechanisms explaining the success of invasive species remain largely unclear and disentangling the effects of genetics from purely demographic processes is challenging (Szűcs *et al.* 2014). From a demographic point of view, the null hypothesis of a successful invasion can be formulated as follows: successful invaders will thrive in an environment with similar climatic conditions as their native area, from which they were historically out of reach. This pre-adaptation hypothesis suggests that a change in dispersal ability of a species could solely explain its spread in a newly available area if the environmental conditions are favorable for source populations (Facon *et al.* 2006). Species distribution models (SDMs) as a tool to estimate the climatic envelope of invasive population in their native range (i.e., their realized niche; Hutchinson 1957) is particularly relevant for predicting suitable habitats for invaders, and the contribution of niche theory to the study of biological invasions has proven to be of great use when assessing the amplitude of environmental changes undergone by invasive populations (Tingley *et al.* 2014; Broennimann *et al.* 2007). However, if the native distribution of a species is large, SDMs will be more prone to over-estimate the potential realized niche in uncolonized areas if only a limited proportion of the native distribution is serving as a source for the invasion (Godefroid *et al.* 2016). Furthermore, if multiple sources are present in the native distribution, intra-specific niche differentiation could be overlooked and therefore bias niche projection on uncolonized areas (Lecocq *et al.* 2015; Godefroid *et al.* 2015).

Such an invasion scenario can be found in the worldwide spread of the spotted-wing *Drosophila suzukii*. This drosophilid species native from south-east Asia has started to colonize Europe and the United States in 2008 (Hauser 2011; Calabria *et al.* 2012) and is now distributed widely across the globe with the latest occurrences in the Southern hemisphere in Brazil (Depra *et al.* 2014; Paula *et al.* 2014). Additionally to its invader

status, *D. suzukii* is a major pest for stone fruits and berries crops causing considerable damage for the fruit production economy in the infested cultures (Lee *et al.* 2011). In a recent paper Fraimout *et al.* (2016) applied population genetics methods to retrace the invasion routes of the species at a global scale. In particular the continental part of Asia was pointed as origins for the two main introduction events in Europe and the USA, and some migration from USA to Europe was also observed.

Here we use multi-correlative SDM and the global spread of *D. suzukii* as a model to address specific questions: 1) What is the realized niche of *D. suzukii* in its native range? 2) Can we observe shifts in realized niche between ancestral and some regions of the invasive ranges? 3) Do invasive populations show more similarity in their realized niche than with their most probable geographical sources? Given the efficiency of the species to have colonized such a broad range of environments is a relatively short time (i.e. less than ten years, Asplen *et al.* 2015) we expect that niche similarity should in fact be of greater importance than local adaptations having favored the establishment of *D. suzukii* populations in areas with less suitable climate. We test the hypothesis that *D. suzukii* might in fact be a generalist species, having colonized environments which were historically out of reach.

Material & Methods

Occurrence data

We gathered a total of 344 presence data of *D. suzukii* from its native range in south-eastern Asia and in its invasive range in continental USA, Hawaii, Europe and the South hemisphere distributions (Brazil and the French island of La Réunion) from published occurrence records, databases, and personal observations from the field. For continental USA occurrences, we specifically used the EDDMapS county checklist (2016, www.eddmaps.org) providing a thorough record of the species, and randomly sampled 150 locations within counties where *D. suzukii* have been recorded.

Environmental data and representation in the climatic space

The selection of the environmental predictors for niche modeling is a source of uncertainty in model predictions that is still poorly studied but that can be reduced with sound statistical methods and ecological knowledge of species (Austin & Van Niel 2011, Barbet-Massin & Jetz, 2014). A low number of predictive variables prevents the issues of model

over-fitting (i.e., when predictors are exceeding the number of occurrences) and of multicollinearity (i.e., when the predictive variables are highly correlated, Heikkinen *et al.*, 2006). Therefore, we focused on proximal variables which constraint directly species physiology and resources (Broennimann *et al.*, 2006) and used eight climatic variables to encompass the mean trends, the variability, and the extremes of the current and future climate (see section *Future projections* below) at a given location: (1) mean annual temperatures, (2) mean annual precipitation, (3) temperature seasonality, (4) precipitation seasonality, (5) temperature of the coldest month and (6) of the warmest month, (7) precipitation of the driest month and (8) of the wettest month. In a recent paper, Petitpierre *et al.* (2016) showed that these eight climatic variables were one of the two best sets of predictors to yield the highest transferability of SDMs at macroclimatic scale when they projected the climatic niche of invasive plants in different environments. We extracted these climatic variables from the Worldclim Global Climate data on the current climate (1961-1990; Hijmans *et al.*, 2005) at a 10 arc-minute resolution to explore the climatic space of *D. suzukii*'s distribution.

In order to get a visual representation of the positions of our occurrences in the climatic space, we first performed a principal component analysis (PCA) on our climatic predictors using the “FactoMineR” library (Husson *et al.*, 2016) in R.

Ecological niche modeling

We first modeled the distribution of *D. suzukii* at a global scale (i.e. including its native and non-native range). considered as the best method to model distribution of one invasive species (Beaumont *et al.* 2009). This approach seems particularly suitable to model distribution of invasive species as it provides more accurate assessment of the potential invasive distribution at that stage of the invasion, where expanding populations are not expected to be at equilibrium with the climate (Beaumont *et al.* 2009). Then, to evaluate the capacity of the models to predict the presence of the invasive populations from the presence records of the ancestral population, we also modeled separately the native niche of *D. suzukii* (south-eastern Asia), its invasive niche in USA, and its invasive niche in Europe. We then projected the native modeled niche on the invasive range (in the USA and in Europe). Conversely, the native range was also projected from the two invasive ranges, separately. To avoid any potential geographic sampling bias within continent, we removed any duplicated presence data within a 50 km radius around each occurrence of the more sampled areas (in Europe and continental Asia).

We performed an ensemble of eight modeling techniques, all implemented in the biomod2 package (Thuiller *et al.*, 2013) in R v3.2.3 (R Core Team, 2015) : (1) generalized linear models (GLM), (2) generalized boosted models (GBM), (3) generalized additive model (GAM), (4) flexible discriminant analysis (FDA), (5) classification tree analysis (CTA), (6) artificial neural networks (ANN), (7) random forests (RF) and (8) multivariate adaptive regression splines (MARS). Because our datasets do not provide true absences, we generated 10 sets of 1000 artificial absence records (pseudo-absences) by randomly sampling from the grid cells of each study area (Barbet-Massin *et al.*, 2012). We replicated ten times the splitting of the data into a random subset including 70% of the data for calibration and one second subset with the remaining for evaluation by cross-validation. We used the true skill statistic (TSS) to evaluate the predictive performance of each modeling technique (Allouche *et al.*, 2006). To account for the inter-model variability in predictive performance, and because they have been shown to provide the best predictive performance, we chose consensus methods (Araújo & New, 2007; Marmion *et al.*, 2009) by averaging the predictions of models from the eight modeling techniques (and across GCMs to get the consensus distribution in the future), removing models with a TSS below the quality threshold of 0.5 (Marmion *et al.*, 2009). We calculated the average values of TSS across pseudo-absences and data-splitting runs to evaluate the ensemble model.

Additionally to test if presence in the invasive range could be better predicted knowing the invasion history of the species, we grouped our occurrence data at a more precise geographical scale corresponding to the genetic groups and of *D. suzukii* described in Fraimout *et al.* (2016) with China, Japan, Hawaii, western-USA, eastern-USA, Brazil, La Réunion and Europe. However, due to the small record of presences in the most recent colonization in Brazil, as well as the relative small areas that represent the two islands (Hawaii and La Réunion) prevent us to include them in the comparative analyses of niche modeling and niche overlap.

Future projections

Global distribution of *D. suzukii* was also projected by the year 2070 under a scenario of future climate change (Representative Concentration Pathways, hereafter called RCPs) with very high greenhouse gas emissions (RCP8.5). We derived future climate projections (2061-2080) from three global circulation models (GCMs; IPSL-CM5A-LR, MIROC-ESM and HadGEM2-ES; Hijmans *et al.*, 2005).

Measurement of niche similarities between ranges

In parallel to the crossed projections from niche modeling, we measured the niche overlap between native and invasive ranges. We used Broennimann's framework (2012) that provides a quantification of the niche similarity between two taxa that occurred in different geographic regions by correcting for bias related to the availability of environmental space for each. For each range, we performed a PCA calibrated on the entire environmental space of the study area (hereafter called PCA-env). We used the same eight climatic predictors than we selected for ecological niche modeling. The scores of the occurrences on the first two axes of PCA-env were projected onto a 100 × 100 grid, and a smoothed density of occurrences in environmental space was calculated using a kernel density function to compare ecological niches independently of sampling effort and of the spatial resolution of the environmental grid (Broennimann *et al.*, 2012).

To quantify the similarities in climatic requirements between ranges, we calculated Schoener's *D* (Schoener, 1968, reviewed in Warren *et al.*, 2008), a measure of niche overlap which varies from 0 (no overlap) to 1 (complete overlap). We then performed statistical tests of niche equivalency and similarity (Warren *et al.*, 2008). The hypothesis of niche equivalency was tested by calculating Schoener's *D* for 100 simulated datasets with a random sample of the same number of occurrences than the observed dataset. Niches are considered statistically equivalent (null hypothesis) if the observed value of *D* falls within the range of 95% of the simulated values (Broennimann *et al.*, 2012). We then performed a two-way test of niche similarity to determine whether niches are more similar than expected at random. We calculated the overlap between one range and the randomly shifted density of occurrences of another range, and *vice versa*. This process was repeated 1000 times. The null hypothesis (i.e., niches are no more similar to one another than expected at random) was rejected if the observed overlap was greater than 95% of the simulated values of overlap.

Results

*Variation in *D. sukuzii* presence data for climatic variables*

We visualized the distributions of the occurrence data in the climatic space by performing a PCA on our dataset (Fig. 1). The two first axes explained moderate amount of variation in the data (48.87%) Occurrences are discriminated along the first axis (25.57%) on average by an increasing gradient of precipitation with a decreasing seasonality and a decline of the maximum temperatures during the warmest month and along the second

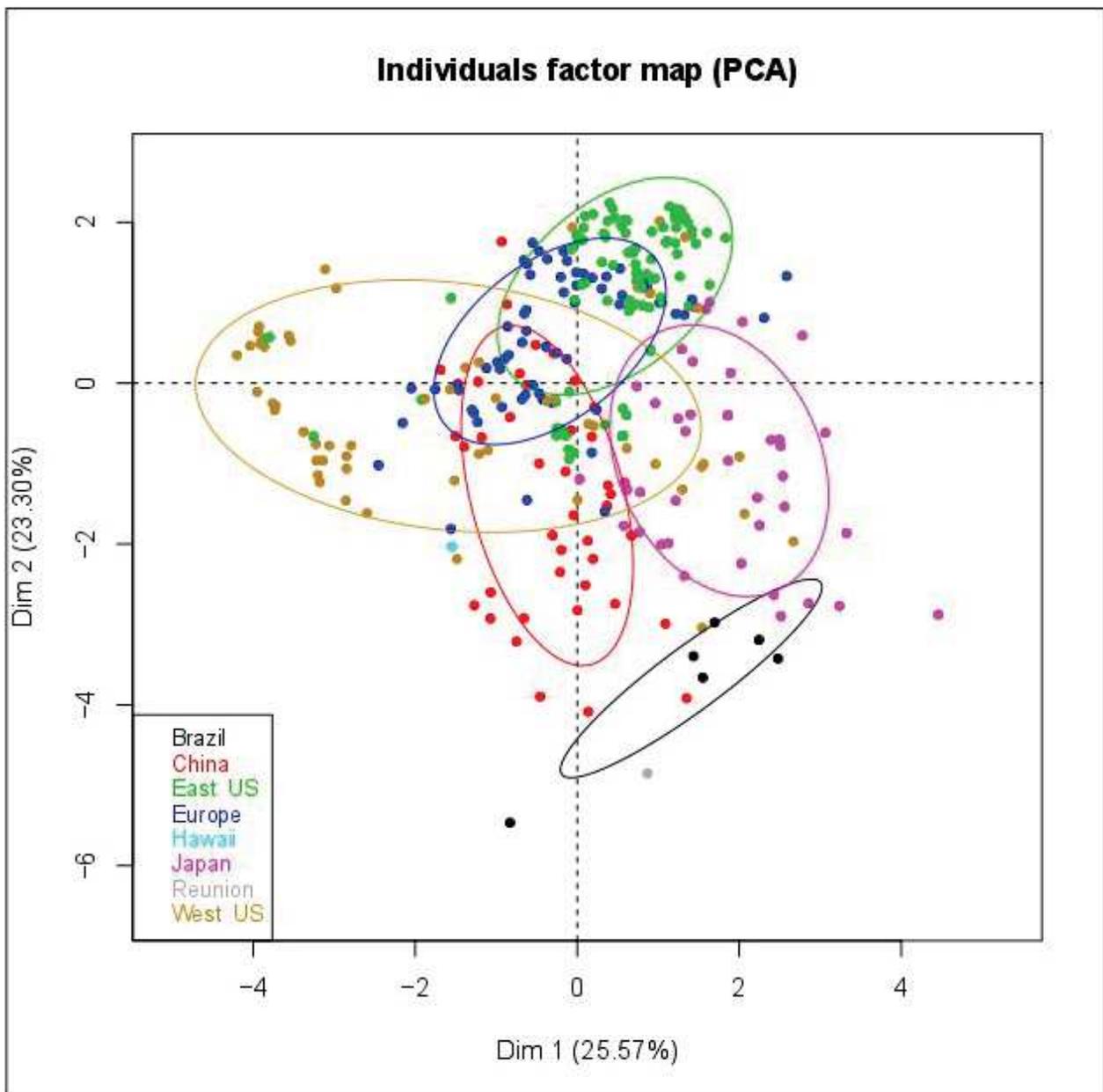


Figure 1. Plot of the PCA on worldwide presence data of *D. sukuzii*

Points represent each presence data gathered from the current native and invasive range of *D. sukuzii*.

axis by the latitudinal gradient and on average an increasing temperature seasonality with an increase of the minimum temperature during the coldest month (Fig 2).

Projections of the niche of D. sukuzii worldwide

The ensemble modeling of the worldwide current distribution of *D. sukuzii* obtained a high evaluation score (TSS = 0.821; Fig. 3). Future projections displayed a northward expansion of the environmental suitability and a decrease at the southern range margin for both the native and invasive ranges (Fig. 4).

Projections of the niche of D. sukuzii across its native and invasive ranges

All models computed for each range also showed a high predictive performance (USA: TSS = 0.842, Europe: TSS = 0.856, south-eastern Asia: TSS = 0.753, continental south-east Asia: TSS = 0.759, and Japan: TSS = 0.94) and presented in figure 5. Models calibrated on the native range were projected across both invasive ranges. These projections yielded presence probabilities on average higher or equal to 0.50 in central and western Europe and included all the described invasive range of *D. sukuzii*. In USA, projections from the native model yielded probability presence on average > 0.65 but only in eastern-USA. All models were reciprocally projected on each other. As expected, the occurrences in each invasive range were better predicted with the reference model (e.g. when European records are used to model European distribution). However, model calibrated on USA yielded fairly high presence probabilities in different part of Europe whereas the reciprocal projection of the European model in USA did not yield presence probabilities > 0.3

Measures of niche overlap between biogeographical regions

The environmental space available in the invasive ranges in USA and Europe were both widely included in the environmental space available in south-eastern Asia (Fig. 6). However, the niche overlaps between the native and the invasive ranges were low ($D = 0.361$ and $D = 0.323$ in USA and in Europe, respectively, Fig. 6). We found also a low overlap between Europe and USA ($D = 0.271$, Figure 7). This classification of the values of Schoener's D followed the suggestion of Rödder & Engler (2011) to facilitate the interpretations of the results: no overlap or very limited overlap for D comprise between 0 and 0.2, low overlap between 0.2 and 0.4 and moderate overlap between 0.4 and 0.6.

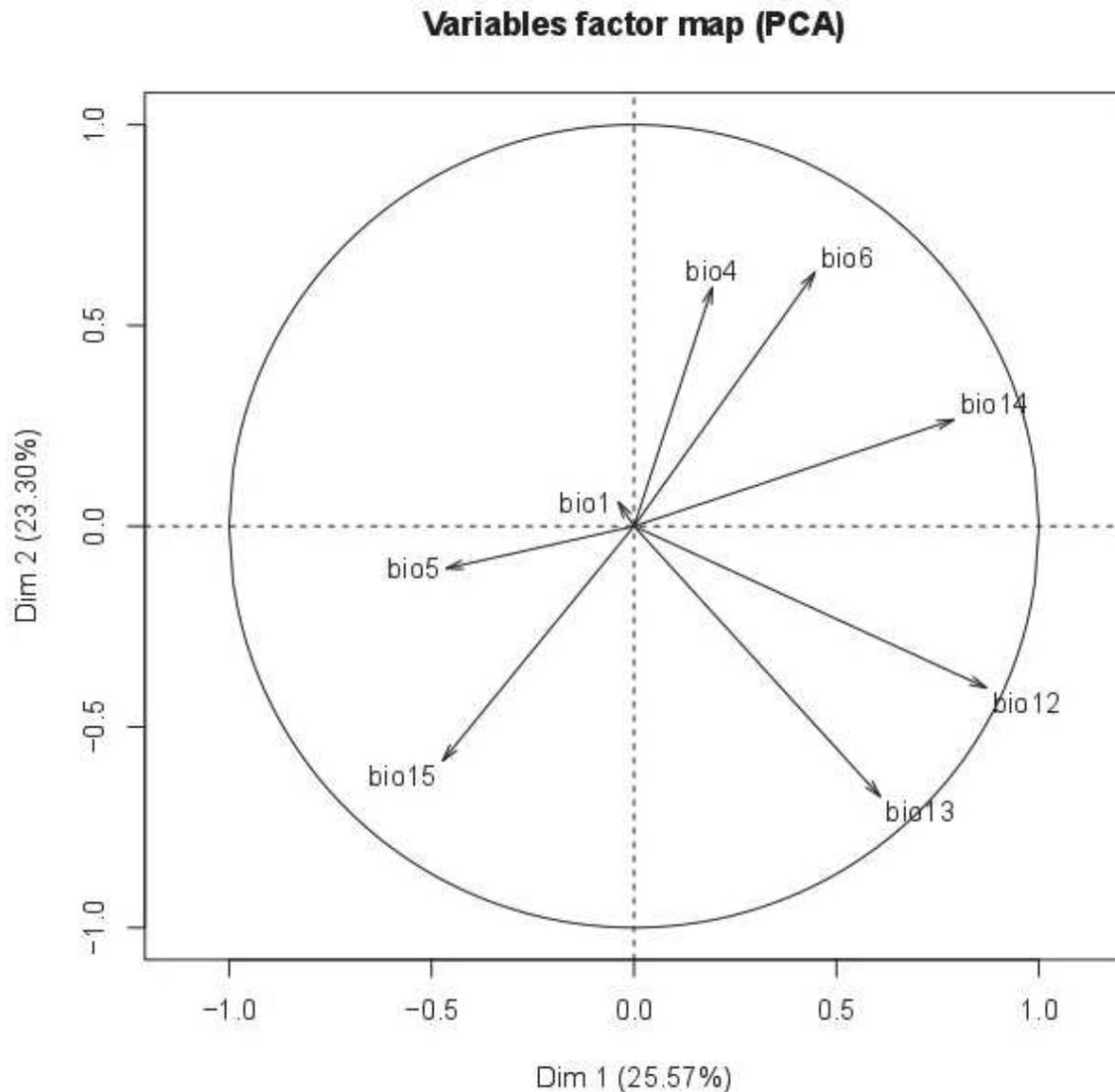


Figure 2. Plot of the PCA on climatic variables explaining the variations in *D. suzukii*'s presence data.

Climatic variables were gathered from the Worldclim Global Climate database (www.worldclim.org) and are coded as follows: bio1 = mean annual temperature; bio4 = temperature seasonality; bio5 = maximum temperature of the warmest month; bio6: minimum temperature of the coldest month; bio12 = annual precipitation; bio13 = precipitation of the wettest month; bio14 = precipitation of the driest month; bio15 = precipitation seasonality.

Figure 3. Plot of the ensemble modeling of the worldwide current distribution of *D. sukii*

Models were calibrated on all 344 presence data from *D. sukii* worldwide distribution and projected back globally. Color scale represent probabilities of presence from low (blue) to high (red).

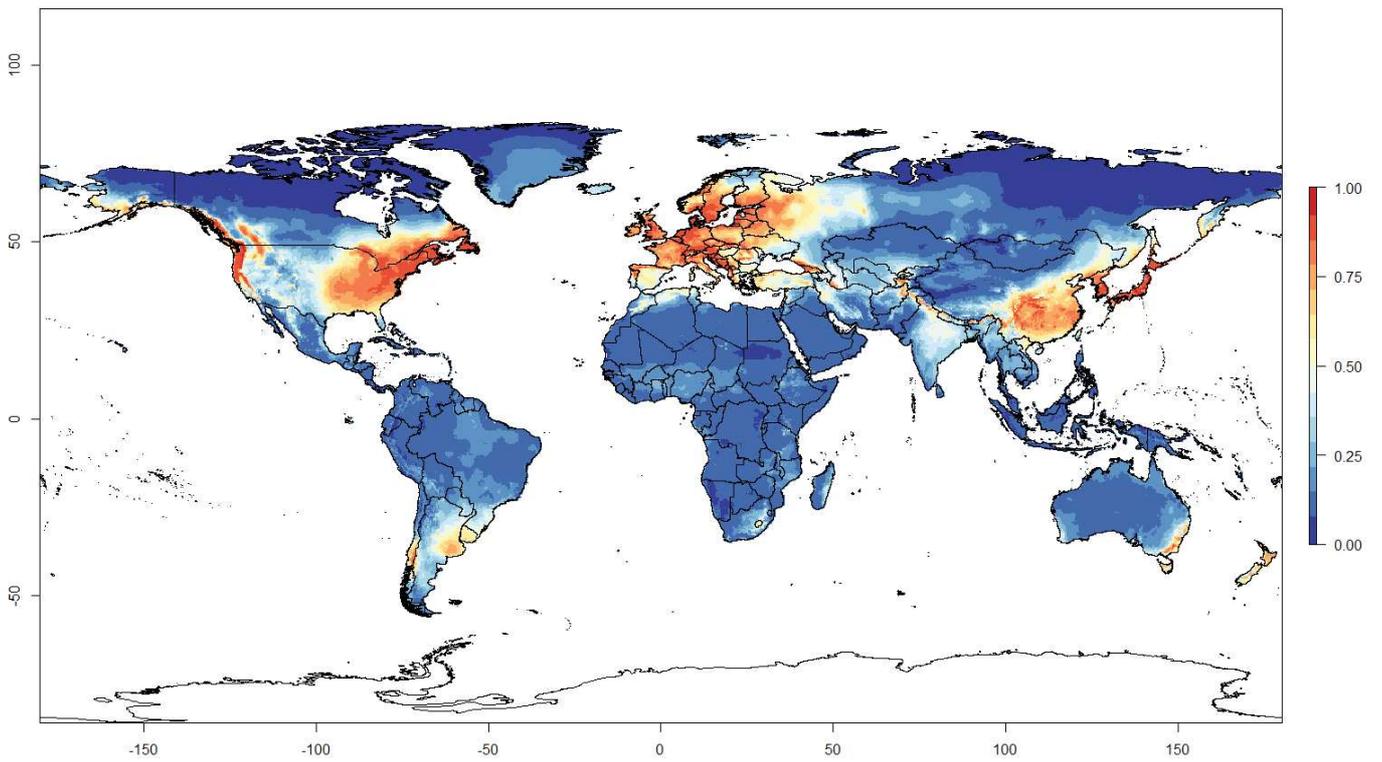


Figure 4. Plot of the ensemble modeling of future worldwide distribution of *D. sukii* under high climate change scenario.

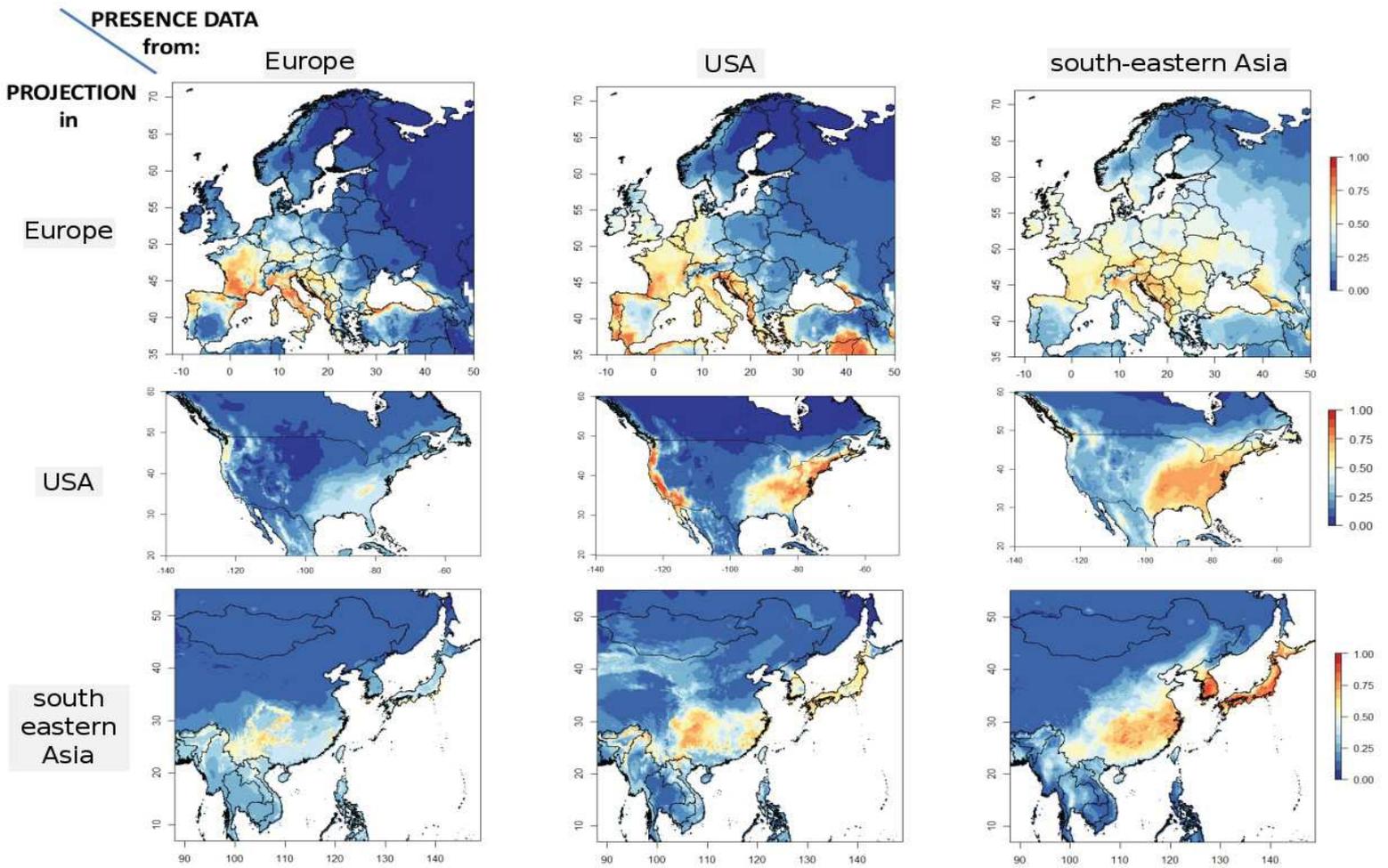


Figure 5. Reciprocal model projection between the Asian native range and both invasive ranges in Europe and USA.

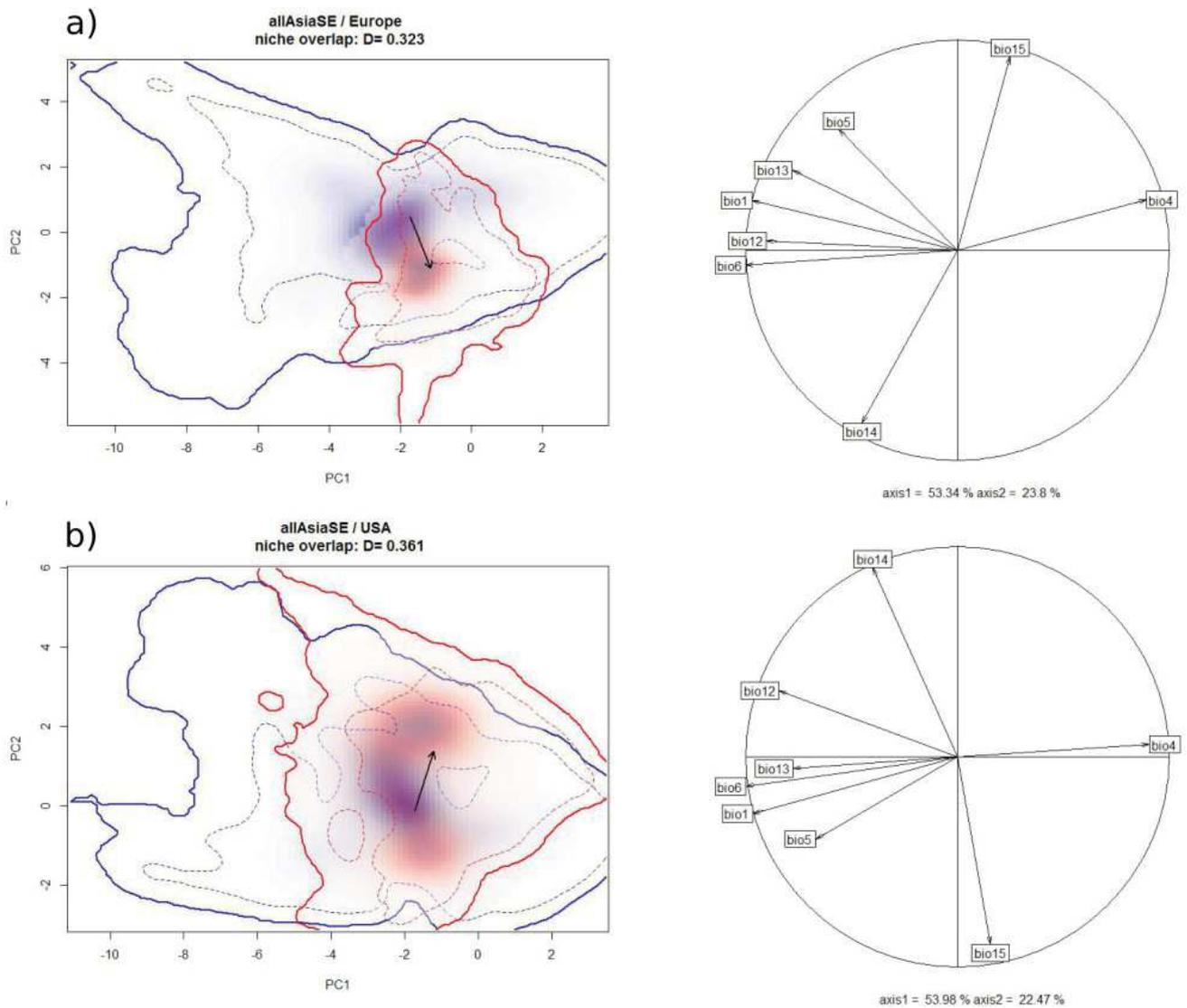


Figure 6. Niche overlap between the Asian native range and both invasive ranges in Europe and USA.

Principal components analysis are calibrated on environmental background (PCA-env) to compare the use of the environmental space in the native range in south-east Asia (in blue) and the invasive range a) in Europe, b) in USA (in red). Shadings correspond to the smoothed density of occurrences and solid and dotted lines correspond to 100% and 50% of the available environment space, respectively. The arrow indicates the change in the centroid from native niche to the invasive niche. Correlation circles in the left column show the contribution of each environmental variable to the two first axes of the principal components analysis. Environmental variables are the mean annual temperature (bio1), mean annual precipitation (bio12), temperature seasonality (bio4), 4) precipitation seasonality (bio15), temperature of the coldest month (bio6) and of the warmest month (bio5), precipitation of the driest month (bio14) and of the wettest month (bio13).

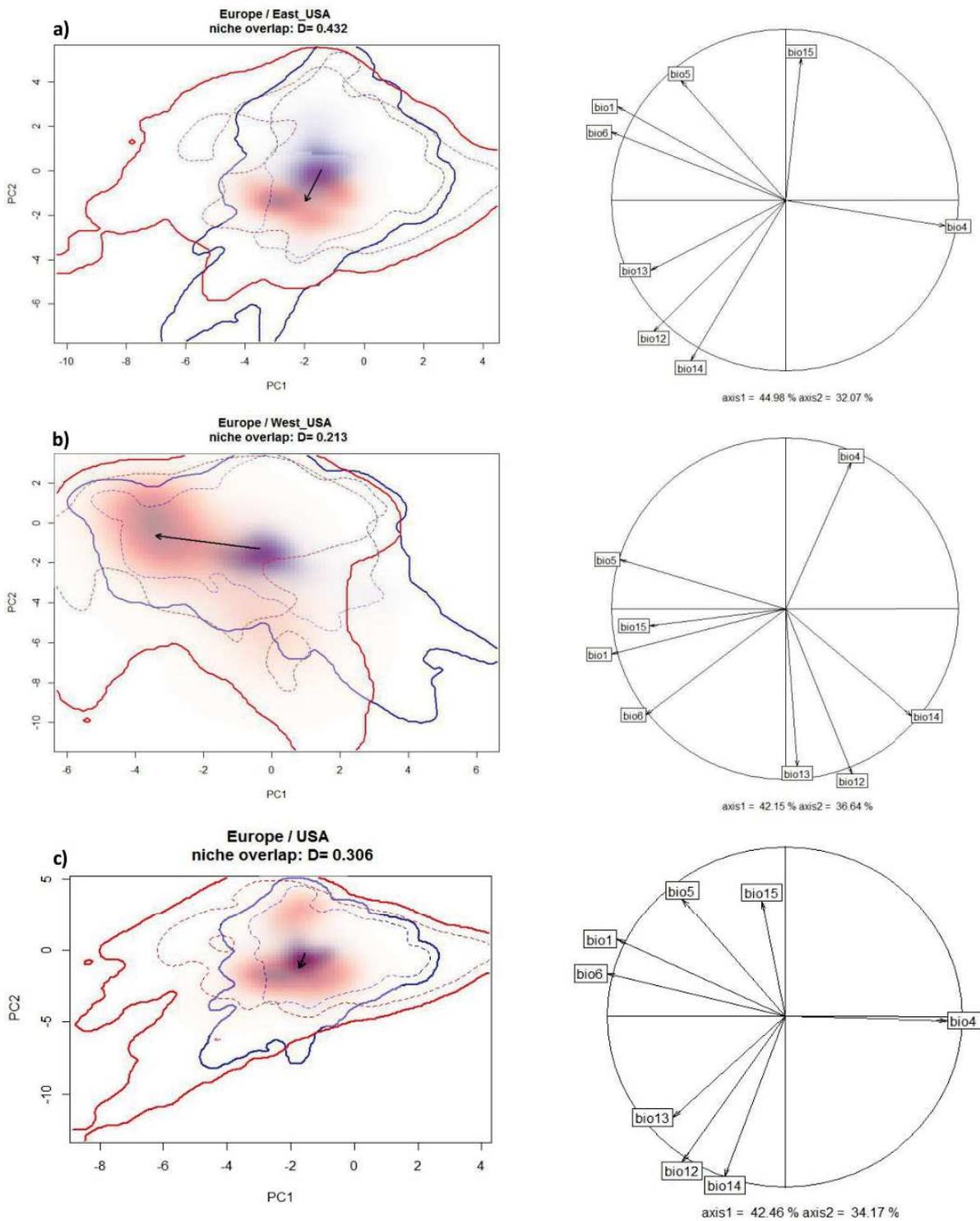


Figure 7. Niche overlap between the two main invasive ranges in Europe and USA. Principal components analysis are calibrated on environmental background (PCA-env) to compare the use of the environmental space in the invasive range in Europe (in blue) and the invasive range a) in eastern USA, b) in western USA, and c) in all USA (in red).

Accounting for origins in the native range

We refined our occurrence data to test if the presence in the invasive range was better predicted when accounting for the invasion history. Particularly, we expected that presence in Europe and USA would be better predicted when models were calibrated only on their continental Asian origin. However, all the measures of niche overlaps between invasive ranges (in North America and Europe) and the native ranges were very limited or low (from 0.174 to 0.354) when considering the continental south-eastern Asia and Japan apart (results not shown). These results suggest that the invasive populations do not share more climatic requirements with some particular part of native range. Moreover, presences on the invasive ranges were better predicted when the models were calibrated on the whole south-eastern Asia than when continental Asia and Japan were considered separately.

The model calibrated on south-eastern Asia predicted better the occurrence on the eastern coast of the USA than on the western coast. The overlap between south-eastern Asia and eastern USA was still moderate ($D = 0.425$) but higher than the one with western-USA ($D = 0.214$). Finally, we found a moderate overlap between Europe and eastern USA, but very limited between overlap between Europe and western USA (Fig. 7).

Discussion

Generalities about our questions and dataset

In the present paper we aimed at assessing the realized niche of *D. suzukii* in its native and invasive range and sought to identify potential niche shifts associated to its dispersion at a worldwide scale.

The generation of niche-based distribution models is inherently relying on the occurrence points used for calibration. These models will thus represent the realized niche of a species and be as close and descriptive of the fundamental niche as possible by increasing sampling efforts (Medley 2010). As for other invasive insects, the native distribution of *D. suzukii* in Asia is broad, and covers a large range of environments and countries (Asplen *et al.* 2015). Because our dataset incorporates occurrences from a wide distribution in the native range, it should encompass a sufficient amount of environmental variation to calibrate our models and predict invasive ranges. One additional challenge in the generation of niche-based distributions from presence data is the lack of consideration for possible biotic interactions. Indeed, absence of occurrence of a species in a geographical area may be the result of out-competition of the latter by a conspecific

population, presence of predators or other patterns of biotic interactions (Soberon & Peterson 2005, Hirzel & Le Lay 2008). It is thus primordial to consider that absence of distribution records may not be due solely to unsuitable environmental conditions. In the case of *D. suzukii* one of its main ecological characteristic is the evolution of its ovipositing behavior from an ancestral '*Drosophila*-like' decaying substrate, to a wide variety of ripening stonefruit and berries (Asplen *et al.* 2015, Keeseey *et al.* 2015). This specialization in egg-laying behavior opened a new niche for the species which remained unoccupied by other *drosophilid* species (Poyet *et al.* 2014), and therefore relaxes the weight of potential biotic interactions in the calibration accuracy of our models.

Calibration of models and projection

The projection of our models calibrated on data from the entire native range (i.e. south-east Asia) revealed high probabilities of presence in all current invasive distribution of *D. suzukii* we could project onto (namely Europe, USA, and Brazil) suggesting a fair relevance of our selected climatic variables. Interestingly, the west coast of USA and the state of California particularly, did not yield high probability values whichever the native niche used for projection. The first outbreak of *D. suzukii* in continental USA was reported in this particular area (Hauser 2011) and was further described as an admixture event between Hawaiian and Asian populations (Fraimout *et al.* 2016). Low probability values in western USA could indicate a shift in *D. suzukii* climatic niche from the native Asian range to this particular region. The variability of climatic conditions along the western coast of USA (Mock 1996, Shinker 2010) and particularly the higher levels of drought in southern California may represent a challenge to *D. suzukii* life cycle. Our result suggest that *D. suzukii* most likely shifted its niche in direction of warmer and much drier climate in western USA. Dessication resistance is a well studied phenotypic trait in *Drosophila* (e.g. Hoffmann & Harshman 1999, Hoffmann *et al.* 2003, Matzkin *et al.* 2009) and could have been selected in *D. suzukii* colonizers in arid regions of western-US. As western-US populations were founded by two genetically differentiated populations, shift in the climatic tolerance of the species possibly benefited from the admixture of Asian and Hawaiian propagules producing higher genetic variance and novel genotypes able to respond to this environmental pressure (Kolbe *et al.* 2007, Uller & Leimu 2011). These newly founded populations most likely served as sources for the subsequent invasion of the eastern US (Fraimout *et al.* 2016). As pointed out in Fraimout *et al.* (2016) eastern US *D. suzukii* populations were founded not by secondary Asian introductions but by dispersing

individuals from southern California (San Diego region). Owing to higher probabilities of presence in eastern US as modeled from data in the Asian native area (Fig. 3), direct invasion of eastern US by Asian propagules could have been historically possible. It is possible that western US, as a major entry point in continental USA for trade shipments (www.usda.gov) served as a geographical (and climatic) buffer for the installment of *D. suzukii* in USA. The putative niche shift in this region could have then unlock the global invasion of USA by the species, which would explain the speed of spread of *D. suzukii* in this country.

Interestingly, reciprocal projections from Europe to USA and *vice versa* revealed an asymmetrical relationship between the two areas. Whereas projection from European presence data to USA did not reveal high presence probabilities, the reciprocal projection from USA to Europe indicated high presence probabilities in Europe (Fig. 5). Furthermore, eastern US and European climatic envelope were not differentiated by ordination in the two first axes of the PCA (see Fig. 1 for graphic representation on PC1 and PC2) This is of particular interest as the most probable global invasion scenario of *D. suzukii* revealed an asymmetrical gene flow from eastern US to the northern part of Europe.

Better results when accounting for invasion routes?

We found no evidence supporting a better prediction of presence in the invasive range by reducing the datasets to the most likely areas of origins in the native range. The separation of the native range into Japan and continental Asia did not yield any higher probabilities of presence in the invasive European or North American range. In fact, probabilities of presence were largely reduced in both the invasive range when models were calibrated onto Japan or continental Asia. This may be due to the reduction in data points used for model calibration thus reducing the efficiency of niche projection in the invasive ranges (Stockwell & Peterson 2002). These results also emphasize the limitations of considering the climatic similarities as the only determinant of the invasiveness of one species in one range.

Uncolonized areas at risk

Although *D. suzukii* has widely colonized most continents of the world, our model projection show that some geographical areas could still be at risk. North and South-Africa especially as well as the eastern coast of Australia should take into account the present results to monitor potential *D. suzukii* infestation. Furthermore, owing to our future

projections (Fig. 4) particular attention should be paid in these areas as well as in the northern range of *D. suzukii* distribution in Europe and USA which could significantly increase along with temperatures in the next fifty years.

Concluding remarks

Our analysis revealed several points of interest in the understanding of the worldwide invasion of *D. suzukii*: i) *D. suzukii* seems to have a broad climatic niche inherent from its large native range in south-eastern Asia, ii) niche breadth was characterized by high probabilities of presence in most of the invasive range in Europe and eastern USA when models were calibrated from Asian presence data. This result points toward a generalist status for *D. suzukii* (at least for environmental conditions) that could explain the pace of worldwide colonization that was historically constrained by a lack of migration means. iii) Colonization of western US, and southern California particularly seems to involved a shift in the climatic niche of *D. suzukii*, as indicated by the very low to null probabilities of presence in this area regardless of the calibrations used for the model projections. iv) Reciprocal projections revealed high presence probabilities in Europe with models calibrated in USA, emphasizing the role of eastern US as a secondary source of introductions in Europe. v) Particular attention should be paid in uncolonized areas such as Australia, New-Zealand, North and South Africa, as well as northern areas in Europe and North-America which could be subjected to future *D. suzukii* introductions within the next decades.

Acknowledgments

AF was supported by state funding by the Agence Nationale de la Recherche, through the LabEx ANR-10-LABX-0003-BCDiv, of the program “Investissements d’avenir” (ANR-11-IDEX-0004-02)

References

- Allouche, O., Tsoar, A., and Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS): Assessing the accuracy of distribution models. *Journal of Applied Ecology* 43, 1223–1232.
- Araujo, M., and New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22, 42–47.

- Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., *et al.* (2015). Invasion biology of spotted wing *Drosophila* (*Drosophila suzukii*): a global perspective and future priorities. *Journal of Pest Science* *88*, 469–494.
- Austin, M.P., and Van Niel, K.P. (2011). Improving species distribution models for climate change studies: variable selection and scale: Species distribution models for climate change studies. *Journal of Biogeography* *38*, 1–8.
- Barbet-Massin, M., and Jetz, W. (2014). A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Diversity and Distributions* *20*, 1285–1295.
- Barbet-Massin, M., Jiguet, F., Albert, C.H., and Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many?: How to use pseudo-absences in niche modelling? *Methods in Ecology and Evolution* *3*, 327–338.
- Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R., and Hughes, L. (2009). Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions* *15*, 409–420.
- Broennimann, O., Thuiller, W., Hughes, G., Midgley, G.F., Alkemade, J.M.R., and Guisan, A. (2006). Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology* *12*, 1079–1093.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T., and Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters* *10*, 701–709.
- Calabria, G., Máca, J., Bächli, G., Serra, L., and Pascual, M. (2012). First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe: First record of *Drosophila suzukii* in Europe. *Journal of Applied Entomology* *136*, 139–147.
- Deprá, M., Poppe, J.L., Schmitz, H.J., De Toni, D.C., and Valente, V.L.S. (2014). The first records of the invasive pest *Drosophila suzukii* in the South American continent. *Journal of Pest Science* *87*, 379–383.
- Facon, B., Genton, B., Shykoff, J., Jarne, P., Estoup, A., and David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* *21*, 130–135.
- Godefroid, M., Cruaud, A., Rossi, J.-P., and Rasplus, J.-Y. (2015). Assessing the Risk of Invasion by Tephritid Fruit Flies: Intraspecific Divergence Matters. *PLOS ONE* *10*, e0135209.
- Harris, A., and Shaw, B. (2014). First record of *Drosophila suzukii* (Matsumura)(Diptera, Drosophilidae) in Great Britain. *Dipterists Digest* *21*, 189–192.

- Hauser, M. (2011). A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Management Science* 67, 1352–1357.
- Heikkinen, R.K., Luoto, M., Araujo, M.B., Virkkala, R., Thuiller, W., and Sykes, M.T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography* 30, 751–777.
- Hirzel, A.H., and Le Lay, G. (2008). Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45, 1372–1381.
- Hoffmann, A.A., and Harshman, L.G. (1999). Desiccation and starvation resistance in *Drosophila*: patterns of variation at the species, population and intrapopulation levels. *Heredity* 83, 637–643.
- Hoffmann, A.A., Sørensen, J.G., and Loeschcke, V. (2003). Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *Journal of Thermal Biology* 28, 175–216.
- Hutchinson, G.E. (1957). Population studies: Animal ecology and demography. *Bulletin of Mathematical Biology* 53, 193–213.
- Keesey, I.W., Knaden, M., and Hansson, B.S. (2015). Olfactory Specialization in *Drosophila suzukii* Supports an Ecological Shift in Host Preference from Rotten to Fresh Fruit. *Journal of Chemical Ecology* 41, 121–128.
- Kolbe, J.J., Larson, A., Losos, J.B., and de Queiroz, K. (2008). Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biology Letters* 4, 434–437.
- Lê, S., Josse, J., Husson, F., and others (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software* 25, 1–18.
- Lecocq, T., Rasmont, P., Harpke, A., and Schweiger, O. (2016). Improving International Trade Regulation by Considering Intraspecific Variation for Invasion Risk Assessment of Commercially Traded Species: The *Bombus terrestris* Case: Subspecies-based invasion risk assessment. *Conservation Letters* 9, 281–289.
- Lee, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17, 386–391.
- Lee, J.C., Bruck, D.J., Curry, H., Edwards, D., Haviland, D.R., Van Steenwyk, R.A., and Yorgey, B.M. (2011). The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest Management Science* 67, 1358–1367.
- Lengyel, G.D., Orosz, S., Kiss, B., Lupták, R., and Kárpáti, Z. (2015). New records and present status of the invasive Spotted Wing *Drosophila*, *Drosophila suzukii* (Matsumura, 1931) (Diptera) in Hungary. *Acta Zoologica Academiae Scientiarum Hungaricae* 61, 73–80.
- Lombaert, E., Guillemaud, T., Lundgren, J., Koch, R., Facon, B., Grez, A., Loomans, A.,

- Malausa, T., Nedved, O., Rhule, E., *et al.* (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology* 23, 5979–5997.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., and Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* 15, 59–69.
- Matzkin, L.M., Watts, T.D., and Markow, T.A. (2009). Evolution of stress resistance in *Drosophila*: interspecific variation in tolerance to desiccation and starvation. *Functional Ecology* 23, 521–527.
- Medley, K.A. (2010). Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models: Niche shifts and global invasion. *Global Ecology and Biogeography* 19, 122–133.
- Peterson, A.T., and Holt, R.D. (2003). Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. *Ecology Letters* 6, 774–782.
- Poyet, M., Eslin, P., Héraude, M., Le Roux, V., Prévost, G., Gibert, P., and Chabrierie, O. (2014). Invasive host for invasive pest: when the Asiatic cherry fly (*Drosophila suzukii*) meets the American black cherry (*Prunus serotina*) in Europe: When *D. suzukii* meets *P. serotina*. *Agricultural and Forest Entomology* 16, 251–259.
- Rödger, D., and Engler, J.O. (2011). Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks: Quantitative metrics of niche overlap. *Global Ecology and Biogeography* 20, 915–927.
- Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A., Baughman, S., Cabin, R.J., Cohen, J.E., Ellstrand, N.C., *et al.* (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics* 305–332.
- Shinker, J.J. (2010). Visualizing Spatial Heterogeneity of Western U.S. Climate Variability. *Earth Interactions* 14, 1–15.
- Soberon, J., and Peterson, A.T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas.
- Stockwell, D.R., and Peterson, A.T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148, 1–13.
- Szücs, M., Melbourne, B.A., Tuff, T., and Hufbauer, R.A. (2014). The roles of demography and genetics in the early stages of colonization. *Proceedings of the Royal Society B: Biological Sciences* 281, 20141073–20141073.
- Tingley, R., Vallinoto, M., Sequeira, F., and Kearney, M.R. (2014). Realized niche shift during a global biological invasion. *Proceedings of the National Academy of Sciences* 111, 10233–10238.
- Uller, T., and Leimu, R. (2011). Founder events predict changes in genetic diversity during

human-mediated range expansions. *Global Change Biology* 17, 3478–3485.

Warren, D.L., Glor, R.E., and Turelli, M. (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62, 2868–2883.

Yoshida, T., Goka, K., Ishihama, F., Ishihara, M., and Kudo, S. (2007). Biological invasion as a natural experiment of the evolutionary processes: introduction of the special feature. *Ecological Research* 22, 849–854.

Conclusion générale

L'invasion de *Drosophila suzukii* a été l'objet d'un grand nombre d'études depuis son introduction aux États-Unis et en Europe en 2008. Cette invasion intrigue par son caractère synchrone et fascine par l'impressionnant succès invasif de l'espèce. Bien que le premier point n'ait jamais été directement abordé par la communauté scientifique, le second à quant à lui bénéficié d'un grand nombre d'études ayant pour but d'identifier les mécanismes biologiques à l'origine d'un tel succès. Qu'il s'agisse de la résistance au froid et de la capacité d'hivernage (Shearer *et al.* 2016, Plantamp *et al.* 2016), de l'association avec les micro-organismes (i.e. levures, Hamby *et al.* 2012, Mori *et al.* 2016, ou *Wolbachia sp.* Cattell *et al.* 2016a,b ; voir Annexes) ou encore de la large gamme de plante hôte dans l'aire envahie (Lee *et al.* 2011, Poyet *et al.* 2014), de nombreuses hypothèses ont été formulées pour expliquer ce succès.

Dans ce travail de thèse, nous nous sommes intéressés au potentiel évolutif de l'espèce, et plus particulièrement, nous avons étudié le rôle potentiel des mécanismes de réponse adaptative (i.e. via sélection sur la variance génétique et plasticité phénotypique) tout au long de l'histoire de cette invasion à grande échelle.

Résultats principaux et perspectives

Concernant son historique tout d'abord : les informations concernant les voies de migration des espèces envahissantes sont capitales dans leur gestion (Kolar & Lodge 2001, Drake & Lodge 2004, Hulme 2009) mais également pour comprendre les dynamiques démo-génétiques mises en place pendant les trois étapes du processus d'invasion (Estoup & Guillemaud 2010). Dans le cas de *D. suzukii*, il semble malheureusement impossible de déterminer avec précision les moyens physiques de migrations de l'espèce (mais voir Cini *et al.* 2014) du fait de la relation étroite entre la biologie de l'espèce et la complexité des réseaux commerciaux de produits agricoles. Toutefois, comme nous l'avons démontré ici, il est possible de retracer l'histoire de l'invasion à l'échelle mondiale, ainsi que d'identifier des événements clés de cette histoire, à partir de données moléculaires et des méthodes d'ABC-Random Forest (Pudlo *et al.* 2016). En résumé : nos résultats indiquent trois événements d'introduction majeurs depuis l'aire native en Asie, à Hawaii, en Europe et à l'Ouest des États-Unis et à Hawaii. Ces zones d'introduction primaire ont ensuite servi elles-mêmes de sources (i.e. populations têtes de ponts, Lombaert *et al.* 2011) pour les introductions secondaires de l'espèce dans l'hémisphère Sud.

Ces résultats ont également soulevé plusieurs questions concernant le succès de cette invasion : comment expliquer la pérennité des populations Hawaïennes ? Ces populations sont

des exemples parfaits du *paradoxe des invasions* (Estoup *et al.* 2016) : comment cette population a pu s'établir à Hawaï après un goulot d'étranglement tel, et la perte de plus de la moitié de sa variation allélique ? Ces populations méritent d'être analysées plus en détails notamment pour l'identification de processus génétiques pouvant expliquer un tel succès (e.g. purge d'allèles délétères, adaptation locale etc.). Quelle a été la dynamique d'invasion au sein de chaque zone envahie ? Nous avons identifié des scénarios complexes d'invasion, notamment sur la côté Ouest des États-Unis, où de multiples événements d'introduction et de dispersion ont participé à l'invasion de cette zone. Toutefois, l'utilisation de marqueurs microsatellites – bien que particulièrement adéquate dans le cas d'une reconstruction à l'échelle globale – semble atteindre ses limites lorsque des dynamiques démographiques plus fines doivent être analysées. Dans le cas de l'Ouest des États-Unis, la séquence chronologique des différents événements d'introduction n'a pas pu être identifiée avec précision. De la même façon, l'histoire de l'invasion en Europe, et plus particulièrement la contribution de propagules Nord-Américaines à la colonisation du nord de l'Europe, pourra être analysée plus finement à l'avenir. Cette échelle d'analyse (i.e. intra-continentale), demande la caractérisation d'un plus grand nombre de marqueurs moléculaires, et pour cela, il semble que l'exploration du génome de *D. sukii* soit une perspective prometteuse (Cristescu 2015, Estoup *et al.* 2016)

La reconstruction de l'histoire d'invasion nous a donc permis d'établir des liens de parenté entre populations envahissantes et populations native de *D. sukii*. Cette information est une donnée essentielle à l'étude comparative des patrons de variations quantitative et de la divergence phénotypique entre populations. Nous avons donc utilisé ces informations apportées par la génétique des populations pour tester les attendus théoriques sur l'évolution de cette variation quantitative. L'évolution de la matrice **G** et son rôle dans les trajectoires évolutives est encore aujourd'hui un sujet théorique central en biologie évolutive et ne peut que bénéficier des résultats des approches empiriques (e.g. Chantepie *et al.* 2015, Paccard *et al.* 2016, Puentes *et al.* 2016, Simon *et al.* 2016). En nous intéressant à l'invasion de *D. sukii* nous nous proposons de tester la stabilité de la matrice **G** dans un contexte incluant potentiellement tous les paramètres pouvant influencer cette stabilité (i.e. dérive génétique, tailles de populations fluctuantes, sélection directionnelle ou stabilisante sur l'aile ; Jones *et al.* 2003, 2004, Revell 2007, Arnold *et al.* 2008). A cet égard, nos résultats indiquent une stabilité de certains aspects de la géométrie de **G** (i.e. forme et orientation) et une évolution de sa taille. Nous avons donc pu confirmer des attendus sur l'évolution du volume de **G** dans un contexte neutre – une réduction du volume dans les populations soumises à goulot d'étranglement, potentiellement compensée par des événements d'admixture – mais n'écartons pas l'importance de la sélection dans sa stabilité entre toutes les populations étudiées. Pour un caractère comme l'aile, ces résultats confirmeraient une forte sélection stabilisante pour ce trait, et le maintien de sa forme dans une zone proche de l'optimum

de valeur sélective sur le paysage adaptatif. Une perspective particulièrement intéressante serait d'estimer la matrice **G** à partir d'autres traits potentiellement impliqués dans le succès de l'invasion. Parmi ces traits, de bon candidats seraient par exemple des caractères liés à la reproduction tels que la fécondité ou l'âge à la reproduction (e.g. Johansson *et al.* 2012, Bacigalupe *et al.* 2013).

Le dernier aspect que nous avons abordé concerne la relation entre l'invasion de *D. suzukii* et l'environnement. La modélisation de la niche climatique de *D. suzukii* a confirmé le caractère tempéré de la niche climatique ancestrale de cette espèce (e.g. Ometto *et al.* 2013). De plus, cette niche ancestrale semble englober une grande partie des climats rencontrés par l'espèce dans l'aire envahie, et pourrait expliquer la rapidité d'expansion de *D. suzukii* dans ces régions. La similarité entre environnements natifs et envahissants pourrait alors expliquer l'absence de différence de plasticité phénotypique que nous avons observée pour les populations Japonaises, Françaises et Américaines. En outre, si l'optimum historique de température de *D. suzukii* se situe dans une gamme relativement basse, il semble réaliste de formuler l'hypothèse selon laquelle les populations ayant envahi des environnements proches de cet optimum, puissent être pré-adaptées à ces conditions. Dans notre cas, ceci pourrait expliquer les meilleures performances de vol des populations utilisées dans l'analyse de plasticité phénotypique à des températures plus basses. Une façon directe de tester cette hypothèse serait de contrôler la température à laquelle les individus volent, afin de tester si par exemple, les mouches élevées au froid sont plus performantes dans un air froid, et *vice versa* (e.g. Frazier *et al.* 2008). L'utilisation de populations provenant de latitudes plus au sud de la distribution, et plus particulièrement du sud de la Californie (où nous avons identifié un changement dans la niche climatique), pourrait également permettre de tester cette hypothèse. Dans ce cas, nous prédirions une meilleure performance de vol des populations du Sud, à des températures plus élevées.

Apports théoriques et méthodologiques de la thèse

Chapitre 1 : génétique des populations

De nombreuses données auront été produites tout au long de cette thèse, qui pourront à l'avenir, faciliter les recherches sur *D. suzukii*. Le développement de marqueurs microsatellites spécifiques ainsi que le génotypage des populations mondiales de l'espèce peuvent servir de référence à la caractérisation génétique de nouvelles populations envahissantes et, possiblement à l'identification de leur source. De plus, nous avons montré que l'application des méthodes d'ABC-Random Forest (Pudlo *et al.* 2016) à la reconstruction des routes d'invasions est une avancée majeure dans ce

domaine. En rendant public nos jeux de données, ainsi que nos *scripts* d'analyses d'ABC-RF nous espérons que ces résultats bénéficieront aux analyses futures de génétique des populations portant sur *D. suzukii* ou sur d'autres modèles d'invasion. Enfin, la collecte et la centralisation d'un grand nombre d'échantillons à l'échelle mondiale de la distribution nous a permis de créer une véritable bibliothèque de données, qu'il est alors possible d'explorer grâce aux méthodes de génomiques et de séquençage nouvelle génération.

Chapitre 2 : génétique quantitative

La stabilité et l'évolution de la matrice **G** est un élément central des modèles prédictifs en biologie évolutive et l'approche comparative entre matrices ancestrales et dérivées pour aborder les questions relatives à cette stabilité semble particulièrement prometteuse (e.g. Stepan *et al.* 2002). Ici, nous avons apporté des éléments de réponse important quand à la stabilité de la structure de **G** dans le contexte particulier des invasions biologiques. De plus, la composition des populations stock de *D. suzukii* et la conservation des individus de chaque lignées utilisées au cours de ce travail, permettra de poursuivre l'analyse de nombreux traits quantitatifs. Bien que l'accès aux traits d'histoire de vie nécessite la mise en place de nouvelles populations de laboratoire, d'importantes questions peuvent être abordées à partir d'autres traits morphologiques issus des données de cette thèse. Par exemple, l'analyse de l'asymétrie fluctuante (e.g. Klingenberg 2015) à partir des deux ailes, est une perspective particulièrement intéressante afin d'évaluer la stabilité de développement des populations envahissantes par rapport aux natives (e.g. Leary & Allendorf 1989, Parsons 1990, Leamy & Klingenberg 2005). Cette approche permettrait en outre l'analyse de la modularité développementale, en relation avec la modularité génétique (e.g. Klingenberg, 2003 ; Klingenberg *et al.* 2010). Un autre trait morphologique méritant d'être analysé de façon détaillée est le caractère « clé » du succès écologique de *D. suzukii* à l'échelle macro-évolutive, son ovipositeur hypertrophié (Atallah *et al.* 2014). Aucune donnée n'est disponible concernant la variation intraspécifique, et bien que certains gènes majeurs aient été identifiés dans le contrôle de son développement (Prudhomme, pers. comm.), l'architecture génétique de ce trait chez *D. suzukii* reste inconnue.

Chapitre 3 : plasticité phénotypique

L'analyse des performances de vol est une approche idéale pour étudier la relation entre forme et fonction du phénotype. De plus, les données de trajectoires de vol obtenues ici pourraient être particulièrement intéressantes pour une analyse comportementale du vol chez *D. suzukii*. Le vol stationnaire est bien connu chez la drosophile (Ellington 1984, Fry *et al.* 2005), mais est-il possible de caractériser des comportements plus fins à partir des trajectoires de vol des individus? Les vols

de fuites sont-ils aérodynamiquement différents des vols stationnaires ? L'analyse des trajectoires obtenues au cours de cette thèse pourrait aider à répondre à ces questions. Enfin, la caractérisation précise de la norme de réaction de *D. sukii* à plusieurs températures pourrait remettre en question l'existence des « morphes saisonniers » décrits pour cette espèce dans la littérature (e.g. Stephens *et al.* 2015, Revadi *et al.* 2015, Shearer *et al.* 2016). En effet, de nombreux articles font état de morphotypes discrets chez *D. sukii* avec une forme d'été (i.e. petite taille, faible pigmentation) et une forme hivernale (i.e. grande taille, forte pigmentation). Ce résultat semble toutefois surprenant pour une espèce de drosophile (e.g. David *et al.* 1994) et pourrait être vérifiée par une analyse expérimentale.

Chapitre 4 : modélisation de niche

Grâce aux modèles de modélisation de niche climatique, nous avons pu établir des prédictions globales quant aux probabilités de présence actuelle et future de *D. sukii*. Ces données sont capitales dans la gestion des futures invasions, particulièrement dans l'identification des zones non-colonisées à risques (Peterson 2003, Herborg *et al.* 2007, Broennimann & Guisan 2008).

Conclusions personnelles

Ce travail n'aurait jamais pu se faire sans tous les collaborateurs qui y ont participé. Le vaste champ théorique balayé – génétique des populations, génétique quantitative, évolution, morphométrie géométrique, physiologie, entomologie, modélisation etc – aurait pu vite devenir insurmontable, ou, a minima, confirmer l'adage du *Jack of all trades, master of none*. Néanmoins, grâce à toutes ces collaborations, nous avons pu mettre en place un ensemble cohérent d'hypothèses et d'expériences, qui ont abouti en une thèse aussi complète que possible, sur l'étude de la variation génétique et phénotypique.

Enfin, à l'heure où la science du génome nous offre des outils permettant d'aller de plus en plus loin dans la compréhension du vivant (e.g. Hsu *et al.* 2014), l'approche naturaliste des études de terrain et du phénotype n'a pour autant pas perdu de son importance. Pour comprendre son modèle biologique il est fondamental de comprendre son environnement, et comme il l'a été formulé par Michaud (1941) « *Whatever the subject may be – taxonomy, morphology, physiology, embryology, genetics, behaviorism, or conservation – there is always much to be gained by a fuller knowledge of how living things exist within their own native habitats* ».

Références bibliographiques

- Ackermann, R.R., and Cheverud, J.M. (2002). Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. *American Journal of Physical Anthropology* 117, 260–271.
- Ackermann, R.R., and Cheverud, J.M. (2004). Detecting genetic drift versus selection in human evolution. *Proceedings of the National Academy of Sciences* 101, 17946–17951.
- Adams, M.D. (2000). *Drosophila melanogaster* The Genome Sequence of. *Science* 287, 2185–2185.
- Adams, D.C., Rohlf, F.J., and Slice, D.E. (2004). Geometric morphometrics: Ten years of progress following the “revolution.” *Italian Journal of Zoology* 71, 5–16.
- Adrion, J.R., Kousathanas, A., Pascual, M., Burrack, H.J., Haddad, N.M., Bergland, A.O., Machado, H., Sackton, T.B., Schlenke, T.A., Watada, M., *et al.* (2014). *Drosophila suzukii*: The Genetic Footprint of a Recent, Worldwide Invasion. *Molecular Biology and Evolution* 31, 3148–3163.
- Agrawal, A.F., and Stinchcombe, J.R. (2009). How much do genetic covariances alter the rate of adaptation? *Proceedings of the Royal Society B: Biological Sciences* 276, 1183–1191.
- Aguirre, J.D., Hine, E., McGuigan, K., and Blows, M.W. (2014). Comparing G: multivariate analysis of genetic variation in multiple populations. *Heredity* 112, 21–29.
- Ahlroth, P., Alatalo, R.V., Holopainen, A., Kumpulainen, T., and Suhonen, J. (2003). Founder population size and number of source populations enhance colonization success in waterstriders. *Oecologia* 137, 617–620.
- Akiner, M.M., Demirci, B., Babuadze, G., Robert, V., and Schaffner, F. (2016). Spread of the Invasive Mosquitoes *Aedes aegypti* and *Aedes albopictus* in the Black Sea Region Increases Risk of Chikungunya, Dengue, and Zika Outbreaks in Europe. *PLOS Neglected Tropical Diseases* 10, e0004664.
- Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E., and Thuiller, W. (2010). Sampling in ecology and evolution - bridging the gap between theory and practice. *Ecography* 33, 1028–1037.
- Albon, S.D., Guinness, F.E., and Clutton-Brock, T.H. (1983). The influence of climatic variation on the birth weights of red deer (*Cervus elaphus*). *Journal of Zoology* 200, 295–298.
- Allendorf, F.W., and Lundquist, L.L. (2003). Introduction: population biology, evolution, and control of invasive species. *Conservation Biology* 17, 24–30.
- Arca, M., Mougel, F., Guillemaud, T., Dupas, S., Rome, Q., Perrard, A., Muller, F., Fossoud, A., Capdevielle-Dulac, C., Torres-Leguizamon, M., *et al.* (2015). Reconstructing the invasion and the demographic history of the yellow-legged hornet, *Vespa velutina*, in Europe. *Biological Invasions* 17, 2357–2371.

- Armbruster, W.S., Pelabon, C., Bolstad, G.H., and Hansen, T.F. (2014). Integrated phenotypes: understanding trait covariation in plants and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 20130245–20130245.
- Arnold, S. J. (1992). Constraints on phenotypic evolution. *American Naturalist*, S85-S107.
- Arnold, S.J., Pfrender, M.E., and Jones, A.G. (2001). The adaptive landscape as a conceptual bridge between micro-and macroevolution. *Genetica* 112, 9–32.
- Arnold, S.J., Bürger, R., Hohenlohe, P.A., Ajie, B.C., and Jones, A.G. (2008). Understanding the evolution and stability of the G-matrix. *Evolution* 62, 2451–2461.
- Ashander, J., Chevin, L.-M., and Baskett, M.L. (2016). Predicting evolutionary rescue via evolving plasticity in stochastic environments. *Proceedings of the Royal Society B: Biological Sciences* 283, 20161690.
- Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., *et al.* (2015). Invasion biology of spotted wing *Drosophila* (*Drosophila suzukii*): a global perspective and future priorities. *Journal of Pest Science* 88, 469–494.
- Atallah, J., Teixeira, L., Salazar, R., Zaragoza, G., and Kopp, A. (2014). The making of a pest: the evolution of a fruit-penetrating ovipositor in *Drosophila suzukii* and related species. *Proceedings of the Royal Society B: Biological Sciences* 281, 20132840–20132840.
- Avery, P.J., and Hill, W.G. (1979). Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. *Genetics* 91, 817–844.
- Avise, J. C. (2012). *Molecular markers, natural history and evolution*. Springer Science & Business Media.
- Azevedo, R.B.R., James, A.C., McCabe, J., and Partridge, L. (1998). Latitudinal Variation of Wing: Thorax Size Ratio and Wing-Aspect Ratio in *Drosophila melanogaster*. *Evolution* 52, 1353.
- Bacigalupe, L.D. (2009). Biological invasions and phenotypic evolution: a quantitative genetic perspective. *Biological Invasions* 11, 2243–2250.
- Bacigalupe, L.D., Barrientos, K., Beckerman, A.P., Carter, M.J., Figueroa, C.C., Foster, S.P., Moore, A.J., Silva, A.X., and Nespolo, R.F. (2013). Can invasions occur without change? A comparison of **G** -matrices and selection in the peach-potato aphid, *Myzus persicae*. *Ecology and Evolution* 3, 5109–5118.
- Baker, H.N., and Stebbins, G.L.,. (1965). *The genetics of colonizing species; proceedings*. New York: Academic Press.
- Barrett, R., and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution* 23, 38–44.
- Barrett, S.C.H., Colautti, R.I., and Eckert, C.G. (2008). Plant reproductive systems and evolution during biological invasion. *Molecular Ecology* 17, 373–383.
- Barton, N.H., and Turelli, M. (1989). Evolutionary quantitative genetics: how little do we know?

Annual Review of Genetics 23, 337–370.

- Basoalto, E., Hilton, R., and Knight, A. (2013). Factors affecting the efficacy of a vinegar trap for *Drosophila suzukii* (Diptera; Drosophilidae). *Journal of Applied Entomology* 137, 561–570.
- Bebber, D.P., Ramotowski, M.A.T., and Gurr, S.J. (2013). Crop pests and pathogens move polewards in a warming world. *Nature Climate Change* 3, 985–988.
- Bégin, M., and Roff, D.A. (2004). From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution* 58, 2287.
- Beldade, P., Mateus, A.R.A., and Keller, R.A. (2011). Evolution and molecular mechanisms of adaptive developmental plasticity: ADAPTIVE DEVELOPMENTAL PLASTICITY. *Molecular Ecology* 20, 1347–1363.
- Bell, G. (2008). *Selection: the mechanism of evolution*. Oxford University Press on Demand.
- Bellen, H.J., Tong, C., and Tsuda, H. (2010). 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. *Nature Reviews Neuroscience* 11, 514–522.
- Birmingham, A.L., Kovacs, E., Lafontaine, J.P., Avelino, N., Borden, J.H., Andreller, I.S., and Gries, G. (2011). A New Trap and Lure for *Drosophila melanogaster* (Diptera: Drosophilidae). *Journal of Economic Entomology* 104, 1018–1023.
- Bitner-Mathé, B. C., & Klaczko, L. B. (1999). Heritability, phenotypic and genetic correlations of size and shape of *Drosophila mediopunctata* wings. *Heredity*, 83(6), 688-696.
- Björklund, M., Husby, A., and Gustafsson, L. (2013). Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. *Journal of Evolutionary Biology* 26, 1–13.
- Blossey, B., and Notzold, R. (1995). Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. *Journal of Ecology* 83, 887–889.
- Blows, M.W. (2007). A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* 20, 1–8.
- Blows, M.W., and Hoffmann, A.A. (2005). A reassessment of genetic limits to evolutionary change. *Ecology* 86, 1371–1384.
- Bock, D.G., Caseys, C., Cousens, R.D., Hahn, M.A., Heredia, S.M., Hübner, S., Turner, K.G., Whitney, K.D., and Rieseberg, L.H. (2015). What we still don't know about invasion genetics. *Molecular Ecology* 24, 2277–2297.
- Bolda, M.P., Goodhue, R.E., and Zalom, F.G. (2010). Spotted wing drosophila: potential economic impact of a newly established pest. *Agricultural and Resource Economics Update* 13, 5–8.
- Bolstad, G.H., Cassara, J.A., Márquez, E., Hansen, T.F., van der Linde, K., Houle, D., and Pélabon, C. (2015). Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences* 112, 13284–13289.
- Bonizzoni, M., Gasperi, G., Chen, X., and James, A.A. (2013). The invasive mosquito species

- Aedes albopictus*: current knowledge and future perspectives. *Trends in Parasitology* 29, 460–468.
- Bookstein, F.L. (1996). Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology* 58, 313–365.
- Bowler PJ (2005) Variation from Darwin to the modern synthesis. In: Variation. A Central Concept in Biology (eds Halgrimsson B, Hall BK), pp. 29–47, Elsevier, Amsterdam
- Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in genetics*, 13, 115-155.
- Bradshaw, C.J.A., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., Barbet-Massin, M., Salles, J.-M., Simard, F., and Courchamp, F. (2016). Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* 7, 12986.
- Brakefield, P.M. (2003). Artificial selection and the development of ecologically relevant phenotypes. *Ecology* 84, 1661–1671.
- Brakefield, P.M. (2006). Evo-devo and constraints on selection. *Trends in Ecology & Evolution* 21, 362–368.
- Brakefield, P.M., Pijpe, J., and Zwaan, B.J. (2007). Developmental plasticity and acclimation both contribute to adaptive responses to alternating seasons of plenty and of stress in *Bicyclus* butterflies. *Journal of Biosciences* 32, 465–475.
- Brodie, E. D., Moore, A. J., & Janzen, F. J. (1995). Visualizing and quantifying natural selection. *Trends in Ecology & Evolution*, 10(8), 313-318.
- Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S., Dekker, C.L., Swan, G.E., Butte, A.J., *et al.* (2015). Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. *Cell* 160, 37–47.
- Broennimann, O., and Guisan, A. (2008). Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters* 4, 585–589.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T., and Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters* 10, 701–709.
- de Buffon, G.L.L.C., and Furne et Compagnie (Paris) (1844). *Oeuvres complètes* (chez Furne et Cie).
- Burrack, H.J., Smith, J.P., Pfeiffer, D.G., Koehler, G., and Laforest, J. (2012). Using Volunteer-Based Networks to Track & *Drosophila suzukii* (Diptera: Drosophilidae) an Invasive Pest of Fruit Crops. *Journal of Integrated Pest Management* 3, 1–5.
- Burrack, H.J., Fernandez, G.E., Spivey, T., and Kraus, D.A. (2013). Variation in selection and utilization of host crops in the field and laboratory by *Drosophila suzukii* Matsumara (Diptera: Drosophilidae), an invasive frugivore: Selection and utilization of host crops by the invasive frugivore *D. suzukii*. *Pest Management Science* 69, 1173–1180.

- Calabria, G., Máca, J., Bächli, G., Serra, L., and Pascual, M. (2012). First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe: First record of *Drosophila suzukii* in Europe. *Journal of Applied Entomology* 136, 139–147.
- Calsbeek, B., Lavergne, S., Patel, M., and Molofsky, J. (2011). Comparing the genetic architecture and potential response to selection of invasive and native populations of reed canary grass: Genetic architecture of invasive populations. *Evolutionary Applications* 4, 726–735.
- Cano, J.M., Laurila, A., Palo, J., and Merilä, J. (2004). Population differentiation in G matrix structure due to natural selection in *Rana temporaria*. *Evolution* 58, 2013–2020.
- Careau, V., Wolak, M.E., Carter, P.A., and Garland, T. (2015). Evolution of the additive genetic variance–covariance matrix under continuous directional selection on a complex behavioural phenotype. *Proceedings of the Royal Society B: Biological Sciences* 282, 20151119.
- Carlton, J. (2003). *Invasive species: vectors and management strategies*. Island Press.
- Carrière, Y., Roff, D.A., and Deland, J.-P. (1995). The Joint Evolution of Diapause and Insecticide Resistance: A Test of an Optimality Model. *Ecology* 76, 1497–1505.
- Cattel, J., Martinez, J., Jiggins, F., Mouton, L., and Gibert, P. (2016a). *Wolbachia* -mediated protection against viruses in the invasive pest *Drosophila suzukii*: *Wolbachia and antiviral protection in D. suzukii*. *Insect Molecular Biology* 25, 595–603.
- Cattel, J., Kaur, R., Gibert, P., Martinez, J., Fraimout, A., Jiggins, F., Andrieux, T., Siozios, S., Anfora, G., Miller, W., *et al.* (2016b). *Wolbachia* in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies. *PLOS ONE* 11, e0147766.
- Cesari, M., Maistrello, L., Ganzerli, F., Dioli, P., Rebecchi, L., and Guidetti, R. (2015). A pest alien invasion in progress: potential pathways of origin of the brown marmorated stink bug *Halyomorpha halys* populations in Italy. *Journal of Pest Science* 88, 1–7.
- Cha, D.H., Adams, T., Rogg, H., and Landolt, P.J. (2012). Identification and Field Evaluation of Fermentation Volatiles from Wine and Vinegar that Mediate Attraction of Spotted Wing *Drosophila*, *Drosophila suzukii*. *Journal of Chemical Ecology* 38, 1419–1431.
- Chabert, S., Allemand, R., Poyet, M., Eslin, P., and Gibert, P. (2012). Ability of European parasitoids (Hymenoptera) to control a new invasive Asiatic pest, *Drosophila suzukii*. *Biological Control* 63, 40–47.
- Chantepie, S., Robert, A., Sorci, G., Hingrat, Y., Charmantier, A., Leveque, G., Lacroix, F., and Teplitsky, C. (2015). Quantitative Genetics of the Aging of Reproductive Traits in the Houbara Bustard. *PLOS ONE* 10, e0133140.
- Charmantier, A., and Garant, D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences* 272, 1415–1425.

- Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B., and Sheldon, B.C. (2008). Adaptive Phenotypic Plasticity in Response to Climate Change in a Wild Bird Population. *Science* 320, 800–803.
- Cheverud, J. M. (1996). Developmental integration and the evolution of pleiotropy. *American Zoologist*, 36(1), 44-50.
- Chevin, L.-M., Lande, R., and Mace, G.M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLoS Biology* 8, e1000357.
- Chiu, J.C., Jiang, X., Zhao, L., Hamm, C.A., Cridland, J.M., Saelao, P., Hamby, K.A., Lee, E.K., Kwok, R.S., Zhang, G., *et al.* (2013). Genome of *Drosophila suzukii*, the Spotted Wing *Drosophila*. *G3; Genes|Genomes|Genetics* 3, 2257–2271.
- Cini, A., Ioriatti, C., Anfora, G., and others (2012). A review of the invasion of *Drosophila suzukii* in Europe and a draft research agenda for integrated pest management. *Bulletin of Insectology* 65, 149–160.
- Cini, A., Anfora, G., Escudero-Colomar, L.A., Grassi, A., Santosuosso, U., Seljak, G., and Papini, A. (2014). Tracking the invasion of the alien fruit pest *Drosophila suzukii* in Europe. *Journal of Pest Science* 87, 559–566.
- Clavero, M., Brotons, L., Pons, P., and Sol, D. (2009). Prominent role of invasive species in avian biodiversity loss. *Biological Conservation* 142, 2043–2049.
- Colautti, R.I., and Barrett, S.C.H. (2013). Rapid Adaptation to Climate Facilitates Range Expansion of an Invasive Plant. *Science* 342, 364–366.
- Colautti, R.I., and Lau, J.A. (2015). Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology* 24, 1999–2017.
- Colautti, R.I., and MacIsaac, H.J. (2004). A neutral terminology to define “invasive” species. *Diversity and Distributions* 10, 135–141.
- Colautti, R.I., Eckert, C.G., and Barrett, S.C.H. (2010). Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. *Proceedings of the Royal Society B: Biological Sciences* 277, 1799–1806.
- Coltman, D.W., Pilkington, J., Kruuk, L.E.B., Wilson, K., and Pemberton, J.M. (2001). Positive genetic correlation between parasite resistance and body size in a free-living ungulate population. *Evolution* 55, 2116–2125.
- Conner, J. K., & Hartl, D. L. (2004). A primer of ecological genetics. Sinauer Associates Incorporated.
- Crawley MJ (1997) *Plant Ecology*. Blackwell Science, Oxford
- Crispo, E. (2007). The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* 61, 2469–2479.
- Cristescu, M.E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology* 24, 2212–2225.

- Cuthbertson, A., Collins, D., Blackburn, L., Audsley, N., and Bell, H. (2014). Preliminary Screening of Potential Control Products against *Drosophila suzukii*. *Insects* 5, 488–498.
- Dalton, D.T., Walton, V.M., Shearer, P.W., Walsh, D.B., Caprile, J., and Isaacs, R. (2011). Laboratory survival of *Drosophila suzukii* under simulated winter conditions of the Pacific Northwest and seasonal field trapping in five primary regions of small and stone fruit production in the United States. *Pest Management Science* 67, 1368–1374.
- Darwin, Ch. 1859. *The Origin of Species by Means of Natural Selection*. Murray, London. [1872, 6th edn. Murray, London]
- Darwin, Ch. 1868. *The Variation of Animals and Plants under Domestication*, 2 vols. Murray, London
- David P., Samadi Sarah. *La théorie de l'évolution : une logique pour la biologie*. Paris : Flammarion, 2011, 354 p. (Champs Sciences). ISBN 978-2-0812-6597-4
- David, J.R., Gibert, P., Legout, H., Pétavy, G., Capy, P., and Moreteau, B. (2005). Isofemale lines in *Drosophila*: an empirical approach to quantitative trait analysis in natural populations. *Heredity* 94, 3–12.
- Davidson, A.M., Jennions, M., and Nicotra, A.B. (2011). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis: Invasive species have higher phenotypic plasticity. *Ecology Letters* 14, 419–431.
- De Carvalho, D., Ingvarsson, P.K., Joseph, J., Suter, L., Sedivy, C., Macaya-Sanz, D., Cottrell, J., Heinze, B., Schanzer, I., and Lexer, C. (2010). Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology* 19, 1638–1650.
- De Jong, G. (2005). Evolution of phenotypic plasticity: patterns of plasticity and the emergence of ecotypes: Research review. *New Phytologist* 166, 101–118.
- De los Santo Ramos, M., Bello, A., Gómez, R.J., Hernández, R., and Iribe, T.J. (2014). Monitoring of spotted wing drosophila (*Drosophila suzukii* Mats.) and assessment of the new attractant SuzukiiTrap® in Tijuana, Baja California, Mexico. *Journal of Food, Agriculture and Environment* 12, 349–355.
- Debat, V., and David, P. (2001). Mapping phenotypes: canalization, plasticity and developmental stability. *Trends in Ecology & Evolution* 16, 555–561.
- Deprá, M., Poppe, J.L., Schmitz, H.J., De Toni, D.C., and Valente, V.L.S. (2014). The first records of the invasive pest *Drosophila suzukii* in the South American continent. *Journal of Pest Science* 87, 379–383.
- Des Marais, D.L., Hernandez, K.M., and Juenger, T.E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics* 44, 5–29.
- Dlugosch, K.M., and Parker, I.M. (2008). Founding events in species invasions: genetic variation,

- adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17, 431–449.
- Dlugosch, K.M., Anderson, S.R., Braasch, J., Cang, F.A., and Gillette, H.D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology* 24, 2095–2111.
- Doherty, T.S., Glen, A.S., Nimmo, D.G., Ritchie, E.G., and Dickman, C.R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences* 113, 11261–11265.
- Doroszuk, A., Wojewodzic, M.W., Gort, G., and Kammenga, J.E. (2008). Rapid Divergence of Genetic Variance-Covariance Matrix within a Natural Population. *The American Naturalist* 171, 291–304.
- Drake, J.M., and Lodge, D.M. (2004). Global hot spots of biological invasions: evaluating options for ballast–water management. *Proceedings of the Royal Society of London B: Biological Sciences* 271, 575–580.
- Dryden, I. L., & Mardia, K. V. (1998). *Statistical analysis of shape*. Wiley.
- Duffy, D.L., Montgomery, G.W., Chen, W., Zhao, Z.Z., Le, L., James, M.R., Hayward, N.K., Martin, N.G., and Sturm, R.A. (2007). A three–single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. *The American Journal of Human Genetics* 80, 241–252.
- Endler, J. A. (1986). *Natural selection in the wild* (No. 21). Princeton University Press.
- Eroukhmanoff, F. (2009). Just How Much is the G-matrix Actually Constraining Adaptation? *Evolutionary Biology* 36, 323–326.
- Eroukhmanoff, F., and Svensson, E.I. (2011). Evolution and stability of the G-matrix during the colonization of a novel environment: G-matrix evolution. *Journal of Evolutionary Biology* 24, 1363–1373.
- Erwin, D.H. (2000). Macroevolution is more than repeated rounds of microevolution. *Evolution & Development* 2, 78–84.
- Estoup, A., and Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: why, how and so what?. *Molecular Ecology* 19, 4113–4130.
- Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., and Facon, B. (2016). Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics* 47.
- Facon, B., Genton, B., Shykoff, J., Jarne, P., Estoup, A., and David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* 21, 130–135.
- Facon, B., Pointier, J.-P., Jarne, P., Sarda, V., and David, P. (2008). High Genetic Variance in Life-History Strategies within Invasive Populations by Way of Multiple Introductions. *Current Biology* 18, 363–367.

- Facon, B., Hufbauer, R.A., Tayeh, A., Loiseau, A., Lombaert, E., Vitalis, R., Guillemaud, T., Lundgren, J.G., and Estoup, A. (2011). Inbreeding Depression Is Purged in the Invasive Insect *Harmonia axyridis*. *Current Biology* 21, 424–427.
- Falconer, D. S. (1960). The genetics of litter size in mice. *Journal of cellular and comparative physiology*, 56(S1), 153-167.
- Falconer, D. S., & Mackay, T. F. (1996). *Introduction to quantitative genetics*.
- Fischer, E.K., Ghalambor, C.K., and Hoke, K.L. (2016). Plasticity and evolution in correlated suites of traits. *Journal of Evolutionary Biology* 29, 991–1002.
- Fisher, R. (1930). *The Genetical Theory of Natural Selection*. Oxford, UK: Oxford University Press.
- Forsman, A. (2014). Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences* 111, 302–307.
- Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D.W., Haywood, J., Lean, J., Lowe, D.C., Myhre, G., *et al.* (2007). Changes in atmospheric constituents and in radiative forcing. Chapter 2. In *Climate Change 2007. The Physical Science Basis*, p.
- Forsyth, D.M., and Duncan, R.P. (2001). Propagule Size and the Relative Success of Exotic Ungulate and Bird Introductions to New Zealand. *The American Naturalist* 157, 583–595.
- Fraimout, A., Loiseau, A., Price, D.K., Xuereb, A., Martin, J.-F., Vitalis, R., Fellous, S., Debat, V., and Estoup, A. (2015). New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (Diptera: Drosophilidae): A promising molecular tool for inferring the invasion history of this major insect pest. *European Journal of Entomology*.
- Fraimout, A., Debat, V., Fellous, S., Hufbauer, R., Foucaud, J., Pudlo, P., Marin, J.-M., Price, D.K., Cattell, J., Chen, X., Deprá, M., Duyck, P.-F., Guedot, C., Loeb, G., Kenis, M., Kimura, M.T., Loiseau, A., Martinez-Sañudo, I., Pascual, M., Polihronakis Richmond, M., Shearer, P., Singh, N., Tamura, T., Xuéreb, A., Zhang, J. & Estoup, A. *Mol. Biol. Evol.* submitted
- Fry, S.N. (2005). The aerodynamics of hovering flight in *Drosophila*. *Journal of Experimental Biology* 208, 2303–2318.
- Futuyma, D.J. (2010). EVOLUTIONARY CONSTRAINT AND ECOLOGICAL CONSEQUENCES. *Evolution* 64, 1865–1884.
- Futuyma, D.J., Keese, M.C., and Scheffer, S.J. (1993). Genetic Constraints and the Phylogeny of Insect-Plant Associations: Responses of *Ophraella communa* (Coleoptera: Chrysomelidae) to Host Plants of its Congeners. *Evolution* 47, 888.
- Geng, Y., van Klinken, R.D., Sosa, A., Li, B., Chen, J., and Xu, C.-Y. (2016). The Relative Importance of Genetic Diversity and Phenotypic Plasticity in Determining Invasion Success of a Clonal Weed in the USA and China. *Frontiers in Plant Science* 7.
- Ghalambor, C.K., McKAY, J.K., Carroll, S.P., and Reznick, D.N. (2007). Adaptive versus non-

- adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21, 394–407.
- Gibert, P., Moreteau, B., & David, J. R. (2000). Developmental constraints on an adaptive plasticity: reaction norms of pigmentation in adult segments of *Drosophila melanogaster*. *Evolution & development*, 2(5), 249-260.
- Gibert, P., Hill, M., Pascual, M., Plantamp, C., Terblanche, J.S., Yassin, A., and Sgrò, C.M. (2016). *Drosophila* as models to understand the adaptive process during invasion. *Biological Invasions* 18, 1089–1103.
- Gibson, G., and Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics* 5, 681–690.
- Gidaszewski, N.A., Baylac, M., and Klingenberg, C. (2009). Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. *BMC Evolutionary Biology* 9, 110.
- Gilchrist, A.S., Azevedo, R.B.R., Partridge, L., and O'higgins, P. (2000). Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. *Evolution & Development* 2, 114–124.
- Gilchrist, G.W., Huey, R.B., Balanyà, J., Pascual, M., and Serra, L. (2004). A time series of evolution in action: a latitudinal cline in wing size in south american *Drosophila subobscura*. *Evolution* 58, 768.
- Goodhue, R.E., Bolda, M., Farnsworth, D., Williams, J.C., and Zalom, F.G. (2011). Spotted wing drosophila infestation of California strawberries and raspberries: economic analysis of potential revenue losses and control costs. *Pest Management Science* 67, 1396–1402.
- Grasswitz, T.R. (2015). First Occurrence of Spotted Wing *Drosophila*, *Drosophila suzukii*¹, in New Mexico, with Notes on Its Host Plant Associations. *Southwestern Entomologist* 40, 851–853.
- Guillaume, F., & Whitlock, M. C. (2007). Effects of migration on the genetic covariance matrix. *Evolution*, 61(10), 2398-2409.
- Haas, H.L., and Tolley, K.A. (1998). Geographic variation of wing morphology in three Eurasian populations of the fruit fly, *Drosophila lummei*. *Journal of Zoology* 245, 197–203.
- Hahn, M.A., van Kleunen, M., and Müller-Schärer, H. (2012). Increased Phenotypic Plasticity to Climate May Have Boosted the Invasion Success of Polyploid *Centaurea stoebe*. *PLoS ONE* 7, e50284.
- Hamby, K.A., Hernandez, A., Boundy-Mills, K., and Zalom, F.G. (2012). Associations of Yeasts with Spotted-Wing *Drosophila* (*Drosophila suzukii*; Diptera: Drosophilidae) in Cherries and Raspberries. *Applied and Environmental Microbiology* 78, 4869–4873.
- Hamby, K.A., Kwok, R.S., Zalom, F.G., and Chiu, J.C. (2013). Integrating Circadian Activity and Gene Expression Profiles to Predict Chronotoxicity of *Drosophila suzukii* Response to Insecticides. *PLoS ONE* 8, e68472.

- Hansen, T.F., and Houle, D. (2004). Evolvability, stabilizing selection, and the problem of stasis. In Pigliucci & Preston. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford University Press (2004)
- Hansen, T.F., and Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* 21, 1201–1219.
- Hartl, D.L., Clark, A.G., and Clark, A.G. (1997). *Principles of population genetics* (Sinauer associates Sunderland).
- Hauser, M. (2011). A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. *Pest Management Science* 67, 1352–1357.
- Hayes, K.R., and Barry, S.C. (2008). Are there any consistent predictors of invasion success? *Biological Invasions* 10, 483–506.
- Herborg, L.-M., Jerde, C.L., Lodge, D.M., Ruiz, G.M., and MacIsaac, H.J. (2007). Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications* 17, 663–674.
- Hill, W.G., Goddard, M.E., and Visscher, P.M. (2008). Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics* 4, e1000008.
- Hine, E. (2006). Determining the Effective Dimensionality of the Genetic Variance-Covariance Matrix. *Genetics* 173, 1135–1144.
- Hoffmann, A. A., & Merilä, J. (1999). Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology & Evolution*, 14(3), 96-101.
- Hoffmann, A.A., and Hercus, M.J. (2000). Environmental stress as an evolutionary force. *Bioscience* 50, 217–226.
- Hoffmann, A.A., and Parsons, P.A. (1988). The analysis of quantitative variation in natural populations with isofemale strains. *Genetique, Selection, Evolution* 20, 1.
- Hoffmann, A.A., Merilä, J., and Kristensen, T.N. (2016). Heritability and evolvability of fitness and nonfitness traits: Lessons from livestock: HERITABILITY AND EVOLVABILITY IN LIVESTOCK. *Evolution* 70, 1770–1779.
- Hohenlohe, P.A., and Arnold, S.J. (2008). MIPoD: A Hypothesis-Testing Framework for Microevolutionary Inference from Patterns of Divergence. *The American Naturalist* 171, 366–385.
- Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* 130, 195–204.
- Houle, D., Mezey, J., Galpern, P., and Carter, A. (2003). Automated measurement of *Drosophila* wings. *BMC Evolutionary Biology* 3, 1.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157, 1262–1278.

- Huey, R.B., Gilchrist, G.W., Carlson, M.L., Berrigan, D., and Serra, L. (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287, 308–309.
- Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O., and Estoup, A. (2012). Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions: Anthropogenically induced adaptation to invade. *Evolutionary Applications* 5, 89–101.
- Hulme, P.E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* 46, 10–18.
- Husband, B. C., & Barrett, S. C. (1993). Multiple origins of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae): inferences from style morph and isozyme variation. *Journal of Evolutionary Biology*, 6(4), 591-608.
- Hutchinson, G.E. (1957). Population studies: Animal ecology and demography. *Bulletin of Mathematical Biology* 53, 193–213.
- Jamniczky, H.A., and Hallgrímsson, B. (2009). A comparison of covariance structure in wild and laboratory murid crania. *Evolution* 63, 1540–1556.
- Jaramillo, S.L., Mehlferber, E., and Moore, P.J. (2015). Life-history trade-offs under different larval diets in *Drosophila suzukii* (Diptera: Drosophilidae). *Physiological Entomology* 40, 2–9.
- Johansson, F., Lind, M.I., Ingvarsson, P.K., and Bokma, F. (2012). Evolution of the G-matrix in life history traits in the common frog during a recent colonisation of an island system. *Evolutionary Ecology* 26, 863–878.
- Jones, A.G., Arnold, S.J., and Bürger, R. (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* 57, 1747–1760.
- Jones, A.G., Arnold, S.J., and Burger, R. (2004). Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* 58, 1639.
- Kaneshiro KY. 1983. *Drosophila* (Sophophora) *suzukii* (Matsumura). *Proc Hawaiian Entomol Soc.* 24:179.
- Keeseey, I.W., Knaden, M., and Hansson, B.S. (2015). Olfactory Specialization in *Drosophila suzukii* Supports an Ecological Shift in Host Preference from Rotten to Fresh Fruit. *Journal of Chemical Ecology* 41, 121–128.
- Keller, S.R., and Taylor, D.R. (2008). History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* 11, 852–866.
- Kilkenny, F.F., and Galloway, L.F. (2013). Adaptive divergence at the margin of an invaded range: adaptive divergence in margin populations.. *Evolution* 67, 722–731.
- Kimura, M.T. (2004). Cold and heat tolerance of drosophilid flies with reference to their latitudinal distributions. *Oecologia* 140, 442–449.

- Kinjo, H., Kunimi, Y., and Nakai, M. (2014). Effects of temperature on the reproduction and development of *Drosophila suzukii* (Diptera: Drosophilidae). *Applied Entomology and Zoology* 49, 297–304.
- Klingenberg, C.P. (2003). Quantitative genetics of geometric shape: heritability and the pitfalls of the univariate approach. *Evolution* 57, 191–195.
- Klingenberg, C.P. (2010). Evolution and development of shape: integrating quantitative approaches. *Nature Reviews Genetics*.
- Klingenberg, C.P., and Leamy, L.J. (2001). Quantitative genetics of geometric shape in the mouse mandible. *Evolution* 55, 2342–2352.
- Klingenberg, C.P., Mebus, K., and Auffray, J.-C. (2003). Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evolution & Development* 5, 522–531.
- Klingenberg, C.P., Leamy, L.J., and Cheverud, J.M. (2004). Integration and modularity of quantitative trait locus effects on geometric shape in the mouse mandible. *Genetics* 166, 1909–1921.
- Kolar, C.S., and Lodge, D.M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* 16, 199–204.
- Kolbe, J.J., Larson, A., Losos, J.B., and de Queiroz, K. (2008). Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biology Letters* 4, 434–437.
- Kopp, A., and True, J.R. (2002). Evolution of male sexual characters in the oriental *Drosophila melanogaster* species group. *Evolution & Development* 4, 278–291.
- Krebs, R.A., and Loeschcke, V. (1997). Estimating heritability in a threshold trait: heat-shock tolerance in *Drosophila buzzatii*. *Heredity* 79.
- Lande, R. (1979). Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. *Evolution* 33, 402.
- Lande, R. (1980). Genetic variation and phenotypic evolution during allopatric speciation. *American Naturalist* 463–479.
- Lande, R. (1981). Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences* 78, 3721–3725.
- Lande, R. (2015). Evolution of phenotypic plasticity in colonizing species. *Molecular Ecology* 24, 2038–2045.
- Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, 1210-1226.
- Landolt, P.J., Adams, T., and Rogg, H. (2012). Trapping spotted wing drosophila, *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae), with combinations of vinegar and wine, and acetic acid and ethanol: Trapping spotted wing drosophila. *Journal of Applied Entomology*

136, 148–154.

- Lavergne, S., and Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences* 104, 3883–3888.
- Leamy, L.J., and Klingenberg, C.P. (2005). The Genetics and Evolution of Fluctuating Asymmetry. *Annual Review of Ecology, Evolution, and Systematics* 36, 1–21.
- Leary, R.F., and Allendorf, F.W. (1989). Fluctuating asymmetry as an indicator of stress: implications for conservation biology. *Trends in Ecology & Evolution* 4, 214–217.
- Ledon-Rettig, C.C., Pfennig, D.W., and Crespi, E.J. (2010). Diet and hormonal manipulation reveal cryptic genetic variation: implications for the evolution of novel feeding strategies. *Proceedings of the Royal Society B: Biological Sciences* 277, 3569–3578.
- Lee, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* 17, 386–391.
- Lee, C.E., and Gelembiuk, G.W. (2008). Evolutionary origins of invasive populations. *Evolutionary Applications* 1, 427–448.
- Lee, C.E., Remfert, J.L., and Chang, Y.-M. (2007). Response to selection and evolvability of invasive populations. *Genetica* 129, 179–192.
- Lee, J.C., Bruck, D.J., Curry, H., Edwards, D., Haviland, D.R., Van Steenwyk, R.A., and Yorgey, B.M. (2011). The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. *Pest Management Science* 67, 1358–1367.
- Leinonen, T., Cano, J.M., and Merilä, J. (2011). Genetics of body shape and armour variation in threespine sticklebacks: Genetics of threespine stickleback body shape. *Journal of Evolutionary Biology* 24, 206–218.
- Leprieur, F., Beauchard, O., Blanchet, S., Oberdorff, T., and Brosse, S. (2008). Fish Invasions in the World's River Systems: When Natural Processes Are Blurred by Human Activities. *PLoS Biology* 6, e28.
- Lester, P.J., Gruber, M.A.M., Brenton-Rule, E.C., Archer, M., Corley, J.C., Dvořák, L., Masciocchi, M., and Van Oystaeyen, A. (2014). Determining the origin of invasions and demonstrating a lack of enemy release from microsporidian pathogens in common wasps (*Vespula vulgaris*). *Diversity and Distributions* 20, 964–974.
- Levis, N.A., and Pfennig, D.W. (2016). Evaluating “Plasticity-First” Evolution in Nature: Key Criteria and Empirical Approaches. *Trends in Ecology & Evolution* 31, 563–574.
- Levis, N.A., Schooler, M.L., Johnson, J.R., and Collyer, M.L. (2016). Non-adaptive phenotypic plasticity: the effects of terrestrial and aquatic herbicides on larval salamander morphology and swim speed. *Biological Journal of the Linnean Society*.
- Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics* 1–18.
- Linnen, C.R., Kingsley, E.P., Jensen, J.D., and Hoekstra, H.E. (2009). On the Origin and Spread of

- an Adaptive Allele in Deer Mice. *Science* 325, 1095–1098.
- Liu, H., and Stiling, P. (2006). Testing the enemy release hypothesis: a review and meta-analysis. *Biological Invasions* 8, 1535–1545.
- Llewellyn, D., Thompson, M.B., Brown, G.P., Phillips, B.L., and Shine, R. (2012). Reduced investment in immune function in invasion-front populations of the cane toad (*Rhinella marina*) in Australia. *Biological Invasions* 14, 999–1008.
- Lockwood, J.L., Cassey, P., and Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution* 20, 223–228.
- Lombaert, E., Guillemaud, T., Cornuet, J.-M., Malausa, T., Facon, B., and Estoup, A. (2010). Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *PLoS ONE* 5, e9743.
- Lombaert, E., Guillemaud, T., Lundgren, J., Koch, R., Facon, B., Grez, A., Loomans, A., Malausa, T., Nedved, O., Rhule, E., *et al.* (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology* 23, 5979–5997.
- Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Vol. 1). Sunderland, MA: Sinauer.
- Matsumura S. 1931: 6000 illustrated insects of Japan-empire (in Japanese). Toko shoin, Tokyo, 1497 pp.
- Maier, C.T. (2012). First Detection and Widespread Distribution of the Spotted Wing *Drosophila*, *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae), in Connecticut in 2011. *Proceedings of the Entomological Society of Washington* 114, 329–337.
- Marroig, G., and Cheverud, J.M. (2005). Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. *Evolution* 59, 1128–1142.
- Martin, A., and Orgogozo, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation: perspective. *Evolution*.
- Martínez-Abadías, N., González-José, R., González-Martín, A., Van der Molen, S., Talavera, A., Hernández, P., and Hernández, M. (2006). Phenotypic evolution of human craniofacial morphology after admixture: A geometric morphometrics approach. *American Journal of Physical Anthropology* 129, 387–398.
- Martínez-Abadías, N., Esparza, M., Sjøvold, T., González-José, R., Santos, M., Hernández, M., and Klingenberg, C.P. (2012). Pervasive genetic integration directs the evolution of human skull shape: genetic integration in the human skull. *Evolution* 66, 1010–1023.
- Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, 536, 143-145.
- Mayr, E. (1959). Darwin and the evolutionary theory in Biology. Pages 1–10 in *Evolution and Anthropology: A Centennial Appraisal* (B J Meggers, Ed) The Anthropological Society of

Washington, Washington DC.

- Mayr, E. (1970). *Populations, Species, and Evolution*. The Belknap Press of Harvard University Press.
- McGuigan, K. (2006). Studying phenotypic evolution using multivariate quantitative genetics: Evolutionary quantitative genetic. *Molecular Ecology* 15, 883–896.
- McKay, J. K., & Latta, R. G. (2002). Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, 17(6), 285-291.
- Medley, K.A. (2010). Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models: Niche shifts and global invasion. *Global Ecology and Biogeography* 19, 122–133.
- Merilä, J., and Crnokrak, P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits: Natural selection and genetic differentiation. *Journal of Evolutionary Biology* 14, 892–903.
- Merilä, J., and Hendry, A.P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* 7, 1–14.
- Merilä, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity*, 83(2), 103-109.
- Merilä, J., Sheldon, B.C., and Kruuk, L.E.B. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica* 112, 199–222.
- Mezey, J.G., and Houle, D. (2005). The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* 59, 1027–1038.
- Michaud, H. (1941). Importance of Field Work for the High School Biology Teacher. *The American Biology Teacher* 3, 205–208.
- Miller, J. (2010). Species Distribution Modeling: Species distribution modeling. *Geography Compass* 4, 490–509.
- Mitsui, H., Takahashi, K.H., and Kimura, M.T. (2006). Spatial distributions and clutch sizes of *Drosophila* species ovipositing on cherry fruits of different stages. *Population Ecology* 48, 233–237.
- Mitsui, H., Beppu, K., and Kimura, M.T. (2010). Seasonal life cycles and resource uses of flower- and fruit-feeding drosophilid flies (Diptera: Drosophilidae) in central Japan. *Entomological Science* 13, 60–67.
- Mitteroecker, P., and Gunz, P. (2009). Advances in Geometric Morphometrics. *Evolutionary Biology* 36, 235–247.
- Monastiriotti, M., Linn Jr, C.E., and White, K. (1996). Characterization of *Drosophila* tyramine β -hydroxylase gene and isolation of mutant flies lacking octopamine. *The Journal of Neuroscience* 16, 3900–3911.
- Montague, J.L., Barrett, S.C.H., and Eckert, C.G. (2007). Re-establishment of clinal variation in

- flowering time among introduced populations of purple loosestrife (*Lythrum salicaria*, Lythraceae). *Journal of Evolutionary Biology* 0, 071119204355004–???
- Mousseau, T. A., & Roff, D. A. (1987). Natural selection and the heritability of fitness components. *Heredity*, 59(Pt 2), 181-197.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., and Worm, B. (2011). How Many Species Are There on Earth and in the Ocean? *PLoS Biology* 9, e1001127.
- Mori, B.A., Whitener, A.B., Leinweber, Y., Revadi, S., Beers, E.H., Witzgall, P., and Becher, P.G. (2016). Enhanced yeast feeding following mating facilitates control of the invasive fruit pest *Drosophila suzukii*. *Journal of Applied Ecology*.
- Morrissey, M.B., Walling, C.A., Wilson, A.J., Pemberton, J.M., Clutton-Brock, T.H., and Kruuk, L.E.B. (2012). Genetic Analysis of Life-History Constraint and Evolution in a Wild Ungulate Population. *The American Naturalist* 179, E97–E114.
- Munday, P.L., Warner, R.R., Monroe, K., Pandolfi, J.M., and Marshall, D.J. (2013). Predicting evolutionary responses to climate change in the sea. *Ecology Letters* 16, 1488–1500.
- Nei, M. (1975). *Molecular population genetics and evolution* (Amsterdam : New York: North-Holland Pub. Co. ; American Elsevier Pub. Co).
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583–590.
- O’Dowd, D.J., Green, P.T., and Lake, P.S. (2003). Invasional “meltdown” on an oceanic island. *Ecology Letters* 6, 812–817.
- Olivieri, I. (2009). Alternative mechanisms of range expansion are associated with different changes of evolutionary potential. *Trends in Ecology & Evolution* 24, 289–292.
- Ometto, L., Cestaro, A., Ramasamy, S., Grassi, A., Revadi, S., Siozios, S., Moretto, M., Fontana, P., Varotto, C., Pisani, D., *et al.* (2013). Linking Genomics and Ecology to Investigate the Complex Evolution of an Invasive *Drosophila* Pest. *Genome Biology and Evolution* 5, 745–757.
- Orhan, A., Aslantaş, R., öNder, B.Ş., and Tozlu, G. (2016). First record of the invasive vinegar fly *Drosophila suzukii* (Matsumura)(Diptera: Drosophilidae) from eastern Turkey. *TURKISH JOURNAL OF ZOOLOGY* 40, 290–293.
- Orr, H.A. (2005). The genetic theory of adaptation: a brief history. *Nature Reviews Genetics* 6, 119–127.
- Orr, H.A., and Betancourt, A.J. (2001). Haldane’s sieve and adaptation from the standing genetic variation. *Genetics* 157, 875–884.
- Paccard, A., Van Buskirk, J., Willi, Y., Eckert, C.G., and Bronstein, J.L. (2016). Quantitative Genetic Architecture at Latitudinal Range Boundaries: Reduced Variation but Higher Trait Independence. *The American Naturalist* 187, 667–677.
- Paini, D.R., Sheppard, A.W., Cook, D.C., De Barro, P.J., Worner, S.P., and Thomas, M.B. (2016).

- Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences* *113*, 7575–7579.
- Parker, J.D., Torchin, M.E., Hufbauer, R.A., Lemoine, N.P., Alba, C., Blumenthal, D.M., Bossdorf, O., Byers, J.E., Dunn, A.M., Heckman, R.W., *et al.* (2013). Do invasive species perform better in their new ranges? *Ecology* *94*, 985–994.
- Parsons, P.A. (1990). Fluctuating asymmetry: an epigenetic measure of stress. *Biological Reviews* *65*, 131–145.
- Pavlicev, M., Norgard, E.A., Fawcett, G.L., and Cheverud, J.M. (2011). Evolution of pleiotropy: epistatic interaction pattern supports a mechanistic model underlying variation in genotype-phenotype map. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* *316B*, 371–385.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., ... & Coops, N. C. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277-278.
- Perrings, C., Dehnen-Schmutz, K., Touza, J., and Williamson, M. (2005). How to manage biological invasions under globalization. *Trends in Ecology & Evolution* *20*, 212–215.
- Peterson, A.T. (2003). Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology* *78*, 419–433.
- Pfennig, D.W., Rice, A.M., and Martin, R.A. (2006). Ecological opportunity and phenotypic plasticity interact to promote character displacement and species coexistence. *Ecology* *87*, 769–779.
- Phillips, P. C., & Arnold, S. J. (1989). Visualizing multivariate selection. *Evolution*, 1209-1222.
- Phillips, P. C., & Arnold, S. J. (1999). Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution*, 1506-1515.
- Phillips, P.C., Whitlock, M.C., and Fowler, K. (2001). Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* *158*, 1137–1145.
- Pigliucci, M. (2005). Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* *20*, 481–486.
- Pigliucci, M. (2009). An Extended Synthesis for Evolutionary Biology. *Annals of the New York Academy of Sciences* *1168*, 218–228.
- Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., *et al.* (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment* *84*, 1–20.
- Plantamp, C., Salort, K., Gibert, P., Dumet, A., Mialdea, G., Mondy, N., and Voituron, Y. (2016). All or nothing: Survival, reproduction and oxidative balance in Spotted Wing *Drosophila* (*Drosophila suzukii*) in response to cold. *Journal of Insect Physiology* *89*, 28–36.
- Polly, P.D. (2008). Developmental Dynamics and G-Matrices: Can Morphometric Spaces be Used to Model Phenotypic Evolution? *Evolutionary Biology* *35*, 83–96.
- Poyet, M., Eslin, P., Héraude, M., Le Roux, V., Prévost, G., Gibert, P., and Chabrierie, O. (2014).

Invasive host for invasive pest: when the Asiatic cherry fly (*Drosophila suzukii*) meets the American black cherry (*Prunus serotina*) in Europe: When *D. suzukii* meets *P. serotina*. *Agricultural and Forest Entomology* 16, 251–259.

- Price, T., and Schuller, D. (1991). On the Low Heritability of Life-History Traits. *Evolution* 45, 853.
- Price, T.D., Qvarnström, A., and Irwin, D.E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 1433–1440.
- Prout, T., and Barker, J.S. (1989). Ecological aspects of the heritability of body size in *Drosophila buzzatii*. *Genetics* 123, 803–813.
- Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., and Robert, C.P. (2016). Reliable ABC model choice via random forests. *Bioinformatics* 32, 859–866.
- Puentes, A., Granath, G., and Ågren, J. (2016). Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*: G MATRIX STRUCTURE IN POPULATIONS OF *A. lyrata*. *Evolution* 70, 2370–2386.
- Puillandre, N., Dupas, S., Dangles, O., Zeddani, J.-L., Capdevielle-Dulac, C., Barbin, K., Torres-Leguizamon, M., and Silvain, J.-F. (2008). Genetic bottleneck in invasive species: the potato tuber moth adds to the list. *Biological Invasions* 10, 319–333.
- Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K., and Whittam, T.S. (2000). Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406, 64–67.
- Rejmánek, M., Richardson, D.M., Barbour, M.G., Crawley, M.J., Hrusa, G.F., Moyle, P.B., Randall, J.M., Simberloff, D., and Williamson, M. (2002). Biological invasions: politics and the discontinuity of ecological terminology. *Bulletin of the Ecological Society of America* 83, 131–133.
- Revadi, S., Lebreton, S., Witzgall, P., Anfora, G., Dekker, T., and Becher, P. (2015). Sexual Behavior of *Drosophila suzukii*. *Insects* 6, 183–196.
- Revell, L.J. (2007). The G matrix under fluctuating correlational mutation and selection. *Evolution* 61, 1857–1872.
- Reznick, D. N., Shaw, F. H., Rodd, F. H., & Shaw, R. G. (1997). Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, 275(5308), 1934-1937.
- Reznick, D.N., and Ricklefs, R.E. (2009). Darwin's bridge between microevolution and macroevolution. *Nature* 457, 837–842.
- Ricciardi, A., Palmer, M.E., and Yan, N.D. (2011). Should Biological Invasions Be Managed as Natural Disasters? *BioScience* 61, 312–317.
- Richards, C.L., Bossdorf, O., Muth, N.Z., Gurevitch, J., and Pigliucci, M. (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* 9, 981–993.
- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D., and West, C.J. (2000).

- Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 6, 93–107.
- Rius, M., and Darling, J.A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution* 29, 233–242.
- Robinson, R. (2008). Human Activity, not Ecosystem Characters, Drives Potential Species Invasions. *PLoS Biology* 6, e39.
- Roff, D. (2000). The evolution of the G matrix: selection or drift? *Heredity* 84, 135–142.
- Roff, D.A. (2007). A centennial celebration for quantitative genetics. *Evolution* 61, 1017–1032.
- Roff, D. A., & Mousseau, T. A. (1999). Does natural selection alter genetic architecture? An evaluation of quantitative genetic variation among populations of *Allenomobius socius* and *A. fasciatus*. *Journal of Evolutionary Biology*, 12(2), 361-369.
- Rohlf, F. J. (1998). On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology*, 47(1), 147-158.
- Rosecchi, E., Thomas, F., and Crivelli, A.J. (2001). Can life-history traits predict the fate of introduced species? A case study on two cyprinid fish in southern France. *Freshwater Biology* 46, 845–853.
- Rushton, S.P., Lurz, P.W.W., Gurnell, J., and Fuller, R. (2000). Modelling the spatial dynamics of parapoxvirus disease in red and grey squirrels: a possible cause of the decline in the red squirrel in the UK? *Journal of Applied Ecology* 37, 997–1012.
- Rutherford, S.L. (2003). Between genotype and phenotype: protein chaperones and evolvability. *Nature Reviews Genetics* 4, 263–274.
- Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A., Baughman, S., Cabin, R.J., Cohen, J.E., Ellstrand, N.C., *et al.* (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics* 305–332.
- Sax, D.F., and Brown, J.H. (2000). The paradox of invasion. *Global Ecology and Biogeography* 9, 363–371.
- Schlichting, C.D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics* 667–693.
- Schmid-Hempel, P., Schmid-Hempel, R., Brunner, P.C., Seeman, O.D., and Allen, G.R. (2007). Invasion success of the bumblebee, *Bombus terrestris*, despite a drastic genetic bottleneck. *Heredity* 99, 414–422.
- Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 1766-1774.
- Schulte, U., Veith, M., Mingo, V., Modica, C., and Hochkirch, A. (2013). Strong genetic differentiation due to multiple founder events during a recent range expansion of an introduced wall lizard population. *Biological Invasions* 15, 2639–2649.
- Scoville, A., Lee, Y.W., Willis, J.H., and Kelly, J.K. (2009). Contribution of chromosomal

- polymorphisms to the G-matrix of *Mimulus guttatus*. *New Phytologist* 183, 803–815.
- Sgrò, C.M., Lowe, A.J., and Hoffmann, A.A. (2011). Building evolutionary resilience for conserving biodiversity under climate change: Conserving biodiversity under climate change. *Evolutionary Applications* 4, 326–337.
- Shaw, F.H., Shaw, R.G., Wilkinson, G.S., and Turelli, M. (1995). Changes in Genetic Variances and Covariances: G Whiz! *Evolution* 49, 1260.
- Shearer, P.W., West, J.D., Walton, V.M., Brown, P.H., Svetec, N., and Chiu, J.C. (2016). Seasonal cues induce phenotypic plasticity of *Drosophila suzukii* to enhance winter survival. *BMC Ecology* 16.
- Simon, M.N., Machado, F.A., and Marroig, G. (2016). High evolutionary constraints limited adaptive responses to past climate changes in toad skulls. *Proceedings of the Royal Society B: Biological Sciences* 283, 20161783.
- Simons, A.M. (2002). The continuity of microevolution and macroevolution. *Journal of Evolutionary Biology* 15, 688–701.
- Slatyer, R.A., Nash, M.A., Miller, A.D., Endo, Y., Umbers, K.D., and Hoffmann, A.A. (2014). Strong genetic structure corresponds to small-scale geographic breaks in the Australian alpine grasshopper *Kosciuscola tristis*. *BMC Evolutionary Biology* 14, 1.
- Stanton, M.L., Roy, B.A., and Thiede, D.A. (2000). Evolution in stressful environments. I. Phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. *Evolution* 54, 93–111.
- Stearns, S.C. (1989). The Evolutionary Significance of Phenotypic Plasticity. *BioScience* 39, 436–445.
- Stebbins, G.L. (1974). *Flowering plants: evolution above the species level*. Cambridge, Mass.: Harvard University Press
- Steppan, S.J., Phillips, P.C., and Houle, D. (2002). Comparative quantitative genetics: evolution of the G matrix. *Trends in Ecology & Evolution* 17, 320–327.
- Takahashi, K. H., & Kimura, M. T. (2005). Intraspecific and interspecific larval interaction in *Drosophila* assessed by integrated fitness measure. *Oikos*, 111(3), 574-581.
- Templeton, A.R. (1980). The theory of speciation via the founder principle. *Genetics* 94, 1011–1038.
- Tingley, R., Vallinoto, M., Sequeira, F., and Kearney, M.R. (2014). Realized niche shift during a global biological invasion. *Proceedings of the National Academy of Sciences* 111, 10233–10238.
- Tochen, S., Dalton, D.T., Wiman, N., Hamm, C., Shearer, P.W., and Walton, V.M. (2014). Temperature-Related Development and Population Parameters for *Drosophila suzukii* (Diptera: Drosophilidae) on Cherry and Blueberry. *Environmental Entomology* 43, 501–510.
- Tompkins, D.M., White, A.R., and Boots, M. (2003). Ecological replacement of native red squirrels

- by invasive greys driven by disease. *Ecology Letters* 6, 189–196.
- Torchin, M.E., Lafferty, K.D., Dobson, A.P., McKenzie, V.J., and Kuris, A.M. (2003). Introduced species and their missing parasites. *Nature* 421, 628–630.
- Trut, L., Oskina, I., and Kharlamova, A. (2009). Animal evolution during domestication: the domesticated fox as a model. *BioEssays* 31, 349–360.
- Tsutsui, N.D., Suarez, A.V., Holway, D.A., and Case, T.J. (2000). Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences* 97, 5948–5953.
- Turelli, M. (1988). Phenotypic Evolution, Constant Covariances, and the Maintenance of Additive Variance. *Evolution* 42, 1342.
- Valéry, L., Fritz, H., Lefeuvre, J.-C., and Simberloff, D. (2008). In search of a real definition of the biological invasion phenomenon itself. *Biological Invasions* 10, 1345–1351.
- Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* 34, 275–305.
- Vilcinskas, A. (2015). Pathogens as biological weapons of invasive species. *PLoS Pathog* 11, e1004714.
- Voje, K.L., Hansen, T.F., Egset, C.K., Bolstad, G.H., and Pélabon, C. (2014). Allometric constraints and the evolution of allometry. *Evolution* 68, 866–885.
- Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811), 563-565.
- Waddington, C.H. (1953). Genetic Assimilation of an Acquired Character. *Evolution* 7, 118.
- Walling, C.A., Morrissey, M.B., Foerster, K., Clutton-Brock, T.H., Pemberton, J.M., and Kruuk, L.E.B. (2014). A Multivariate Analysis of Genetic Constraints to Life History Evolution in a Wild Population of Red Deer. *Genetics* 198, 1735–1749.
- Walsh, B., and Blows, M.W. (2009). Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and Systematics* 40, 41–59.
- Walsh, J.R., Carpenter, S.R., and Vander Zanden, M.J. (2016). Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proceedings of the National Academy of Sciences* 113, 4081–4085.
- Watts, P.C., Keat, S., and Thompson, D.J. (2010). Patterns of spatial genetic structure and diversity at the onset of a rapid range expansion: colonisation of the UK by the small red-eyed damselfly *Erythromma viridulum*. *Biological Invasions* 12, 3887–3903.
- Weber, K. E. (1992). How small are the smallest selectable domains of form?. *Genetics*, 130(2), 345-353.
- West-Eberhard, M.J. (1989). Phenotypic plasticity and the origins of diversity. *Annual Review of*

Ecology and Systematics 249–278.

West-Eberhard, M.J. (2003). Developmental plasticity and evolution. Oxford University Press.

Whitlock, M.C., Phillips, P.C., and Fowler, K. (2002). Persistence of changes in the genetic covariance matrix after a bottleneck. *Evolution* 56, 1968.

Wilkinson, G.S., Fowler, K., and Partridge, L. (1990). Resistance of Genetic Correlation Structure to Directional Selection in *Drosophila melanogaster*. *Evolution* 44, 1990.

Wittenberg, R., & Cock, M. J. (Eds.). (2001). Invasive alien species: a toolkit of best prevention and management practices. CABI.

Wood, C.W., and Brodie, E.D. (2016). Evolutionary response when selection and genetic variation covary across environments. *Ecology Letters* 19, 1189–1200.

Wright, S. (1932). *The roles of mutation, inbreeding, crossbreeding, and selection in evolution* (Vol. 1, pp. 356-366).

Yeh, P.J., and Price, T.D. (2004). Adaptive Phenotypic Plasticity and the Successful Colonization of a Novel Environment. *The American Naturalist* 164, 531–542.

Yoshida, T., Goka, K., Ishihama, F., Ishihara, M., and Kudo, S. (2007). Biological invasion as a natural experiment of the evolutionary processes: introduction of the special feature. *Ecological Research* 22, 849–854.

Zenni, R.D., and Nuñez, M.A. (2013). The elephant in the room: the role of failed invasions in understanding invasion biology. *Oikos* 122, 801–815.

Zhou, S., Campbell, T.G., Stone, E.A., Mackay, T.F.C., and Anholt, R.R.H. (2012). Phenotypic Plasticity of the *Drosophila* Transcriptome. *PLoS Genetics* 8, e1002593.

Annexes

Annexe du Chapitre 1: Supplementary material for the manuscript „Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC-Random Forest“

Annexe du Chapitre 3: Supplementary material for the manuscript „Phenotypic plasticity of *Drosophila suzukii* wing to developmental temperature: implications for flight performance“

Article en collaboration : Wolbachia in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies

Supplementary material for the manuscript: “Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest”

Antoine Fraimout¹, Vincent Debat¹, Simon Fellous², Ruth Hufbauer^{2,3}, Julien Foucaud², Pierre Pudlo⁴, Jean-Michel Marin⁵, Donald K. Price⁶, Julien Cattel⁷, Xiao Chen⁸, Marindia Depra⁹, Pierre François Duyck¹⁰, Christelle Guedot¹¹, Gregory Loeb¹², Marc Kenis¹³, Masahito T. Kimura¹⁴, Anne Loiseau², Isabel Martinez-Sañudo¹⁵, Marta Pascual¹⁶, Maxi Polihronakis Richmond¹⁷, Peter Shearer¹⁸, Nadia Singh¹⁹, Koichiro Tamura²⁰, Anne Xuéreb², Jinping Zhang²¹ & Arnaud Estoup².

Corresponding author: Arnaud Estoup, UMR Centre de Biologie pour la Gestion des Populations (INRA / IRD / Cirad / Montpellier SupAgro), 755 avenue du Campus Agropolis, CS 30 016, 34988 Montferrier-sur-Lez cedex, France; email: estoup@supagro.inra.fr

Short title: Reconstructing invasion routes using ABC random forest

Key words: *Drosophila suzukii*, invasion routes, random forest, approximate Bayesian computation, population genetics.

Appendix S1: Characterization of genetic diversity and structure within and between all genotyped sample sites and definition of genetic groups and representative sample sites for ABC analyses

Motivation

A necessary step prior to conducting any type of ABC analysis is to define genetic groups for use in analyses. If every sample site is used as a genetic group, the number of different possible invasion scenarios becomes computationally prohibitive, and thus this first step is an important one for facilitating coherent analysis (Estoup & Guillemaud 2010; Lombaert *et al.* 2014). Genetic diversity and structure within and between all genotyped sample sites is thus characterized using traditional statistics and clustering methods.

Materials and methods

We genotyped a total of 685 individuals from 23 localities representing the worldwide distribution of *D. sukuzii* (table S1 and fig. 1 in main text) at 25 microsatellite loci (Framout *et al.* 2015). The complete data file including the multilocus 685 genotypes coded in a 3-digit GENEPOP format (Raymond & Rousset 1995) is available under request from the corresponding author (AE). We estimated the level of genetic variability within samples by calculating the mean expected heterozygosity H_e (Nei 1987) and the allelic richness (A) corrected for 15 individuals with FSTAT version 2.9.3.2 (Goudet 2002). We assessed genetic structure between sample sites by computing pairwise F_{ST} values as well as exact test for genotypic differentiation for all pairs of sample sites, using GENEPOP (Raymond & Rousset 1995). Genetic structure was further studied using three complementary clustering methods. First, we used STRUCTURE v2.3.3 (Pritchard *et al.* 2000) assuming a model of admixture with correlated allele frequencies, using parameters set as default and a burn-in period of 10^5 MCMC iterations followed by 10^5 MCMC iterations. The number of possible clusters (K) was set from 1 to 15 and analyses were conducted for 20 replicate runs. Further processing of the STRUCTURE outputs was carried following the methodology described in Lombaert *et al.* (2014) using CLUMP and DISTRUCT softwares (Jakobsson & Rosenberg 2007; Wang *et al.* 2007). Second we used BAPS (Corander *et al.* 2003) analyses carried out on groups of individuals (i.e. sample sites) rather than individuals, with more simple model assumptions (i.e. no admixture and uncorrelated allele frequencies). Third we used the Discriminant Analysis of Principal Components method (DAPC, Jombart *et al.* 2010) implemented in the adegenet R package (v.3.1.0, R core team 2008). DAPC seeks synthetic variables among samples as well as the discriminant functions showing differences between groups as best as possible, while minimizing variation within clusters.

Results

We found substantial polymorphism in our dataset. Heterozygosity was lowest in the invasive Hawaiian sample site (US-Haw, $H_e = 0.623$) and highest in the native Chinese sample site (CN-Lan, $H_e = 0.834$; fig. S5). The same populations had the lowest and highest (corrected) allelic diversity ($A_{\text{US-Haw}} = 3.392$, $A_{\text{CN-Lan}} = 9.347$). All pairwise comparisons indicated statistically significant genotypic differentiation except for between the two Japanese sample sites and two Chinese sample sites (CN-Lan & CN-Lia; table S5). We found an overall low level of genetic differentiation in the native Asian area with a mean $F_{\text{ST}} = 0.014$ (F_{ST} values between all population pairs are given in table S5). Similarly low levels of differentiation were found among European populations with a mean F_{ST} of 0.015. We found higher levels of genetic differentiation among continental US populations (mean $F_{\text{ST}} = 0.044$). Eastern US populations showed overall homogeneity (mean $F_{\text{ST}} = 0.012$) whereas the three western US samples revealed a substantially higher level of differentiation (mean $F_{\text{ST}} = 0.051$). North American samples were the least differentiated from the native Asian group (mean $F_{\text{ST}} = 0.039$, between Asia and the western US; mean $F_{\text{ST}} = 0.056$ between Asia and the eastern US) and also showed relatively small differentiation with the Brazilian sample site (mean $F_{\text{ST}} = 0.028$ between the western US and Brazil; mean $F_{\text{ST}} = 0.031$ between the eastern US and Brazil). European sample sites were more differentiated from the native ones ($F_{\text{ST}} = 0.075$) and shown the lowest level of differentiation with La Réunion sample site (mean $F_{\text{ST}} = 0.046$). The Hawaiian sample site showed substantial genetic differentiation with all other samples, with mean F_{ST} values ranging from 0.079 (between Hawaii and the western US) to 0.196 (between Hawaii and La Réunion).

Although genetic clustering analyses using BAPS, STRUCTURE and DAPC did not converge on the same number of genetic groups (K), a strong, shared cluster identity could be observed among clustering methods. The Hawaiian sample site formed a single group in all clustering analyses, consistent with the high F_{ST} values found between Hawaii and all other sample sites. All native Asian sample sites were included in a single cluster in both BAPS and DAPC analyses (fig. 1 in main text and fig. S1, respectively). STRUCTURE results indicated a single cluster in Asia for $K = 2$ to 5 and Japanese sample sites formed a second Asian cluster for $K > 5$ (fig. S2). European sample sites systematically formed a single cluster in BAPS, DAPC, and STRUCTURE. A higher heterogeneity in clustering results was observed for North American sample sites. Analyses using BAPS indicated the presence of three genetic groups in North America, which consisted of one eastern US group (including the sample sites US-Col, US-Gen, US-Wis and US-NC), one south-western group (US-Wat and US-SD) and a third group including a single sample site from Oregon state (US-Sok; fig. 1 in main text). Results based on DAPC and STRUCTURE indicated a homogeneous grouping of all sample sites from the eastern US, and distinguished the Oregon sample site (US-Sok) from other continental USA samples, grouping this

sample site with Hawaii (fig. S1 and S2, respectively). STRUCTURE analyses showed signatures of genetic introgression from Hawaii in Western US samples, with a stronger introgression pattern in the North than in the South. The Brazilian sample site was assigned along with the eastern US populations by DAPC and STRUCTURE, but to a south-western US group (including the US-Wat and US-SD sample sites) in BAPS. Finally, individuals from La Réunion belonged to the European genetic group in DAPC and STRUCTURE, but formed a single group in BAPS.

We used the results obtained from the three genetic clustering methods along with historical (i.e. dates of first record ; table S1) and geographical information to define seven main genetic groups from our set of 23 sample sites that will be further considered for ABC analyses: Asia (sample sites CN-Lan, CN-Lia, CN-Nin, CN-Shi, JP-Tok and JP-Sap), Hawaii (sample site US-Haw), western US (sample sites US-Wat, US-Sok and US-SD), eastern US (sample sites US-Col, US-NC, US-Wis and US-Gen), Europe (sample sites GE-Dos, FR-Par, FR-Bor, FR-Mon, SW-Del, SP-Bar and IT-Tre), Brazil (BR-PA) and La Réunion (FR-Reu). This allowed us to define eleven nested sets of competing invasion scenarios that we analyzed sequentially using ABC model choice methodologies to reconstruct the worldwide routes of invasion of *D. suzukii* (table 1 in main text).

To reduce computational resources required, ABC treatments were carried out using only one representative sample site for each genetic group per replicate analysis. In other words, when a genetic group included more than two sample sites, we considered the two most differentiated sample sites (i.e. with the highest F_{ST} values in table S5) as representative of the group: JP-Tok and CN-Shi for the group Asia, all three sample sites for the genetically heterogeneous western US group, US-NC and US-Wis for the group eastern US, IT-Tre and GE-Dos for the group Europe (but see the Results section *Origins of invasive genetic groups* of the main text for a further sub-partitioning of the latter group). Each of the two retained sample sites (i.e. the representative sample sites) were included alternatively in the sample set analyzed using ABC, hence providing the possibility to replicate analyses of scenario choice on different sample sets to verify the robustness of our inferences.

References cited in Appendix S1

- Corander J, Waldmann P, Sillanpaa MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics*.163: 367–374.
- Estoup A, Guillemaud T. 2010. Reconstructing routes of invasion using genetic data: why, how and so what?. *Mol Ecol*. 19, 4113–4130.
- Fraimout A, Loiseau A, Price DK, Xuereb A, Martin JF, Vitalis R, Fellous S, Debat V, Estoup A. 2015. New set of microsatellite markers for the spotted-wing *Drosophila suzukii* (*Diptera: Drosophilidae*): A promising molecular tool for inferring the invasion history of this major

- insect pest. *Eur J Entomol.* 112(4): 855.
- Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* 23: 1801–1806.
- Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics.* 11: 94.
- Goudet J. 2002. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Updated from Goudet (1995).
- Lombaert E, Guillemaud T, Lundgren J, Koch R, Facon B, Grez A, Loomans A, Malausa T, Nedved O, Rhule E, Staverlokk A, Steenberg T, Estoup A. 2014. Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Mol Ecol.* 23: 5979–5997.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics.* 155: 945–959.
- Raymond M, Rousset F. 1995. Genepop (version. 1.2), a population genetics software for exact tests and ecumenicism. *J Hered.* 86: 248–249.
- Robert CP, Cornuet JM, Marin JM, Pillai NS. 2011. Lack of confidence in approximate
- Wang S, Lewis CM, Jakobsson M. *et al.* 2007. Genetic variation and population structure in Native Americans. *Plos Genetics.* 3, 2049–2067.

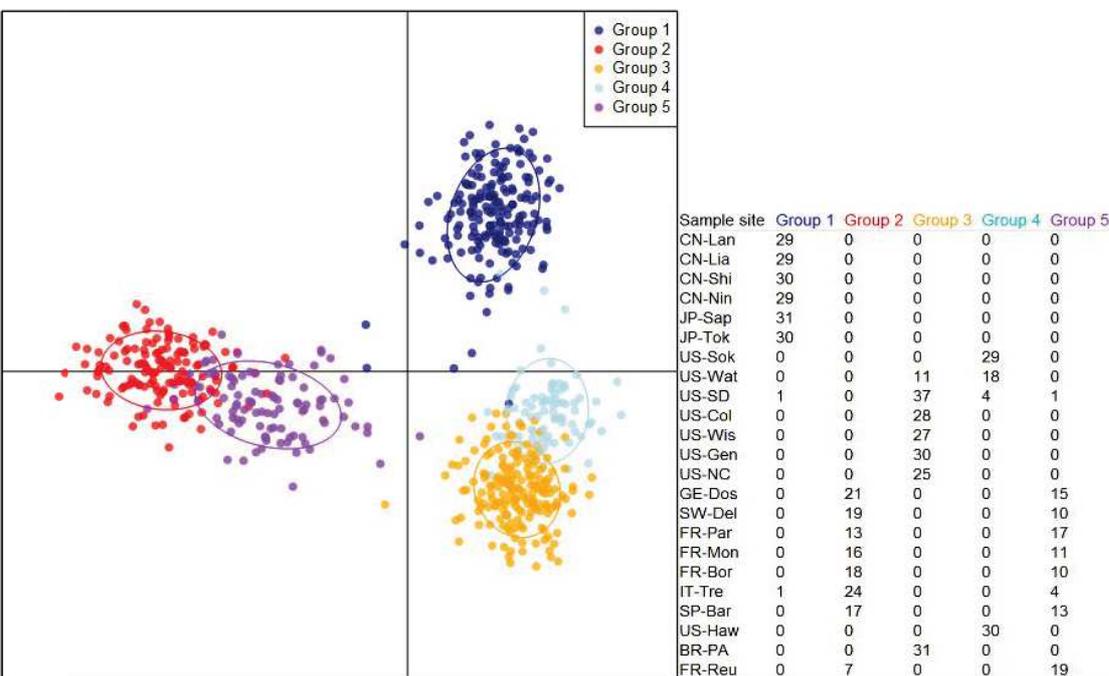


Figure S1. Results from analysis using Discriminant Analysis of Principal Components (DAPC).

Genetic groups, as inferred by DAPC for each individual, are indicated with different colors. Contributions from all sample sites to each genetic group is detailed in the table at the right, with the number of individuals per sample site assigned to each genetic group. See table S1 and fig. S1 for information on sample sites.

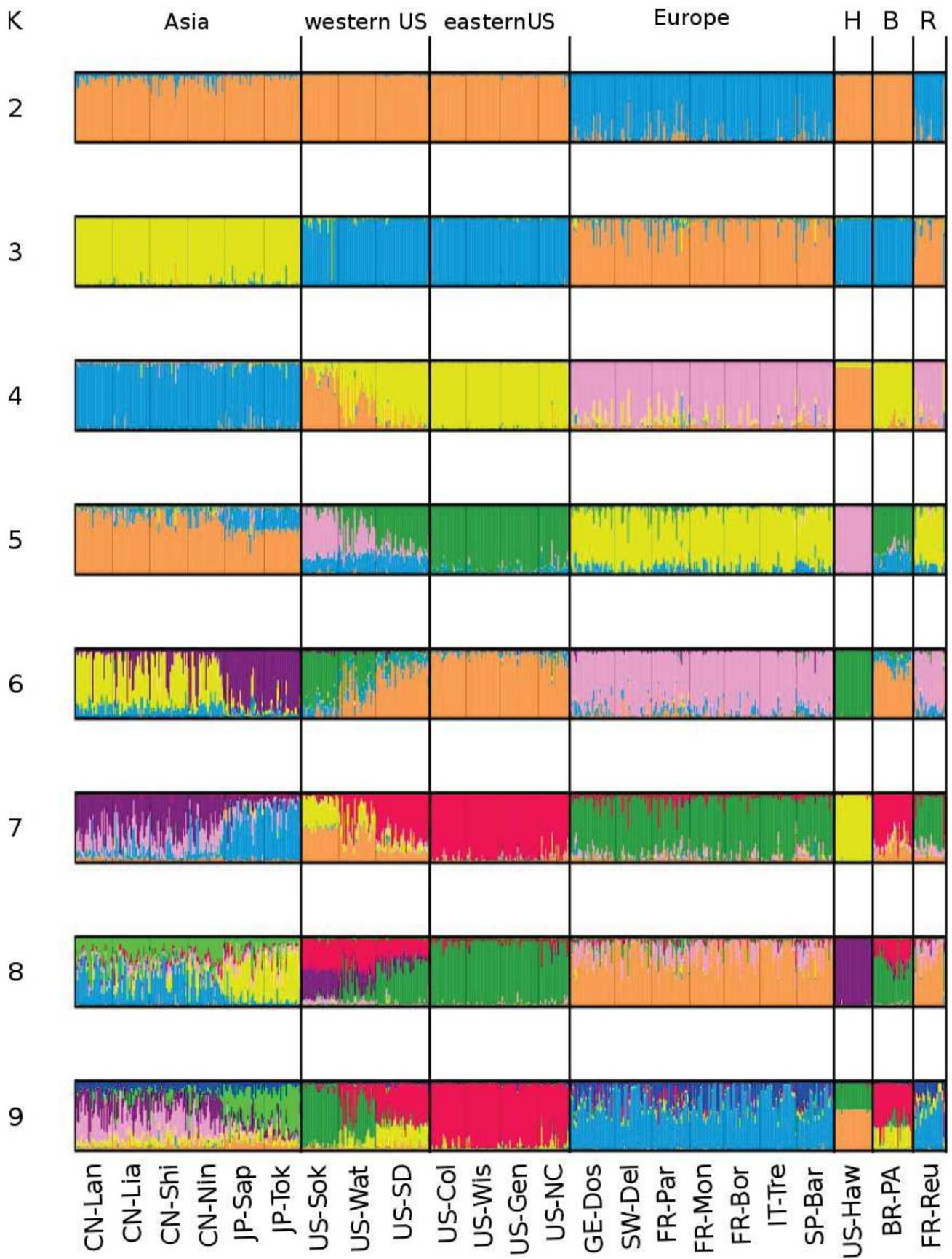


Figure S2. Genetic groups identified by STRUCTURE.

Results are given for a number of genetic group, K, ranging from 2 to 9. Each vertical line represents an individual and each colour represents a genetic group. Individuals are grouped by sample site along the x-axis. The names at the top of the figure refers to the genetic groups used in ABC analyses. H = Hawaii, B = Brazil, R = La Reunion. The names at the bottom of the figure correspond to the sample sites described in table S1 and fig. 1.

Best model = Admixture between Asia and eastern US

Best model = Single introduction event from Asia

Source sample site of the eastern US group used for each replicated ABC-RF analyses

1: US-NC

2: US-Wis

3: US-Gen

4: US-Col

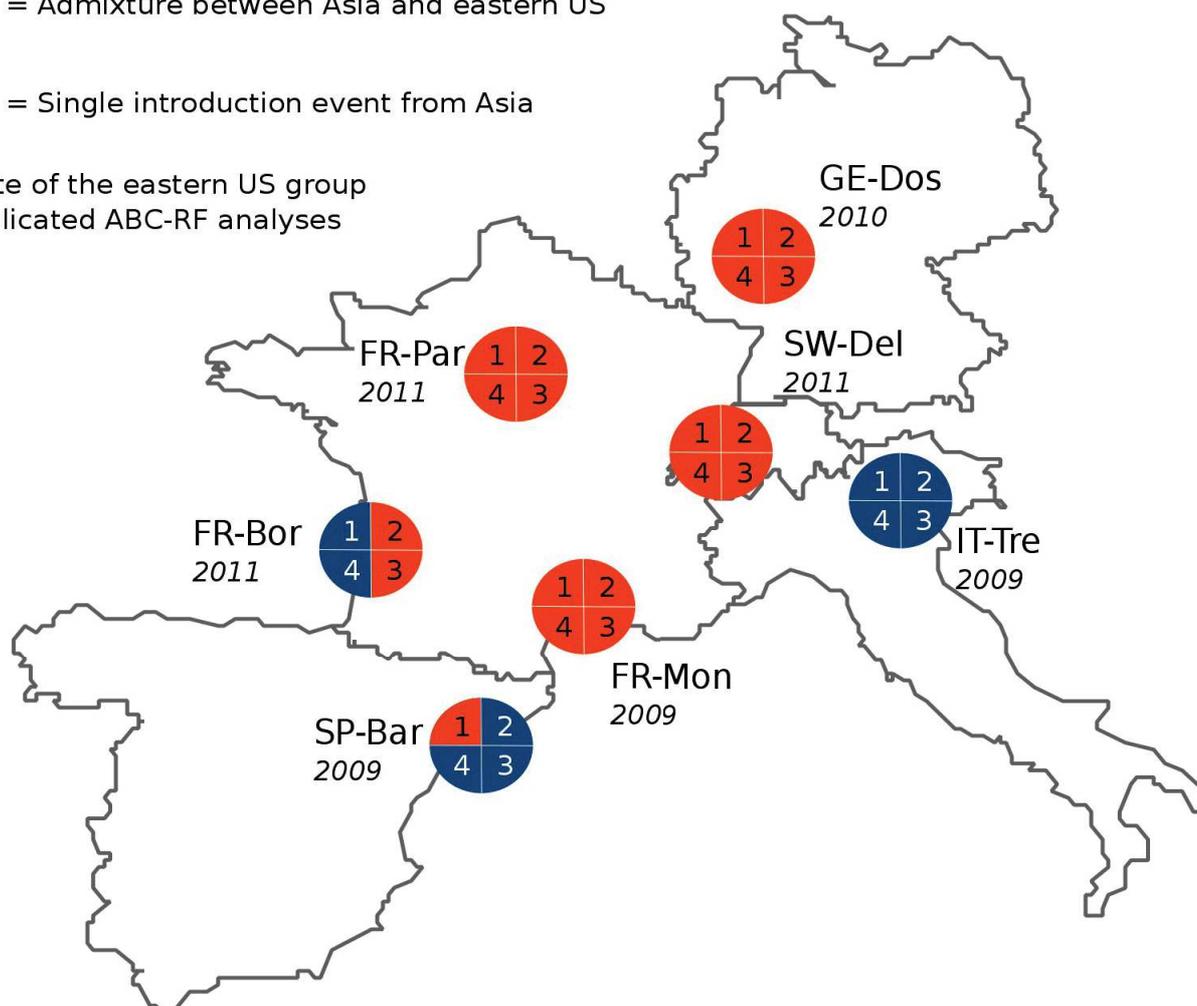


Figure S3: Admixed or non-admixed origin of *D. sukuzii* in Europe inferred for each sample sites, using an alternative representative sample site for the native range.

Potential source populations include (among others) the invasive genetic group from eastern US represented by four invasive sample sites (see fig. 1) and the Asian native range. In contrast to fig. 2 of the main text, the Asian native range is here represented by the Chinese sample site CN-Shi, which shows the highest F_{st} values with other Asian sample sites (table S3). Four replicate independent ABC-RF treatments corresponding to the analysis 3a (table 1) were carried out for each targeted European sample site using one of the four eastern US sample sites for each treatment (US-NC, US-Wis, US-Gen and US-Col for the treatments labeled 1, 2, 3 and 4 in the pies of the figure, respectively). A pie quarter in blue indicates that the best scenario corresponds to a single introduction event from Asia. A pie quarter in red indicates that the best scenario corresponds to an admixture event between Asia and eastern US.

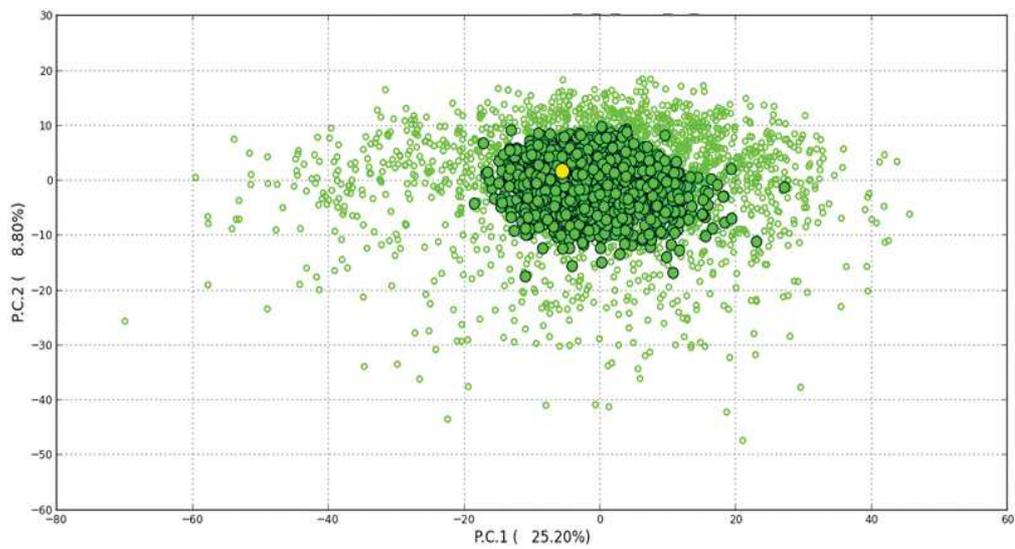


Figure S4. Principal component analysis of 1141 test quantities when processing ABC model-posterior checking for the final worldwide invasion scenario detailed in fig. 1.

The two first axes from the principal component analysis are shown. Small green circles show datasets simulated from priors (subset of 2,000 plots). Large green circles show datasets simulated from posteriors (subset of 2,000 plots). The large yellow circle is the observed dataset. See Materials and methods text section for details.

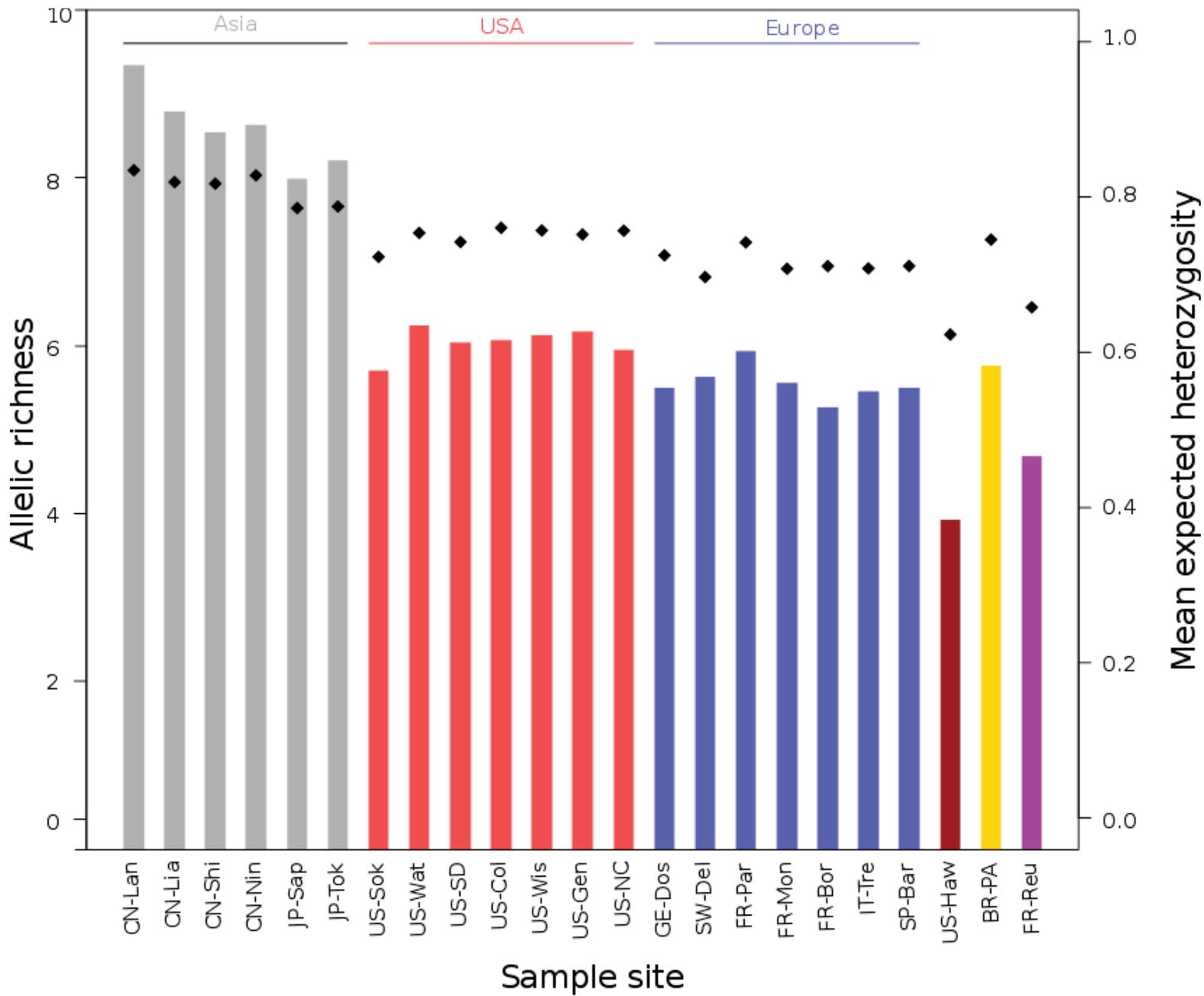


Figure S5. Genetic diversity in the invasive and native sample sites of *D. sukuzii*.

Mean expected heterozygosity (diamonds) and average genetic diversity estimated as allelic richness at 25 microsatellite loci corrected for 15 individuals (bars) are shown. Code names of sample sites given at bottom are the same as those in fig. 1 and table S1.

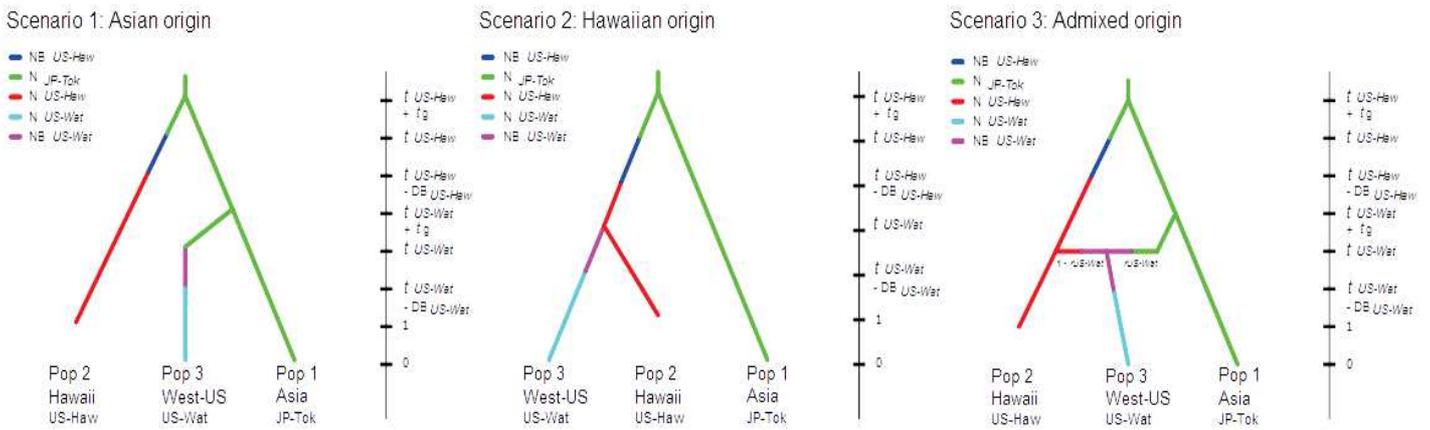


Figure S6. Graphic representation of three ABC scenarios to infer the origin of the US-Wat sample (analysis 1a in table S1 and 2).

The western US target genetic unit is represented by the site sample US-Wat and putative sources by US-Haw and JP-Tok (genetic group Hawaii and Asia respectively). Site sample code names are as in table 1 and fig. 1. All parameters with associated prior distributions are described in table S8. Time 0 is the sampling year 2014 and time 1 is the sampling year 2013. Time is not to scale. Scenario 1 corresponds to an Asian native origin (Pop 1) of the target genetic unit (Pop 3). Scenario 2 corresponds to a Hawaiian origin (Pop 2). In scenario 3, the target genetic unit is the result of an admixture between individuals from the Asian group (Pop 1) at a rate r_{US-Wat} (table 3 and S4) and from the Hawaiian group (Pop 2) at a rate $1 - r_{US-Wat}$. Note the modeling of a native "ghost population" (i.e., unsampled population) by a green line diverging from Pop 1 at a time $t_{US-Haw} + t_g$, with t_g drawn into a loose flat prior, which includes zero at the lower bound (table S8)

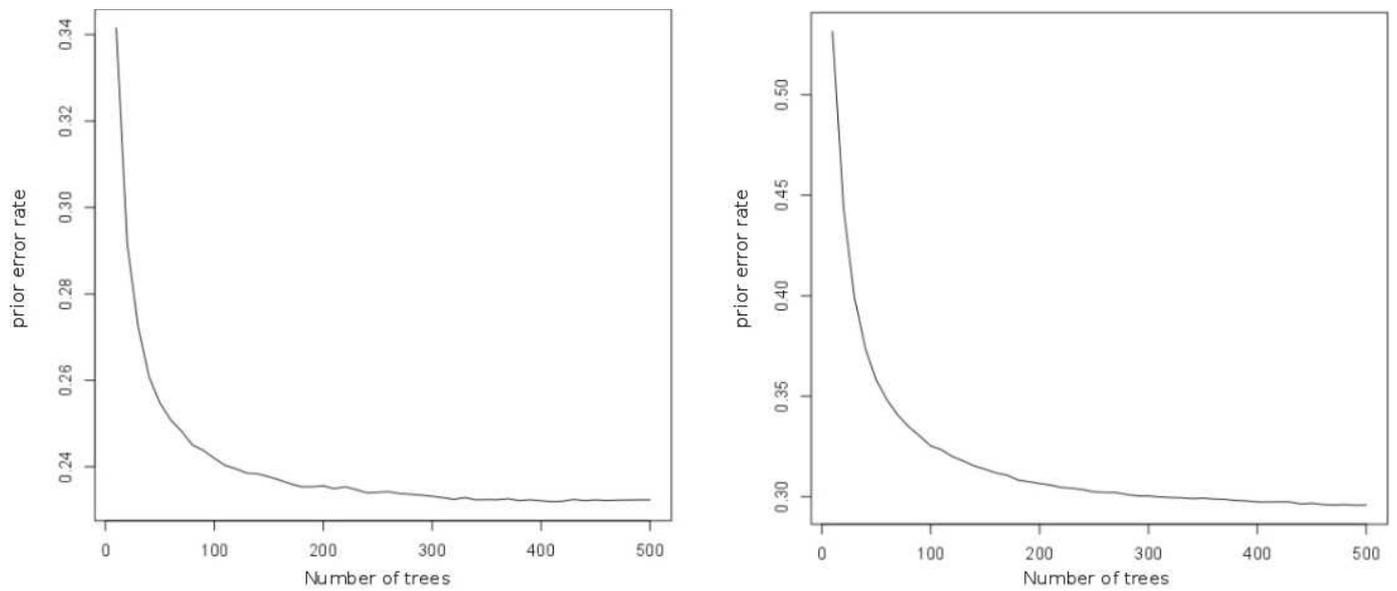


Figure S7. Evolution of the ABC-RF prior error rate with respect to the number of trees in the forest.

Graphs represent the decrease of the ABC-RF prior error rate with the number of trees in the forest for a model setting including six competing scenarios as in analyses 2a and 2b (left panel) and for a more complex setting including 21 competing scenarios as in analyses 5a and 5b (right panel).

Sample site name	Historical status	Sampling location	Geographic coordinates	Date of first observation (year)	Sampling year	Number of genotyped individuals
CN-Lan	Native	Langfang city, Hebei province, China	39° 32' 16.969" N 116° 41' 1.507" E	NA	2014	29
CN-Lia	Native	Liaoyuan city, Jilin province, China	42° 53' 15.961" N 125° 8' 37.175" E		2014	29
CN-Shi	Native	Shiping county, Yunnan province, China	23° 42' 21.37" N 102° 29' 41.942 " E		2014	30
CN-Nin	Native	Ningbo city, Zhejiang province, China	29° 52' 6.01" N 121° 32' 38.364" E		2014	29
JP-Tok	Native	Tokyo, Honshu, Japan	35° 41' 22.155" N 139° 41' 30.143" E		2014	30
JP-Sap	Native	Sapporo, Hokkaido, Japan	43° 3' 43.545" N 141° 21' 15.754" E		2014	31
US-Haw	Invasive	Hilo Hawaii, USA	19° 49' 5.305" N 155° 38' 2.743" E	1980	2013	30
US-Wat	Invasive	Watsonville, California, USA	36° 54' 36.832" N 121° 45' 24.82" E	2008	2014	29
US-SD	Invasive	San-Diego, California, USA	32° 42' 56.657" N 117° 9' 39.901" E	2009	2013	43
US-Sok	Invasive	Dayton, Oregon, USA	45° 13' 14.422" N 123° 4' 34.368" E	2009	2014	29
US-NC	Invasive	Raleigh, North Carolina, USA	35° 46' 46.523" N 78° 38' 17.443" E	2010	2014	25
US-Wis	Invasive	Barneveld, Wisconsin, USA	43° 0' 55.989" N 89° 53' 43.443" E	2010	2014	27
US-Gen	Invasive	Geneva, New-York state, USA	42° 52' 4.741" N 76° 59' 9.005" E	2010	2014	30
US-Col	Invasive	Fort Collins, Colorado, USA	40° 35' 6.937" N 105° 5' 3.922" E	2012	2015	28
SP-Bar	Invasive	Barcelona, Spain	41° 23' 6.23" N 2° 10' 24.252" E	2008	2014	30
IT-Tre	Invasive	Trento, Italy	46° 3' 4.321" N 11° 7' 3.141" E	2008	2014	29
FR-Mon	Invasive	Montpellier, France	43° 36' 38.768" N 3° 52' 36.177" E	2009	2013	27
FR-Bor	Invasive	Sauternes, France	44° 31' 51.715" N 0° 20' 35.764" E	2010	2013	28
GE-Dos	Invasive	Dossenheim, Germany	49° 26' 54.02" N 8° 40' 23.714" E	2010	2013	36
FR-Par	Invasive	Paris, France	48° 51' 23.81" N 2° 21' 7.998" E	2010	2015	30
SW-Del	Invasive	Delémont, Switzerland	47° 21' 57.013" N 7° 20' 42.559" E	2011	2013	29
BR-PA	Invasive	Porto Alegre, Brazil	30° 2' 4.763" N 51° 13' 3.57" E	2014	2014	31
FR-Reu	Invasive	La Réunion France	21° 6' 54.508" N 55° 32' 10.982" E	2014	2014	26

Table S1. Locations from which *D. sukuzii* individuals were collected for these analyses, including the date of first observation at the site, and the year sampled.

Analysis	Scenario number	Specificities of the compared scenarios (origin of the focal population)	Focal genetic group
1a+1b+1c	1	Single introduction event from Asia	western US
	2	Single introduction event from Hawaii	
	3	Admixture event between Asia & Hawaii	
1d	1	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-Sok and US-SD	western US
	2	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-SD, independent admixture event in US-Sok between Hawaii and Asia	
	3	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-Sok, independent admixture event in US-SD between Hawaii and Asia	
	4	Three independent admixture events between Hawaii and Asia in each populations	
	5	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-Sok and US-SD and secondary introduction event from Asia to US-SD	
	6	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-Sok and US-SD and secondary introduction event from Hawaii in US-Sok	
	7	Single admixture event in US-Wat between Hawaii and Asia, followed by dispersion to US-Sok and US-SD and secondary introduction events from Hawaii in US-Sok and from Asia to US-SD	
2a + 2b	1	Single introduction event from Asia	eastern US (2a) or Europe (2b)
	2	Single introduction event from Hawaii	
	3	Single introduction event from western US	
	4	Admixture event between Asia & Hawaii	
	5	Admixture event between Asia & western US	
	6	Admixture event between Hawaii & western US	
3a + 3b	1	Single introduction event from Asia	eastern US (3a) or Europe (3b)
	2	Single introduction event from Hawaii	
	3	Single introduction event from western US	
	4	Single introduction event from eastern US (3a) or Europe (3b)	
	5	Admixture event between Asia & Hawaii	
	6	Admixture event between Asia & western US	
	7	Admixture event between Asia & eastern US (3a) or Europe (3b)	
	8	Admixture event between Hawaii & western US	
	9	Admixture event between Hawaii & eastern US (3a) or Europe (3b)	
	10	Admixture event between western US & eastern US (3a) or Europe (3b)	
4	1	Admixture event between eastern US & secondary Asian population	northern Europe
	2	Admixture event between eastern US & southern Europe	
5a + 5b	1	Single introduction event from Asia	Brazil (5a) or La Réunion (5b)
	2	Single introduction event from Hawaii	
	3	Single introduction event from western US	
	4	Single introduction event from eastern US	
	5	Single introduction event from southern Europe	
	6	Single introduction event from northern Europe	
	7	Admixture event between Asia & Hawaii	
	8	Admixture event between Asia & western US	
	9	Admixture event between Asia & eastern US	
	10	Admixture event between Asia & southern Europe	
	11	Admixture event between Asia & northern Europe	
	12	Admixture event between Hawaii & western US	
	13	Admixture event between Hawaii & eastern US	
	14	Admixture event between Hawaii & southern Europe	
	15	Admixture event between Hawaii & northern Europe	
	16	Admixture event between western US & eastern US	
	17	Admixture event between western US & southern Europe	
	18	Admixture event between western US & northern Europe	
	19	Admixture event between eastern US & southern Europe	
	20	Admixture event between eastern US & northern Europe	
	21	Admixture event between South and northern Europe	

Table S2. Focal genetic group and origin of its source population(s) for each competing scenarios considered for each eleven ABC analyses.

Note: We formalized competing scenarios in a way that a unique introduction event is defined by a single split from one of the possible sources or in the case of admixture, a split from two different

source populations. See table S1 for information on sample sites.

Analysis	Total n° of scenarios	Number of Sum. Stats.	prior error rate			Posterior probability of the best model		Origin of the focal population using either ABC-RF or ABC-LDA (i.e. best model)
			ABC-RF (s.d.)	ABC-LDA (large reference table)	ABC-LDA (small reference table)	ABC-RF (s.d.)	ABC-LDA (large reference table)	
1a	3	39	0.079 (\pm 0.001)	0.059	0.080	0.999 (\pm 0.001)	1.000 [1.000, 1.000]	Asia + Hawaii
1b	3		0.081 (\pm 0.001)	0.071	0.084	0.999 (\pm 0.001)	0.999 [0.999, 0.999]	
1c	3		0.077 (\pm 0.001)	0.062	0.073	1.000 (\pm 0.000)	0.999 [0.999, 0.999]	
1d	7	130	0.346 (\pm 0.001)	0.310	0.397	0.665 (\pm 0.018)	0.612 [0.562, 0.661]	Asia + Hawaii
2a	6	130	0.229 (\pm 0.001)	0.183	0.269	0.859 (\pm 0.019)	0.826 [0.802, 0.849]	western US
2b	6	130	0.229 (\pm 0.001)	0.179	0.305	0.701 (\pm 0.020)	0.773 [0.760, 0.786]	Asia
3a	10	204	0.350 (\pm 0.001)	0.314	0.413	0.783 (\pm 0.016)	0.769 [0.740, 0.797]	western US
3b	10	204	0.354 (\pm 0.001)	0.341	0.412	0.546 (\pm 0.023)	0.845 [0.826, 0.864]	Asia
4	2	301	0.156 (\pm 0.001)	0.217	0.254	0.993 (\pm 0.004)	0.995 [0.994, 0.996]	southern Europe + eastern US
5a	21	424	0.217 (\pm 0.001)	0.162	0.417	0.704 (\pm 0.035)	0.802 [0.769, 0.834]	western US + eastern US
5b	21	424	0.221 (\pm 0.001)	0.172	0.445	0.594 (\pm 0.026)	0.292 [0.204, 0.380]	northern Europe + southern Europe

Table S3. Results of model choice analyses using ABC-RF and ABC-LDA, using the prior set 2.

Note: All model choice analyses were carried out considering the prior set 2 (table S8) and a single set of sample sites representative of the pre-defined genetic groups (sample site JP-Tok for the native Asian group, US-Haw for the Hawaiian group, US-Wat, US-Sok and US-SD for the western US group, US-NC for the eastern US group, IT-Tre and GE-Dos for the southern Europe and northern Europe group respectively, BR-PA for the Brazil group and FR-Reu for the La Reunion group; fig. 1). Datasets were summarized using the whole set of summary statistics proposed by the software DIYABC. The total number of summary statistics ("Number of Sum. Stats.") as well as the total number of compared scenarios ("Total n° of scenarios") are indicated for each analysis. Prior error rates and posterior probabilities of the best model chosen using ABC-RF were averaged over

ten replicate analyses. ABC-RF and ABC-LDA treatments yielded the same best model choice for all analyses and is denoted in the column "Origin of the focal population using either ABC-RF or ABC-LDA (i.e. best model)". Admixture between two source populations are represented by a "+" sign. ABC-LDA posterior probabilities of the best models were estimated using "large reference tables" with 500,000 simulated datasets per scenario. ABC-LDA prior error rates were computed using reference tables of two different sizes: "large reference table" (i.e., 500,000 simulated datasets per scenario) and "small reference table" (i.e., 10,000 simulated datasets per scenario as for ABC-RF analyses). S.d. and CI stand for standard deviation and 95% confidence interval, respectively.

Analysis	Total n° of scenarios	Number of Sum. Stats.	Prior set 1		Origin of the focal population (best model)	Prior set 2		Origins of the focal population (best model)
			prior error rate	posterior probability		prior error rate	posterior probability	
1a	3	39	0.100	1.000	Asia + Hawaii	0.100	1.000	Asia + Hawaii
1b			0.094	0.989		0.094	0.989	
1c			0.096	0.998		0.096	0.998	
1d	7	130	0.327	0.690	Asia + Hawaii	0.327	0.690	Asia + Hawaii
2a	6	130	0.232	0.766	western US	0.227	0.864	western US
2b	6	130	0.224	0.692	Asia	0.224	0.718	Asia + western US
3a	10	204	0.324	0.653	western US	0.355	0.585	western US
3b	10	204	0.393	0.488	Asia + eastern US	0.350	0.501	Asia + eastern US
4	2	301	0.118	1.000	southern Europe + eastern US	0.159	1.000	southern Europe + eastern US
5a	21	424	0.295	0.619	western US + eastern US	0.216	0.674	western US + eastern US
5b	21	424	0.300	0.475	northern Europe + southern Europe	0.223	0.538	northern Europe + southern Europe

Table S4. Results of the ABC-RF analyses defined in table 2 for an alternative set of representative sample sites using prior sets 1 and 2.

Note: The alternative sample sites were chosen considering the most differentiated sites (i.e. sites showing the highest pairwise F_{st} value) for each of the ABC groups (see main text and Appendix S1 for details). Here the alternative sample sites are: CN-Shi for Asia, US-Wis for eastern US, GE-Dos for northern Europe and SP-Bar for southern Europe. ABC-RF posterior probabilities of the best model and prior error rates are averaged over three replicate analyses on the same reference table.

Samples	CN-Lan	CN-Li	CN-Shi	CN-Nin	JP-Sap	JP-Tok	US-Sok	US-Wat	US-SD	US-Col	US-Wis	US-Gen	US-NC	GE-Dos	SW-Del	FR-Par	FR-Mon	FR-Bor	IT-Tre	SP-Bar	US-Haw	BR-PA
CN-Li	0.003																					
CN-Shi	0.014	0.017																				
CN-Nin	0.001	0.002	0.011																			
JP-Sap	0.020	0.017	0.027	0.016																		
JP-Tok	0.022	0.015	0.032	0.018	0.002																	
US-Sok	0.076	0.071	0.084	0.072	0.060	0.069																
US-Wat	0.054	0.049	0.066	0.049	0.042	0.047	0.026															
US-SD	0.067	0.060	0.081	0.058	0.062	0.067	0.093	0.034														
US-Col	0.074	0.063	0.074	0.067	0.061	0.067	0.095	0.044	0.037													
US-Wis	0.081	0.074	0.085	0.076	0.067	0.077	0.102	0.052	0.046	0.003												
US-Gen	0.071	0.061	0.074	0.063	0.056	0.056	0.097	0.047	0.035	0.009	0.021											
US-NC	0.063	0.054	0.067	0.058	0.052	0.050	0.079	0.034	0.033	0.011	0.022	0.005										
GE-Dos	0.072	0.068	0.086	0.068	0.081	0.081	0.135	0.104	0.103	0.107	0.119	0.099	0.092									
SW-Del	0.088	0.090	0.104	0.090	0.108	0.105	0.158	0.135	0.138	0.141	0.145	0.135	0.131	0.023								
FR-Par	0.067	0.063	0.078	0.063	0.083	0.079	0.125	0.094	0.098	0.094	0.109	0.090	0.086	0.019	0.021							
FR-Mon	0.080	0.081	0.100	0.081	0.104	0.096	0.150	0.125	0.126	0.131	0.144	0.125	0.119	0.021	0.015	0.023						
FR-Bor	0.081	0.085	0.099	0.081	0.105	0.101	0.152	0.128	0.133	0.140	0.147	0.134	0.127	0.019	0.008	0.026	0.010					
IT-Tre	0.087	0.088	0.103	0.086	0.111	0.107	0.149	0.132	0.140	0.145	0.152	0.142	0.138	0.030	0.005	0.020	0.017	0.009				
SP-Bar	0.077	0.076	0.094	0.075	0.094	0.088	0.141	0.119	0.120	0.128	0.138	0.121	0.115	0.020	0.010	0.015	0.008	0.004	0.003			
US-Haw	0.114	0.124	0.136	0.120	0.103	0.119	0.089	0.090	0.119	0.155	0.161	0.163	0.159	0.169	0.187	0.168	0.184	0.183	0.182	0.169		
BR-PA	0.074	0.068	0.084	0.068	0.071	0.074	0.084	0.031	0.032	0.035	0.041	0.041	0.028	0.115	0.148	0.104	0.132	0.144	0.149	0.135	0.143	
FR-Reu	0.103	0.103	0.118	0.101	0.121	0.120	0.149	0.126	0.137	0.142	0.155	0.127	0.121	0.061	0.059	0.048	0.058	0.056	0.056	0.045	0.197	0.150

Table S5. F_{ST} values between all pair of sample sites.

Note: Code names of the population sites are similar to those in fig.1. Values in blue bold text indicate pairs of samples showing no significant genotypic differentiation. Red bold text indicates the highest F_{ST} value among sample sites within the genetic groups Asia, eastern US and Europe defined using genetic clustering methods along with historical records information (see Appendix S1 for details). See table S1 for information on sample sites.

Introduction type	Sample site	Prior 1					Prior 2				
		mean	median	mode	q5%	q95%	mean	median	mode	q5%	q95%
Extra continental	US-Haw	0.502	0.491	0.481	0.316	0.714	0.482	0.478	0.473	0.374	0.606
	US-Wat	0.171	0.130	0.107	0.039	0.425	0.136	0.119	0.108	0.057	0.264
	IT-Tre	0.204	0.122	0.117	0.034	0.685	0.203	0.159	0.147	0.077	0.496
	BR-PA	0.161	0.129	0.108	0.021	0.405	0.196	0.184	0.167	0.064	0.371
	FR-Reu	0.258	0.211	0.164	0.049	0.628	0.248	0.220	0.186	0.077	0.512
Intra continental	US-SD	0.148	0.135	0.119	0.048	0.284	0.108	0.094	0.083	0.046	0.211
	US-NC	0.124	0.107	0.082	0.033	0.257	0.120	0.108	0.094	0.053	0.225
Extra + intra continental	US-Sok	0.116	0.098	0.086	0.027	0.248	0.112	0.099	0.093	0.048	0.217
	GE-Dos	0.132	0.115	0.081	0.030	0.278	0.162	0.151	0.135	0.069	0.289
Prior values		0.617	0.200	NA	0.020	1.900	0.351	0.200	NA	0.050	0.822

Table S6. Bottleneck severity in invasive populations of *D. sukuzii* for an alternative set of representative sample sites.

Note: The set of population sites used for the ABC estimations presented here include: (i) for the native area: Japan (JP-Tok+JP-Sap), South-East China (CN-Nin) and North-East China (CN-Lan+CN-Lia), and (ii) for the invaded range : US-Wat, US-Sok and US-SD for western US, US-Wis (instead of US-NC) for eastern US, SP-Bar (instead of IT-Tre) for southern Europe, FR-Par (instead of GE-Dos) for northern Europe, BR-PA for South-America (Brazil) and FR-Reu for La Reunion island. Extra-continental introductions correspond to a long distance introduction from a source located apart from the continent of the focal population, intra-continental introduction corresponds to an introduction event from a source located on the same continent than the focal population, and Extra+Intra continental introduction corresponds to a combination of the two types of sources. Mean, median and mode estimates as well as bounds of 90% credibility intervals (q5% and q95%) are indicated for each bottleneck severity parameter. We roughly classified the estimated bottleneck severity values into three classes: weak (i.e. median value of bottleneck severity < 0.12, in light gray), moderate (0.12 < bottleneck severity < 0.22, in gray) and strong (i.e. bottleneck severity > 0.3, in dark gray). Code names of the sample sites are the same as in table S1 and figure 1.

	Admixture event	Admixture rate (gene fraction from pop x)	mean	median	mode	q5%	q95%
Prior 1	A1	r_{US-Wat} (China – O2)	0.763	0.766	0.780	0.662	0.856
	A2	r_{US-Sok} (USA – US-Wat)	0.232	0.226	0.241	0.090	0.393
	A3	r_{FR-Par} (USA – US-Wis)	0.248	0.238	0.216	0.094	0.431
	A4	r_{BR-PA} (USA – SD)	0.689	0.710	0.722	0.420	0.886
	A5	r_{FR-Reu} (Europe – FR-Par)	0.543	0.558	0.599	0.220	0.815
Prior 2	A1	r_{US-Wat} (China – O2)	0.752	0.753	0.756	0.673	0.825
	A2	r_{US-Sok} (USA – US-Wat)	0.222	0.216	0.204	0.111	0.348
	A3	r_{FR-Par} (USA – US-Wis)	0.257	0.248	0.241	0.121	0.415
	A4	r_{BR-PA} (USA – US-SD)	0.620	0.626	0.641	0.434	0.789
	A5	r_{FR-Reu} (Europe – FR-Par)	0.563	0.576	0.606	0.327	0.753

Table S7. Admixture rates estimated for an alternative set of representative sample sites.

Note: Admixture events are denoted as in fig. 1. Each admixture rate parameter r points to the name of the admixed population site and corresponds to the fraction of genes originating from the source population site in parentheses ($1-r$ genes originate from the other source population site). See table 1 for details on sample sites. Mean, median and mode estimates as well as bounds of 90% credibility intervals (q5% and q95%) are indicated for each admixture parameter. Estimations assuming the prior set 1 and the prior set 2 (table S8) are provided. The alternative set of representative sample sites is the same as the one described in the legend of table S6.

Interpretation	Parameter	Prior distribution set 1	Prior distribution set 2
Stable effective population size	N_i	Uniform [10,000 ; 1,000,000]	Normal [10,000 ; 1,000,000]
Divergence time for the ancestral unsampled "ghost" population	t_{gi}	Uniform [0 ; 20,000]	Normal [25,000 ; 12,500]
Introduction times (in generations)	t_{US-Haw}	Uniform [400 ; 450]	Normal [400 ; 450]
	$t_{US-Wat, SP-Bar}$	Uniform [80 ; 130]	Normal [80 ; 130]
	$t_{US-Sok, US-SD, IT-Tre, FR-Mon}$	Uniform [70 ; 120]	Normal [70 ; 120]
	$t_{US-NC, US-Wis, GE-Dos}$	Uniform [60 ; 110]	Normal [60 ; 110]
	$t_{US-Gen, FR-Par, SW-Del, FR-Bar}$	Uniform [50 ; 100]	Normal [50 ; 100]
	t_{US-Col}	Uniform [40 ; 90]	Normal [40 ; 90]
	$t_{BR-PA, FR-Reu}$	Uniform [10 ; 60]	Normal [10 ; 60]
	Bottleneck durations	DB_i	Uniform [1 ; 200]
Founding effective population sizes	Nb_i	Uniform [2 ; 1000]	Normal [2 ; 1000]
Admixture rate	r_i	Uniform [0.05 ; 0.95]	Normal [0.500 ; 0.225]
Mutation parameters for microsatellites	μ	Uniform [1E-6; 1E-5]	LogUniform [1E-6; 1E-5]
	P	Uniform [0; 3E-1]	LogUniform [0; 3E-1]
	SNI	Uniform [1E-8; 1E-5]	LogUniform [1E-8; 1E-5]

Table S8. Prior sets 1 and 2: historical, demographic and mutation parameters used for ABC analyses, with their interpretation and their prior distributions.

Note: For a population i , effective population sizes (N_i) are expressed in number of diploid individuals and times of events (t) in number of generations going back to the past. t_{gi} corresponds to the time of split in number of generations of the unsampled "ghost" population from an ancestral Asian

population. Bottleneck durations DB for a population i indicates the time at which the population will undergo reduction in its effective size from a founding effective population size NB . We set the admixture rate r as the proportion of genes received by a population i from two source populations providing r and $1-r$ genes (table 4). The microsatellite loci were assumed to follow a generalized stepwise mutation model with a possible range of 40 contiguous allelic states, and characterized by three parameters: the mean mutation rate (μ), the mean parameter of the geometric distribution of mutation length (P), and the mean mutation rate for single nucleotide instability (SNI). Each locus is characterized by individual μ_{loc} , P_{loc} and μSNI_{loc} drawn from Gamma distributions. Note that for Uniform $[x; y]$ and logUniform $[x; y]$, x and y are the bounds of the distributions, whereas than for Normal $[x; y]$, x is the mean and y the standard deviation. Prior distribution set 2 was used in order to evaluate the sensitivity of ABC inferences to prior assumptions.

Supplementary material for the manuscript: “Phenotypic plasticity of *Drosophila suzukii* wing to developmental temperature: implications for flight performance”

Antoine Fraimout, Pauline Jacquemart, David J. Aponte, Thierry Descamps, Anthony Herrel, Raphaël Cornette & Vincent Debat.

Figure S1. Schematic representation of the flight acquisition setup.

Figure S2. Graphical representation of the low-pass quintic spline trajectory smoothing.

Table S1. Results from the MANOVA on the effect of temperature and population on wing shape.

Table S2. Results from the MANOVA on the non-allometric part of wing shape.

Table S3. Results from the MANOVA on the effect of temperature and population on flight performance.

Table S4. Results from the permutational MANOVA on the effect of temperature and population on flight performance.

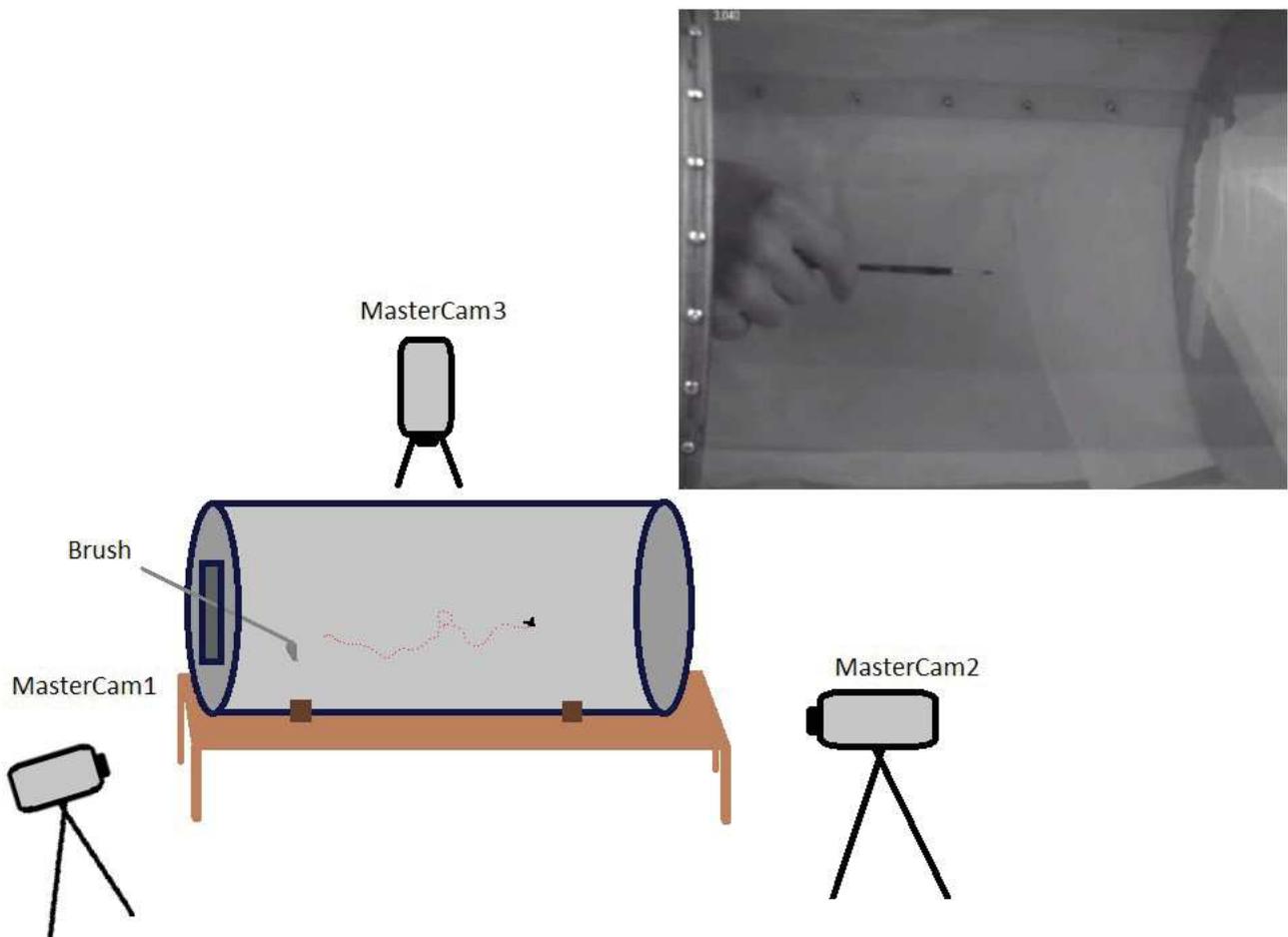


Figure S1. Schematic representation of the flight acquisition setup.

Individual flies were released in the flight chamber and recorded with 3 high-speed cameras. A thin brush was used for non-responsive individuals.

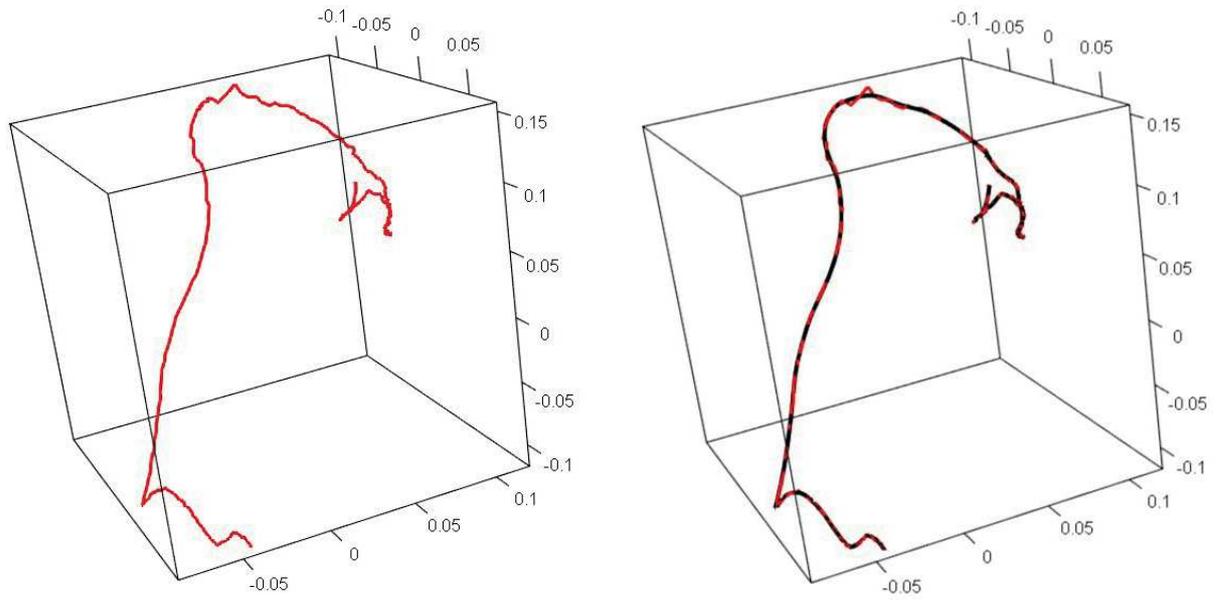


Figure S2. Graphical representation of the low-pass quintic spline trajectory smoothing. The red curve represent the raw trajectory of a single individual for a single flight. Black curve represent the smoothed trajectory.

	Df	Pillai	approx F	num Df	den Df	Pr(>F)	
pop	2	1.31	4.06	52	112	2.708e-10	***
temp	2	1.28	3.83	52	112	1.440e-09	***
pop:temp	4	1.28	1.04	104	232	0.39	
Residuals	80						

Table S1. Results from the MANOVA on the effect of temperature and population on wing shape.

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
pop	2	0.590	0.997	52	124	0.492
Residuals	86					

Table S2. Results from the MANOVA on the non-allometric part of wing shape.

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
temp	2	0.471	1.987	22	142	0.009 **
pop	2	0.381	1.517	22	142	0.077 .
temp:pop	4	0.782	1.614	44	292	0.012 *
Residuals	80					

Table S3. Results from the MANOVA on the effect of temperature and population on flight performance.

	Df	SumOfSqs	MeanSqs	F.model	R2	Pr(>F)
temp	1	16.275	16.275	5.258	0.058	0.003 **
pop	2	4.636	2.318	0.749	0.017	0.636
temp:pop	2	3.206	1.603	0.518	0.011	0.804
Residuals	83	256.933	3.096	0.914		
Total	88	281.050	1.000			

Table S4. Results from the permutational MANOVA on the effect of temperature and population on flight performance.

RESEARCH ARTICLE

Wolbachia in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies

Julien Cattel¹*, Rupinder Kaur^{2,3}*, Patricia Gibert¹, Julien Martinez⁴, Antoine Fraimout⁵, Francis Jiggins⁴, Thibault Andrieux¹, Stefanos Siozios^{2,6}, Gianfranco Anfora², Wolfgang Miller³, Omar Rota-Stabelli^{2‡}, Laurence Mouton^{1‡}

1 Université de Lyon, Université Lyon1, Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558, 43 Bd du 11 Novembre 1918, 69622 Villeurbanne Cedex, France, **2** Chemical Ecology lab, Department of Sustainable Agro-ecosystems and Bio-resources, Fondazione Edmund Mach, San Michele all'Adige (TN), Italy, **3** Laboratory of Genome Dynamics, Department of Cell and Developmental Biology, Center for Anatomy and Cell Biology, Medical University of Vienna, Vienna, Austria, **4** Department of Genetics, University of Cambridge, Cambridge, United Kingdom, **5** Institut de Systématique, Évolution, Biodiversité ISYEB - UMR 7205 – CNRS, MNHN, UPMC, EPHE Muséum national d'Histoire naturelle, Sorbonne Universités 57 rue Cuvier, CP50F-75005, Paris, France, **6** Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom



OPEN ACCESS

Citation: Cattel J, Kaur R, Gibert P, Martinez J, Fraimout A, Jiggins F, et al. (2016) *Wolbachia* in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies. PLoS ONE 11(1): e0147766. doi:10.1371/journal.pone.0147766

Editor: Wolfgang Arthofer, University of Innsbruck, AUSTRIA

Received: October 6, 2015

Accepted: January 7, 2016

Published: January 25, 2016

Copyright: © 2016 Cattel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded partly by CNRS (IFR41-UMR5558) and Office National de l'Eau et des Milieux Aquatiques (ONEMA) (Ecophyto project) for France, and FEM Accordo di Programma Provincia Autonoma di Trento and research funds for GrandiProgetti, Project LEXEM (Laboratory of excellence for epidemiology and modeling, <http://www.lexem.eu>), Italy. Julien Cattel is the recipient of a PhD studentship from the Rhône-Alpes region ("ARC Program" Grant) and Rupinder Kaur has obtained a

* These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* juliencattel@gmail.com (JC); rupinderkaur@fmach.it (RK)

Abstract

The invasive pest *Drosophila suzukii* is characterized by a specific fresh-fruit targeting behavior and has quickly become a menace for the fruit economy of newly infested North American and European regions. *D. suzukii* carries a strain of the endosymbiotic bacterium *Wolbachia*, named *wSuz*, which has a low infection frequency and no reproductive manipulation capabilities in American populations of *D. suzukii*. To further understand the nature of *wSuz* biology and assess its utility as a tool for controlling this pest's populations, we investigated the prevalence of *Wolbachia* in 23 European *D. suzukii* populations, and compared our results with those available in American populations. Our data showed a highly variable infection frequency with a mean prevalence of 46%, which is significantly higher than the 17% found in American populations. Based on Multilocus Sequence Typing analysis, a single *wSuz* strain was diagnosed in all European populations of *D. suzukii*. In agreement with American data, we found no evidence of cytoplasmic incompatibility induced by *wSuz*. These findings raise two questions: a) why *Wolbachia* is maintained in field populations of *D. suzukii* and b) what are the selective forces responsible for the variation in prevalence within populations, particularly between European and American continents? Our results provide new insights into the *D. suzukii*-*Wolbachia* association and highlight regional variations that await further investigation and that should be taken into account for using *Wolbachia*-based pest management programs.

PhD fellowship from the FIRS>T (FEM International Research School-Trentino) programme at Fondazione Edmund Mach, Italy.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Drosophila suzukii Matsumura (Diptera: Drosophilidae), first described in Japan in the early 1900s [1], is an invasive pest of Southeast Asian origin. Since its early detection in California (USA), Spain and Italy (Europe) in 2008, *D. suzukii* has rapidly spread through these two continents aided by global trading and absence of niche competitors [2,3,4,5,6]. Although the vast majority of *Drosophila* species are not fruit pests (larvae developing only in damaged or rotten fruits), *D. suzukii* is able to lay eggs on healthy ripening fruits thanks to female's serrated ovipositor [7]. A wide range of soft and stone fruits including raspberry, strawberry, blueberry, plums and grapes come under *D. suzukii*'s damage range [4,8]. Damage is caused by developing larvae inside the fruit, leading to millions of dollars of annual economic losses worldwide [9,10]. Control of *D. suzukii* populations in the field mainly relies on the use of chemical pesticides, a practice with serious drawbacks because of its use close to harvest and the consequent risk of high residues left on fruits. Management agendas are therefore in permanent search for alternative strategies including those based on bio-control [4,11,12,13,14,15].

Previous studies revealed the presence of the bacterium *Wolbachia* in *D. suzukii* (strain named *wSuz*) [16,17,18,19,20] but of no other heritable bacterial symbionts [21]. This is not surprising since *Wolbachia* is widespread among terrestrial arthropods with an estimate of 52% of species infected at variable prevalence [22]. *Wolbachia* is a maternally inherited bacterium that has established a wide range of relationships with their hosts, from mutualism to parasitism. It is particularly known for its ability to manipulate host reproduction through different strategies to maximize its spread and maintenance in host populations [23]. The most common strategy is cytoplasmic incompatibility (CI), a sperm-egg incompatibility expressed in crosses between infected males and uninfected females leading to the death of embryos in diploid species, thereby facilitating *Wolbachia* to spread throughout host populations [24,25]. However, in the absence of CI or other means of reproductive manipulation, *Wolbachia* can also provide direct host fitness benefits (e.g. increasing survival time, enhancing reproduction, provisioning nutrition or protection against viruses), which may sometimes allow infected individuals to outcompete uninfected counterparts [26]. For example, in *D. mauritiana*, *Wolbachia* infection leads to a decrease of apoptosis and an increase of mitotic divisions in germ line stem cells of infected female ovaries, resulting in more eggs produced than in uninfected ovaries [27]. Lack of any of the suggested effects on the host can potentially lead to loss of infection from a population [28,29,30]. Overall, the infection dynamics of *Wolbachia* depend on the bacterial strain, host species, genetic background, environmental conditions and interactions among these factors [31]. Recently, a field survey in North America indicated a low prevalence of *Wolbachia* in field populations of *D. suzukii* from four localities sampled over two years (prevalence ranged between 7 to 58%), with a mean infection rate of 17% [19]. This study showed no signs of CI or any other reproductive phenotype induced by *wSuz* in *D. suzukii*. Moreover, the imperfect transmission of *Wolbachia* from wild-caught females to their offspring (20 to 95%) suggested possible direct fitness benefits provided by *wSuz* to its American *D. suzukii* host but the phenotypic nature of these benefits has not yet been determined.

In order to improve our understanding of the global infection dynamics of *wSuz* in natural *D. suzukii* populations and thereby assess the utility of *Wolbachia* as a tool for controlling this insect pest population in the field, we investigated the prevalence of *Wolbachia* in 23 localities among eight European countries, typed *Wolbachia* strains through a Multilocus Sequence Typing approach and performed mating experiments on two different European populations.

Materials and Methods

Field sampling

561 individuals were sampled between 2010 and 2014 in 23 locations from eight different countries (Austria, France, Germany, Italy, Slovenia, Spain, Switzerland and United Kingdom) of the European continent (Fig 1 and Table 1). Adults were caught directly in orchards or from attractive traps, placed alive in ethanol (96%) and stored at -20°C until DNA extraction. For *Wolbachia* screening, we gave priority to females because they are known to be responsible for *Wolbachia* transmission.



Fig 1. *Wolbachia* prevalence and collection sites for *D. sukuzii* individuals. A number was assigned for each sampling site and details for each locality are given in the Table 1. Reprinted from <http://d-maps.com> under a CC BY license, with permission from Daniel Dalet, original copyright 2007–2016.

doi:10.1371/journal.pone.0147766.g001

Table 1. Sampling information and percentage of individuals infected by *Wolbachia*.

	Country	Locality	Collection period	Total No. (M+F)	No. of infected individuals	Infection rate (95% CI)
1	Austria	Neustift am Walde	Oct, 2014	19 (6+13)	12	0.63 (0.38–0.84)
2	United Kingdom	East Malling	Sep, 2014	24 (6+18)	4	0.17 (0.05–0.37)
3	France	Montauban	Aug, 2010	9 (7+2)	4	0.44 (0.14–0.79)
4	France	Finestret	Sep, 2010	13 (10+3)	8	0.62 (0.32–0.86)
5	France	Santa Maria Poggio	Jan, 2011	43 (25+18)	12	0.28 (0.15–0.44)
6	France	Mirabel	Aug, 2011	21 (15+6)	6	0.29 (0.11–0.52)
7	France	Bellegarde	Oct, 2012	17 (0+17)	0	0.00 (0.00–0.20)
8a	France	Mougins	Oct, 2012	13 (0+13)	6	0.46 (0.19–0.75)
8b	France	Mougins	Oct, 2013	28 (0+28)	16	0.57 (0.37–0.76)
9	France	Chaussan	Oct, 2013	10 (0+10)	3	0.30 (0.07–0.65)
10	France	Montpellier	Oct, 2013	19 (0+19)	11	0.58 (0.33–0.80)
11	France	Sauternes	Jan, 2014	30 (0+30)	21	0.70 (0.51–0.85)
12	France	Prignonrieux	Jan, 2014	44 (0+44)	16	0.36 (0.22–0.52)
13	France	Concourson sur Layon	Jun, 2014	8 (0+8)	4	0.50 (0.16–0.84)
14	France	Gotheron	Jul, 2014	17 (5+12)	6	0.35 (0.14–0.62)
15	France	Carrière sur Seine	Jan, 2014	35 (15+20)	8	0.23 (0.10–0.40)
16	France	Saint Germain d'Esteuil	Jan, 2014	22 (0+22)	8	0.36 (0.17–0.59)
17	Germany	Dossenheim	Oct, 2013	41 (0+41)	22	0.54 (0.37–0.59)
18	Italy	Vigolo Vattaro	Sep, 2013	16 (0+16)	16	1.00 (0.79–1.00)
19a	Italy	San Michele	Sep, 2013	9 (0+9)	2	0.22 (0.15–0.59)
19b	Italy	San Michele	Sep, 2014	20 (8+12)	13	0.65 (0.41–0.85)
20	Italy	Bari	Feb, 2014	11 (0+11)	8	0.73 (0.39–0.94)
21	Slovenia	Izola	Oct, 2013	35 (0+35)	16	0.46 (0.29–0.63)
22	Spain	Girona	Mar, 2014	37 (3+34)	29	0.85 (0.62–0.90)
23	Switzerland	Gottfrey	Oct, 2013	20 (0+20)	7	0.35 (0.15–0.59)

The sample number corresponds to the code of the sampling indicated in Fig 1. CI: 95% confidence interval.

doi:10.1371/journal.pone.0147766.t001

DNA extraction and *Wolbachia* detection by Polymerase Chain Reaction (PCR)

DNA from whole individuals was extracted using the Nucleo Spin Tissue kit (Macherey-Nagel). The quality of extraction was checked by PCR targeting the arthropod-specific locus Internal Transcribed Spacer *ITS2* gene (*Its2U/Its2L*) [32]. *Wolbachia* detection was performed using *Wolbachia* specific primers for two different genes: *ftsZ*, a bacterial cell division gene (*F2/R2*) [33], and *wsp* (*Wolbachia surface protein*) (*81F/691R*) [34]. PCR reactions were performed in 25 µL volumes containing 100 µM dNTP, 200 nM primers, 20mM Dream Taq Green Buffer, 0.5 IU *Taq* DNA polymerase (Eurobio) and 1µL of DNA template. Cycling conditions were 94°C (2 min), 94°C (30 sec), 55°C (*ITS2* and *ftsZ* primers) or 52°C (*wsp* primers) (30 sec), 72°C (45 sec), 72°C (10 min) for 34 cycles. All the primer sequences were indicated in S1 File. PCR products were visualized in 1% agarose gels. An individual was considered infected when the two *Wolbachia*-specific primers produced fragments of the appropriate size. To reduce the possibility of false negatives, all samples for which we failed to amplify *Wolbachia* signal with normal PCR were verified by a more sensitive real-time quantitative PCR method using the *wsp* primers 81F/691R. The 10µL reaction mixture contained 200 nM of each primer, 5 µL of Light Cycler® 480 SYBR Green I Master (Roche) and 1 µL of DNA sample. The amplification consisted of 10 min at 95°C followed by 40 cycles of 10 sec at 95°C, 20 sec at 53°C and 30 sec at

72°C. Only for two individuals, *Wolbachia* infection was not detected by classic PCR but detected by real-time quantitative PCR.

MLST approach

PCRs targeting six genes of *Wolbachia* (*wsp*, *gatB*, *coxA*, *hcpA*, *ftsZ*, *fbpA*) [34,35] were carried out on two randomly selected *wSuz*-infected samples per locality (Fig 1 and Table 1). All PCR reactions were performed as described before. Annealing temperature was 55°C except for *wsp* (52°C) and all the PCR products were sequenced. Sequences were aligned using MUSCLE [36] algorithm implemented in CLC DNA Workbench 6.9.1. (CLC Bio) and inspected by eye. On an average, we sequenced and compared approximately 3200 nucleotides per individual (between 470 and 610 nt per gene). All primers used in this study are presented in S1 File.

Cytoplasmic Incompatibility (CI) assays

Two different populations of *D. sukii* were used for CI experiments, one from France (Compiègne) and another from Italy (Valsugana). The French population was sampled in 2011 and reared in mass population in Lyon on a cornmeal diet (agar: 1%, dextrose: 8.75%, maize: 8.75%, yeast extract: 2%, nipagin: 3%) under constant lab conditions of 21±1°C temperature with a 12 hours light/dark cycle at 70% relative humidity. The Italian population was established from individuals collected in Valsugana region in 2011, and subsequently reared in laboratory at San Michele all'Adige on standard corn meal diet at a temperature of 23±1°C, 65% relative humidity with a 12:12 light/dark cycle. Maternal transmission of *Wolbachia* in the laboratory-established lines was perfect and thus the infection was stably maintained in the lab. Before starting the CI experiment, both lines were acclimatized under the same laboratory conditions for three generations in Cambridge Lab at a constant temperature of 22°C and 65% relative humidity with 12:12 light/dark cycle, thereby avoiding any bias due to lab environments.

To obtain *Wolbachia*-infected and uninfected fly lines with the same genetic background, antibiotic treatments were performed on infected lines using 0.25mg mL⁻¹ tetracycline for three consecutive generations in mass populations. After treatments, flies were fed on normal standard diet for two further generations to recover them from any side effect caused by antibiotic treatment. Due to different lab practices, flies from Italian populations were further allowed to feed on natural fecal material from infected individuals to re-acquire their loss of gut-associated microbiota but flies from France were not gut-flora restored. Ten isofemale lines were established and the presence of *Wolbachia* was checked by PCR in mothers after laying eggs. The absence of *Wolbachia* was re-confirmed by real-time PCR as described above. This was repeated for three generations and then, for each population, one isofemale line was retained for crossing experiments. Their infection status was also checked and confirmed just before the CI assays (n = 20 for each line), which were performed on the 10th generation after antibiotic treatments were stopped.

All types of crosses (infected female x infected male, infected female x uninfected male, uninfected female x infected male and uninfected female x uninfected male) were performed for each genetic background at Cambridge. Freshly hatched individuals were sexed and placed separately into cornmeal diet tubes to ensure the virginity of flies. A 3-days old virgin male and a 5-days old virgin female were allowed to mate in food vial for 24h. Females were then individually allowed to oviposit for 48h on grape-juice agar petri dish. Every 48h, females were transferred to a new petri dish, following 48h period for oviposition, and the total number of eggs along with the number of hatched larvae per female were recorded. Experiment was performed in replicates for each type of cross as mentioned in Table 2. The incompatibility relationship

Table 2. Cytoplasmic incompatibility assays for *D. sukii* from France and Italy.

Origin	Female	Male	N	Number of eggs laid	Mean hatch rate (±sd)
France	UN	UN	21	1113	0.691 (±0.041) b
	UN	IN	13	763	0.404 (±0.087) a
	IN	UN	15	588	0.573 (±0.075) ab
	IN	IN	15	795	0.718 (±0.069) b
Italy	UN	UN	28	1199	0.634 (±0.059) α
	UN	IN	38	2000	0.502 (±0.044) α
	IN	UN	25	1606	0.505 (±0.064) α
	IN	IN	27	1370	0.729 (±0.045) α

doi:10.1371/journal.pone.0147766.t002

was determined by the rate of eggs that hatched. Females that laid less than 10 eggs were not included in the analysis (percentages are indicated in [S2 File](#)).

Statistical analysis

We tested the difference in *Wolbachia* prevalence between American and European continents by fitting a generalized linear model with a quasi-binomial error (data for North America was taken from [19]). Samples originated from the same population but collected in several years were pooled. We tested for the existence of differences between samplings by using Fisher’s exact test. Estimates of the mean prevalence on each continent were transformed from logits into percentages. Statistical analysis of the CI data was done using a generalized linear model (quasi-binomial family). An ANOVA was done on this model and Tukey’s (HSD) tests were used for two by two comparisons. R 3.2.2 version [37] was used to perform all the analysis.

Ethics Statement

Sampling of *D. sukii* was carried out on private lands with owners’ permission. All individuals were sampled in experimental stations and did not require specific permission. The field studies did not involve endangered or protected species.

Results

Variable *w*Suz infection frequencies in *D. sukii* populations

We detected *Wolbachia* in 22 populations out of 23 surveyed from eight European countries ([Fig 1](#)). Out of the 561 *D. sukii* flies tested, 258 were found infected with *Wolbachia* ([Table 1](#)). We found a highly variable infection prevalence (from 0 to 100%) between localities ($\chi^2 = 192$, d.f = 42, $P < 0.001$). The two extreme values (0 and 100) have been observed only in two populations (Bellegarde in France and Vigolo Vattaro in Italy respectively); in all the other localities, prevalence ranged between 15% and 85% with a mean of 46%.

In two European localities, Mougins (in France) and San Michele (in Italy), we sampled for two consecutive years at the same period (October, 2012–13 and September, 2013–14 respectively) and found an increasing trend of the *Wolbachia* prevalence over time for both localities (Mougins: from 0.46 to 0.57; San Michele: from 0.22 to 0.65; [Table 1](#)) though the trend was not significant (Fisher’s exact test, $P = 0.74$ and $P = 0.05$ respectively).

Hamm [19] screened 929 individuals sampled from four American localities of *D. sukii* over two years and found infection frequencies ranging from 7 to 58% with a mean of 17%. Therefore, in both continents, the prevalence of *w*Suz is highly variable. Overall, the European

infection frequencies is found almost three times significantly higher than the ones in North America (GLM: $t = 4.54$, $d.f = 26$, $P < 0.001$).

European *D. sukuzii* flies carry one single sequence type of *Wolbachia* “wSuz”

MLST analysis, assayed from two individuals per population (Fig 1 and Table 1), indicated that all European *D. sukuzii* individuals carry the same *Wolbachia* sequence type (100% nucleotides identity for all genes). Sequences were identical to those of wSuz previously found and characterized in *D. sukuzii* individuals originated from Italy [18] and North America [19]. The sequences obtained in the present study are recorded in Genbank as KS308222-7.

Cytoplasmic incompatibility assays

To test whether wSuz is able to induce CI, which could facilitate its spread throughout European *D. sukuzii* populations, we performed mating experiments on flies originating from France and Italy (Table 2). In diploid species, CI leads to the death of embryos when infected males mate with uninfected females; therefore in case of wSuz induced CI expression, we expected a lower egg hatch rate in these crosses. We found a significant effect of the type of cross on the hatching rates only for the French population ($F = 3.86$, $d.f = 3$, $P = 0.01$; for Italian populations: $F = 2.14$, $d.f = 3$, $P = 0.10$). Globally, the hatching rate was higher in crosses that involved individuals of the same infection status (control crosses) than that of different infection status (infected and uninfected). The “incompatible” crosses (infected male x uninfected female) tended to have the lowest hatching rate. For the French population, the difference was significant only when the incompatible cross was compared to the two control crosses: infected male and female (Tukey HSD test, $P = 0.02$) and uninfected male and uninfected female (Tukey HSD test, $P = 0.01$). For the Italian population there is no significant difference between the incompatible cross and the three others. Therefore, like in the American populations, these results do not allow us to conclude that wSuz can induce CI in European *D. sukuzii* populations.

Mean eggs hatch rate (only females that lay at least 10 eggs) \pm standard deviation (sd). N: number of crosses. IN: infected with *Wolbachia*, UN: uninfected. Statistical analysis was performed on lines from the two origins separately. One-way ANOVA indicated that the hatching rate differed significantly between the mating types. One-way ANOVA was followed by Tukey’s test to compare between crossing types (means marked with the same letter are not significantly different; $P = 0.05$).

Discussion

We found that the *Wolbachia* infection in *D. sukuzii* is highly variable in European populations with a mean prevalence of 46%. This infection rate differs significantly from the one found in the North America (17% in average) [19]. Several parameters could explain this variability such as the efficiency of vertical transmission or the costs/benefits of infection, which might differ between continents. Moreover, variability can also be due to differences in the host and/or *Wolbachia* genetic backgrounds as indicated in other insect species by trans-infection experiments using a given *Wolbachia* strain injected into different hosts (for example, see [38]). Apart from different host-*Wolbachia* association dynamics, the variability can also be dependent on environmental factors (*i.e.* temperature, diet, larvae density. . .) [39,40,28,41,42] that might vary locally and across time with their colonization dynamics, hence providing higher fitness advantage for *Wolbachia* persistence in European *D. sukuzii* populations than in the American ones. If this is true, we should expect persistence at most/all localities to a certain

infection rate representing the fitness advantage based on that particular locality's environmental and habitat specificity. Although we found a (non-significant) increasing trend for the *Wolbachia* prevalence in two localities, more sampling time points over longer durations are needed to draw a clearer picture.

Infection dynamics of vertically transmitted symbionts depend on the phenotypic effect induced on the host, the fitness cost and/or benefits of the infection, and how reliably symbionts are inherited. Our results clearly showed that *wSuz* is not fixed in *D. sukuzii* populations as otherwise expected in cases of obligatory mutualism. We did not detect any significant evidence for CI induction by *wSuz* in Italian or French *D. sukuzii* population. An imperfect vertical transmission of endosymbiotic bacteria (20 to 95% in American populations) [19] associated with an absence of strong reproductive manipulation would theoretically lead to elimination of the symbiont from host populations unless it provides the other host fitness advantages [29]. The stable presence of *wSuz* in all but one of the populations studied in America and Europe would thus suggest that *wSuz* may provide direct benefits to its host allowing its maintenance in natural populations of *D. sukuzii*. Another possibility would be frequent intraspecific horizontal transfers of *Wolbachia* [43]. However, we favored the first hypothesis because the second one would imply that horizontal transfers occur in all populations from two continents, which seems unlikely. In addition, *Wolbachia* are also known as a mutualist in insects [44] providing beneficial effects to their native hosts, such as an increase of host survival [45] or fecundity under nutritional stress [46], or protection against viral, microbial, and fungal pathogens [47,48,49,50,51,52]. While no fitness benefits have been detected in American populations [19], a recent study suggested a beneficial effect of *Wolbachia* infection on female fecundity in an Italian population of *D. sukuzii* [53]. Since both of the studies used different protocols to perform the assays, different lab practices might impair the significance of this difference. For example, Hamm [19] crossed native *Wolbachia* infected and uninfected flies from the wild: thereby we cannot exclude the confounding effect of different host genetic backgrounds on the fly fecundity. On the other hand, Mazzetto [53] compared infected individuals to uninfected ones that were immediately obtained after antibiotic treatments. Therefore the reduced fecundity of uninfected females can also be explained by side effects of the antibiotics on mitochondria and/or host metabolism [54]. Although with caution, excluding any of these confounding factors suggest that the beneficial effect induced by *wSuz* may depend on the fly genetic background and/or that *D. sukuzii* from Europe and North America are infected by different *Wolbachia* strains. Highly sophisticated approaches, for example, Next Generation Sequencing (NGS) based tools are needed to have more insights into the *Wolbachia* strain diversity in *D. sukuzii* populations. Moreover, future experiments involving exchange or introgression of European and American *Wolbachia* strains to study the life history traits of host organisms in the presence and absence of *Wolbachia* would help to unveil whether differences in *wSuz*-mediated fitness benefits in different continental *D. sukuzii* populations are *Wolbachia* strain-dependent or host dependent.

There is a growing interest in exploring symbiotic microorganisms for biocontrol management programs [55,56,57]; in the case of *Wolbachia*, the idea is to exploit its capability to induce CI to control natural populations of arthropod pests in an Incompatible Insect Technique (IIT) fashion [58,59], where males infected with a CI-inducing strain of *Wolbachia* are released in the field to mate with incompatible females leading to embryo mortality and ultimate population suppression. This procedure is somewhat analogous to the Sterile Insect Technique (SIT), a species-specific method of insect control that relies on the release of large numbers of sterile males instead of incompatible males [60], but the advantage is that insects do not have to be irradiated or genetically modified before release. The inconvenience is, however, that IIT and SIT require solid methods allowing efficient males and females separation,

but *D. sukuzii* males are easily recognizable because of their spotted wings [1]. Considerable research efforts have already demonstrated that *Wolbachia*-inducing CI could be used as a tool for population control [61,62,63]. However, it might be more complicated if *Wolbachia* naturally infects the insect pest like in the case of *D. sukuzii* since it would require finding an incompatible CI-inducing *Wolbachia* strain, which is not rescued by *wSuz* in its original host. Hamm [19] and our results are concordant in showing that *D. sukuzii* is infected by only one strain of *Wolbachia*, which is present in almost all populations screened and does not induce CI. Trans-infection experiments using closely related *Wolbachia* strains such as the highly CI-inducing *wRi* strain from *D. simulans* will thus be needed to assess modification and rescue capabilities of different strain combinations towards a bidirectional CI-based control strategy.

Supporting Information

S1 File. Primers used in this study.

(PDF)

S2 File. Percentage of excluded females that laid less than 10 eggs in crosses experiment.

(PDF)

Acknowledgments

We thank Fabrice Vavre and Sylvain Charlat for useful discussion and H el ene Henri for help in molecular biology. We thank Adriana Escudero (Spain), Gabrijel Seljak (Slovenia), Marco Valerio Rossi Stacconi (Italy), Darren Obbard (UK), Traude Kehrer (Austria), Heidrum Vogt, Serge Fischer, Lionel Delbac, Nicolas Ris, Christophe Plantamp and the CTIFL from Lanxade and Balandran for providing us *D. sukuzii* from the field. We are also grateful to the INRA centre of Gotheron and to the landowners who allowed us to collect insects in their orchards.

Author Contributions

Conceived and designed the experiments: JC RK PG JM FJ SS WM ORS LM. Performed the experiments: JC RK JM TA. Analyzed the data: JC RK JM FJ ORS. Contributed reagents/materials/analysis tools: PG FJ SS WM ORS LM AF GA. Wrote the paper: JC RK PG JM FJ SS WM ORS LM.

References

1. Kanzawa T. Studies on *Drosophila sukuzii* mats. Rev Appl Entomol. 1936; 24:315.
2. Hauser M. A historic account of the invasion of *Drosophila sukuzii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. Pest Manag Sci. 2011; 67:1352–1357. doi: [10.1002/ps.2265](https://doi.org/10.1002/ps.2265) PMID: [21898759](https://pubmed.ncbi.nlm.nih.gov/21898759/)
3. Calabria G, M aca J, B achli G, Serra L, Pascual M. First records of the potential pest species *Drosophila sukuzii* (Diptera: Drosophilidae) in Europe. J Appl Entomol. 2012; 136(1):139–147.
4. Cini A, Ioriatti C, Anfora G. A review of the invasion of *Drosophila sukuzii* in Europe and a draft research agenda for integrated pest management. Bull Insectology. 2012; 65(1):149–160.
5. Rota-Stabelli O, Blaxter M, Anfora G. *Drosophila sukuzii*. Curr Biol. 2013; 23(1):R8–R9. doi: [10.1016/j.cub.2012.11.021](https://doi.org/10.1016/j.cub.2012.11.021) PMID: [23305672](https://pubmed.ncbi.nlm.nih.gov/23305672/)
6. Cini A, Anfora G, Escudero-Colomar LA, Grassi A, Santosuosso U, Seljak G, et al. Tracking the invasion of the alien fruit pest *Drosophila sukuzii* in Europe. J Pest Sci. 2014; 87:559–566.
7. Mitsui H, Takahashi KH, Kimura MT. Spatial distributions and clutch sizes of *Drosophila* species ovipositing on cherry fruits of different stages. Popul Ecol. 2006; 48(3):233–237.
8. Dreves A, Walton V, Fisher G. A New Pest Attacking Healthy Ripening Fruit in Oregon. Oregon Univ Serv. 2009; EM 8991:1–6.

9. Goodhue RE, Bolda M, Farnsworth D, Williams JC, Zalom FG. Spotted wing drosophila infestation of California strawberries and raspberries: economic analysis of potential revenue losses and control costs. *Pest Manag Sci*. 2011; 67(11):1396–1402. doi: [10.1002/ps.2259](https://doi.org/10.1002/ps.2259) PMID: [21815244](https://pubmed.ncbi.nlm.nih.gov/21815244/)
10. Lee JC, Bruck DJ, Dreves AJ, Ioriatti C, Vogt H, Baufeld P. In focus: spotted wing drosophila, *Drosophila sukuzii*, across perspectives. *Pest Manag Sci*. 2011; 67:1349–1351. doi: [10.1002/ps.2271](https://doi.org/10.1002/ps.2271) PMID: [21990168](https://pubmed.ncbi.nlm.nih.gov/21990168/)
11. Chabert S, Allemand R, Poyet M, Eslin P, Gibert P. Ability of European parasitoids (Hymenoptera) to control a new invasive Asiatic pest, *Drosophila sukuzii*. *Biol Control*. 2012; 63(1):40–47.
12. Gargani E, Tarchi F, Frosinini R, Mazza G, Simoni S. Notes on *Drosophila sukuzii* Matsumura (Diptera Drosophilidae): field survey in Tuscany and laboratory evaluation of organic products. *Redia*. 2013; 96:85–90.
13. Poyet M, Havard S, Prevost G, Chabrerie O, Doury G, Gibert P, et al. Resistance of *Drosophila sukuzii* to the larval parasitoids *Leptopilina heterotoma* and *Asobara japonica* is related to haemocyte load. *Physiol Entomol*. 2013; 38(1):45–53.
14. Cuthberston AGS, Collins DA, Blackburn LF, Audsley NA, Bell HA. Preliminary screening of potential control products against *Drosophila sukuzii*. *Insects*. 2014; 5:488–498. doi: [10.3390/insects5020488](https://doi.org/10.3390/insects5020488) PMID: [26462696](https://pubmed.ncbi.nlm.nih.gov/26462696/)
15. Woltz JM, Donahue KM, Bruck DJ, Lee JC. Efficacy of commercially available predators, nematodes and fungal entomopathogens for augmentative control of *Drosophila sukuzii*. *J Appl Entomol*. 2015. doi: [10.1111/jen.12200](https://doi.org/10.1111/jen.12200)
16. Cordaux R, Pichon S, Ling A, Pérez P, Delaunay C, Vavre F, et al. Intense transpositional activity of insertion sequences in an ancient obligate endosymbiont. *Mol Biol Evol*. 2008; 25(9):1889–1896. doi: [10.1093/molbev/msn134](https://doi.org/10.1093/molbev/msn134) PMID: [18562339](https://pubmed.ncbi.nlm.nih.gov/18562339/)
17. Ometto L, Cestaro A, Ramasamy S, Grassi A, Revadi S, Siozios S, et al. Linking genomics and ecology to investigate the complex evolution of an invasive *Drosophila* pest. *Genome Biol Evol*. 2013; 5(4):745–757. doi: [10.1093/gbe/evt034](https://doi.org/10.1093/gbe/evt034) PMID: [23501831](https://pubmed.ncbi.nlm.nih.gov/23501831/)
18. Siozios S, Cestaro A, Kaur R. Draft genome sequence of the *Wolbachia* endosymbiont of *Drosophila sukuzii*. *Genome Announc*. 2013; 1(1):1–2.
19. Hamm CA, Begun DJ, Vo A, Smith CCR, Saelao P, Shaver AO, et al. *Wolbachia* do not live by reproductive manipulation alone: infection polymorphism in *Drosophila sukuzii* and *D. subpulchrella*. *Mol Ecol*. 2014; 23:4871–4885. doi: [10.1111/mec.12901](https://doi.org/10.1111/mec.12901) PMID: [25156506](https://pubmed.ncbi.nlm.nih.gov/25156506/)
20. Tochen S, Dalton DT, Wiman N, Hamm C, Shearer PW, Walton VM. Temperature-Related Development and Population Parameters for *Drosophila sukuzii* (Diptera: Drosophilidae) on Cherry and Blueberry. *Environ Entomol*. 2014; 43(2):501–510. doi: [10.1603/EN13200](https://doi.org/10.1603/EN13200) PMID: [24612968](https://pubmed.ncbi.nlm.nih.gov/24612968/)
21. Chandler JA, James PM, Jospin G, Lang JM. The bacterial communities of *Drosophila sukuzii* collected from undamaged cherries. *PeerJ*. 2014; 2:e474. doi: [10.7717/peerj.474](https://doi.org/10.7717/peerj.474) PMID: [25101226](https://pubmed.ncbi.nlm.nih.gov/25101226/)
22. Weinert LA, Araujo EV, Ahmed MZ, Welch JJ. The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc R Soc B*. 2015; 282:1807. doi: [10.1098/rspb.2015.0249](https://doi.org/10.1098/rspb.2015.0249)
23. Werren JH. Biology of *Wolbachia*. *Annu Rev Entomol*. 1997; 42(124):587–609.
24. Dobson SL, Marsland EJ, Rattanadechakul W. Mutualistic *Wolbachia* infection in *Aedes albopictus*: accelerating cytoplasmic drive. *Genetics*. 2002; 160(3):1087–1094. PMID: [11901124](https://pubmed.ncbi.nlm.nih.gov/11901124/)
25. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol*. 2008; 6(10):741–751. doi: [10.1038/nrmicro1969](https://doi.org/10.1038/nrmicro1969) PMID: [18794912](https://pubmed.ncbi.nlm.nih.gov/18794912/)
26. Schneider EL, Miller WJ, Riegler M. Arthropods shopping for *Wolbachia*. *Manipulative Tenants: Bacteria Associated with Arthropods*. CRC Press. 2011; 149–174.
27. Fast EM, Toomey ME, Panaram K, Desjardins D, Kolaczyk ED, Frydman HM. *Wolbachia* enhance *Drosophila* stem cell proliferation and target the germline stem cell niche. *Science*. 2011; 334:990–992. doi: [10.1126/science.1209609](https://doi.org/10.1126/science.1209609) PMID: [22021671](https://pubmed.ncbi.nlm.nih.gov/22021671/)
28. Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA. From parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PLoS Biol*. 2007; 5(5):0997–1005.
29. Hoffmann AA, Hercus M, Dagher H. Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. *Genetics*. 1998; 148(1):221–31. PMID: [9475734](https://pubmed.ncbi.nlm.nih.gov/9475734/)
30. Turelli M, Hoffmann AA. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. *Genetics*. 1995; 140(4):1319–1338. PMID: [7498773](https://pubmed.ncbi.nlm.nih.gov/7498773/)
31. Jaenike J. Coupled population dynamics of endosymbionts within and between hosts. *Oikos*. 2009; 118(3):353–362.
32. Campbell HD, Schimansky T, Claudianos C, Ozsarac N, Kasprzak AB, Cotsell JN, et al. The *Drosophila melanogaster* flightless-I gene involved in gastrulation and muscle degeneration encodes gelsolin-like

- and leucine-rich repeat domains and is conserved in *Caenorhabditis elegans* and humans. *Proc Natl Acad Sci of the U S A*. 1993; 90(23):11386–11390.
33. Vavre F, Fleury F, Lepetit D, Fouillet P, Boulétreau M. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Mol Biol Evol*. 1999; 16(12):1711–1723. PMID: [10605113](#)
 34. Braig HR, Zhou W, Dobson SL, O'Neill SL. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J Bacteriol*. 1998; 180(9):2373–2378. PMID: [9573188](#)
 35. Baldo L, Hotopp JCD, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR, et al. Multilocus Sequence Typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol*. 2006; 72(11):7098–7110. PMID: [16936055](#)
 36. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32(5):1792–1797. PMID: [15034147](#)
 37. R Core Team. R: a language and environment for statistical computing. 2013.
 38. Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Frentiu FD, McMeniman CJ. The *wMel* *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature*. 2011; 476:450–453. doi: [10.1038/nature10355](#) PMID: [21866159](#)
 39. Reynolds KT, Hoffmann AA. Male age, host effects and the weak expression or non-expression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. *Genet Res*. 2002; 80(2):79–87. PMID: [12534211](#)
 40. Mouton L, Henri H, Bouletreau M, Vavre F. Effect of temperature on *Wolbachia* density and impact on cytoplasmic incompatibility. *Parasitology*. 2006; 132(1):49–56.
 41. Yamada R, Floate KD, Riegler M, O'Neill SL. Male development time influences the strength of *Wolbachia*-induced cytoplasmic incompatibility expression in *Drosophila melanogaster*. *Genetics*. 2007; 177(2):801–808. PMID: [17660578](#)
 42. Pontier SM, Schweisguth F. A *Wolbachia*-sensitive communication between male and female pupae controls gamete compatibility in *Drosophila*. *Curr Biol*. 2015; 25(18):2339–2348. doi: [10.1016/j.cub.2015.07.052](#) PMID: [26344089](#)
 43. Morrow JL, Frommer M, Royer JE, Shearman DCA, Riegler M. *Wolbachia* pseudogenes and low prevalence infections in tropical but not temperate Australian tephritid fruit flies: manifestations of lateral gene transfer and endosymbiont spillover? *BMC Evol Biol*. 2015; 15(202).
 44. Zug R, Hammerstein P. Bad guys turned nice? A critical assessment of *Wolbachia* mutualisms in arthropod hosts. *Biol Rev*. 2015; 90:89–111. doi: [10.1111/brv.12098](#) PMID: [24618033](#)
 45. Fry AJ, Rand DM. *Wolbachia* interactions that determine *Drosophila melanogaster* survival. *Evolution*. 2002; 56(10):1976–1981. PMID: [12449484](#)
 46. Brownlie JC, Johnson KN. Symbiont-mediated protection in insect hosts. *Trends in Microbiol*. 2009; 17(8):348–354.
 47. Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. *Wolbachia* and virus protection in insects. *Science*. 2008; 322(5902):702. doi: [10.1126/science.1162418](#) PMID: [18974344](#)
 48. Teixeira L, Ferreira Á, Ashburner M. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol*. 2008; 6(12):2753–2763.
 49. Osborne SE, Leong YS, O'Neill SL, Johnson KN. Variation in antiviral protection mediated by different *Wolbachia* strains in *Drosophila simulans*. *PLoS Pathog*. 2009; 5(11):e1000656. doi: [10.1371/journal.ppat.1000656](#) PMID: [19911047](#)
 50. Glaser RL, Meola MA. The native *Wolbachia* endosymbionts of *Drosophila melanogaster* and *Culex quinquefasciatus* increase host resistance to west Nile virus infection. *PLoS One*. 2010; 5(8):e11977. doi: [10.1371/journal.pone.0011977](#) PMID: [20700535](#)
 51. Ferreira ÁG, Naylor H, Esteves SS, Teixeira L. The Toll-Dorsal pathway is required for resistance to viral oral infection in *Drosophila*. *PLoS Pathog*. 2014; 10(12):e1004507. doi: [10.1371/journal.ppat.1004507](#) PMID: [25473839](#)
 52. Martinez J, Longdon B, Bauer S, Chan Y-S, Miller WJ, Bourtzis K, et al. Symbionts commonly provide broad spectrum resistance to viruses in insects: a comparative analysis of *Wolbachia* strains. *PLoS Pathog*. 2014; 10(9):e1004369. doi: [10.1371/journal.ppat.1004369](#) PMID: [25233341](#)
 53. Mazzetto F, Gonella E, Alma A. *Wolbachia* infection affects female fecundity in *Drosophila sukuii*. *Bull Insectology*. 2015; 68(1): 153–157.
 54. Ballard JWO, Melvin RG. Tetracycline treatment influences mitochondrial metabolism and mtDNA density two generations after treatment in *Drosophila*. *Insect Mol Biol*. 2007; 16(6):799–802. PMID: [18093008](#)

55. Douglas AE. Symbiotic microorganisms: untapped resources for insect pest control. *Trends in Biotechnol.* 2007; 25(8):338–342.
56. Iturbe-Ormaetxe I, Walker T, O'Neill SL. *Wolbachia* and the biological control of mosquito-borne disease. *EMBO Rep.* 2011; 12(6):508–518. doi: [10.1038/embor.2011.84](https://doi.org/10.1038/embor.2011.84) PMID: [21546911](https://pubmed.ncbi.nlm.nih.gov/21546911/)
57. Bourtzis K. *Wolbachia*-based technologies for insects pest population control. *Adv Exp Med Biol.* 2008; 627:104–113. doi: [10.1007/978-0-387-78225-6_9](https://doi.org/10.1007/978-0-387-78225-6_9) PMID: [18510018](https://pubmed.ncbi.nlm.nih.gov/18510018/)
58. Boller EF, Russ K, Vallo V, Bush GL. Incompatible races of European cherry fruit fly, *Rhagoletis cerasi* (Diptera: Tephritidae), their origin and potential use in biological control. *Entomol Exp Appl.* 1976; 20(3):237–247.
59. Riegler M, Stauffer C. *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). *Mol Ecol.* 2002; 11(11):2425–2434. PMID: [12406252](https://pubmed.ncbi.nlm.nih.gov/12406252/)
60. Dyck VA, Hendrichs J, Robinson AS. *Sterile Insect Technique: principles and practice in area-wide integrated pest management.* Springer. 2005; 978:4020–4051.
61. Zabalou S, Riegler M, Theodorakopoulou M, Stauffer C, Savakis C, Bourtzis K. *Wolbachia*-induced cytoplasmic incompatibility as a means for insect pest population control. *Proc Natl Acad Sci U S A.* 2004; 101(42):15042–15045. PMID: [15469918](https://pubmed.ncbi.nlm.nih.gov/15469918/)
62. O'Connor L, Plichart C, Sang AC, Brelsfoard CL, Bossin HC, Dobson SL. Open release of male mosquitoes infected with a *Wolbachia* biopesticide: field performance and infection containment. *PLoS Negl Trop Dis.* 2012; 6(11):e1797. doi: [10.1371/journal.pntd.0001797](https://doi.org/10.1371/journal.pntd.0001797) PMID: [23166845](https://pubmed.ncbi.nlm.nih.gov/23166845/)
63. Atyame CM, Cattell J, Lebon C, Flores O, Dehecq J-S, Weill M, et al. *Wolbachia*-based population control strategy targeting *Culex quinquefasciatus* mosquitoes proves efficient under semi-field conditions. *Plos One.* 2015; 10(3):e0119288. doi: [10.1371/journal.pone.0119288](https://doi.org/10.1371/journal.pone.0119288) PMID: [25768841](https://pubmed.ncbi.nlm.nih.gov/25768841/)

CORRECTION

Correction: *Wolbachia* in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies

Julien Cattel, Rupinder Kaur, Patricia Gibert, Julien Martinez, Antoine Fraimout, Francis Jiggins, Thibault Andrieux, Stefanos Siozios, Gianfranco Anfora, Wolfgang Miller, Omar Rota-Stabelli, Laurence Mouton

The image for [Fig 1](#) is incorrect. Please see the corrected [Fig 1](#) here.



OPEN ACCESS

Citation: Cattel J, Kaur R, Gibert P, Martinez J, Fraimout A, Jiggins F, et al. (2016) Correction: *Wolbachia* in European Populations of the Invasive Pest *Drosophila suzukii*: Regional Variation in Infection Frequencies. PLoS ONE 11(2): e0150050. doi:10.1371/journal.pone.0150050

Published: February 19, 2016

Copyright: © 2016 Cattel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

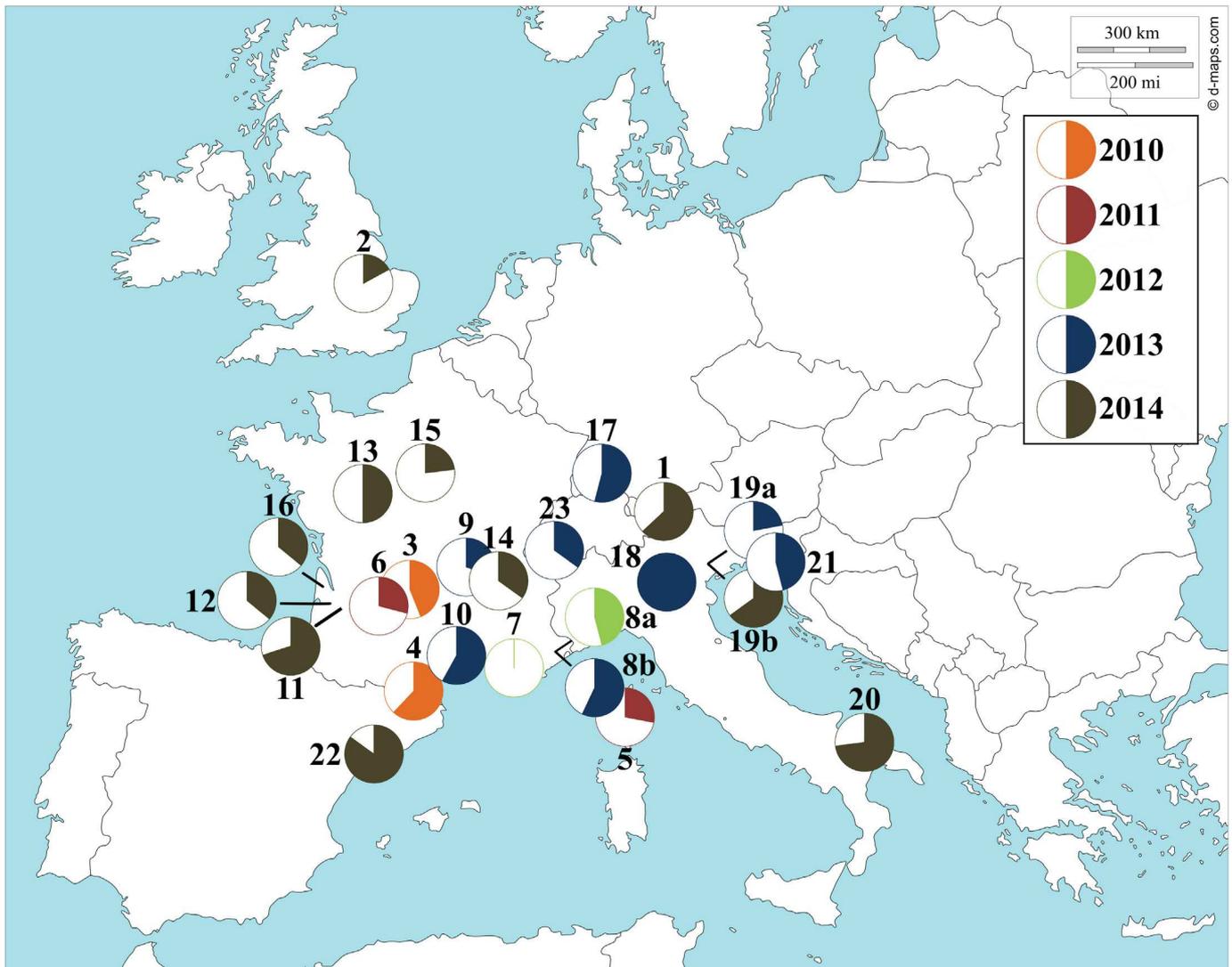


Fig 1. *Wolbachia* prevalence and collection sites for *D. sukuii* individuals. A number was assigned for each sampling site and details for each locality are given in the Table 1. Reprinted from <http://d-maps.com> under a CC BY license, with permission from Daniel Dalet, original copyright 2007–2016.

doi:10.1371/journal.pone.0150050.g001

Reference

1. Cattel J, Kaur R, Gibert P, Martinez J, Frainout A, Jiggins F, et al. (2016) *Wolbachia* in European Populations of the Invasive Pest *Drosophila sukuii*: Regional Variation in Infection Frequencies. PLoS ONE 11(1): e0147766. doi: [10.1371/journal.pone.0147766](https://doi.org/10.1371/journal.pone.0147766) PMID: [26809119](https://pubmed.ncbi.nlm.nih.gov/26809119/)

Résumé

Les invasions biologiques sont une composante du changement global et ont des impacts importants sur les écosystèmes, les agrosystèmes et la santé humaine. Néanmoins, ces situations biologiques particulières offrent la possibilité d'étudier l'évolution phénotypique et génétique sur des temps écologiques. En effet, les invasions biologiques impliquent de fortes contraintes environnementales et démographiques sur les populations envahissantes et, en conséquence, de forts effets de la sélection et de la dérive. La capacité des espèces envahissantes à répondre à ces contraintes est remarquable. Qu'il s'agisse de processus génétiques d'adaptations (i.e. des changements de fréquences d'allèles) ou plastiques (i.e. un ajustement par plasticité phénotypique en réponse à un stimulus environnemental), la capacité de réponse à la sélection des espèces envahissantes les placent au centre des études de la biologie évolutive moderne. Ici, nous utilisons la récente invasion mondiale de la drosophile à ailes tachetées *Drosophila suzukii* pour étudier en détail les mécanismes impliqués dans le succès de cette invasion. Par l'analyse des patrons de variation moléculaire neutre nous avons dans un premier temps retracé l'histoire complexe de cette invasion mondiale. Ce contexte historique nous a alors permis d'évaluer la divergence phénotypique et l'évolution de la variation génétique quantitative entre populations ancestrales et dérivées de *D. suzukii*. En prenant l'aile comme caractère central d'étude, nous avons pu ainsi estimer les effets de la sélection et de la dérive génétique au cours de cette invasion, et discuter leur importance au regard de l'évolution de la forme de l'aile dans cette espèce. Enfin des protocoles expérimentaux d'analyse de la plasticité phénotypique ainsi que des méthodes de modélisation de niche climatique nous ont permis de discuter l'influence de la fluctuation des conditions environnementales sur le succès de cette invasion.

Summary

Biological invasions are a component of global change and have dramatic effects on ecosystems, agrosystems and human health. Nonetheless, these singular biological situations offer a great opportunity for the study of rapid phenotypic and genetic evolution at an ecological timescale. Biological invasions often involve environmental and demographic constraints on populations imposing strong effects of selection and drift. These species are nonetheless among the most successful colonialists in nature, and their ability to respond to these constraints is remarkable. The evolutionary and adaptive potential of invasive populations have been proposed as facilitating factors of the success of invasions. Processes of genetic (i.e. changes in allele frequencies) and plastic (i.e. adjustment to environmental fluctuation through phenotypic plasticity) adaptation involved in the success of biological invasions are at the center of modern evolutionary biology. Here, we use the recent spread of the spotted-wing *Drosophila suzukii* to study the underlying mechanisms of response to selection potentially involved in the success of this global invasion. Analyzing patterns of neutral genetic variation allowed us to decipher the complex history of this worldwide invasion, and subsequently evaluate phenotypic divergence and evolution of quantitative genetic variation among ancestral and derived populations. We thus estimated the effects of selection and drift throughout this invasion and discuss their importance regarding the evolution of wing shape in this species. Finally, experimental protocols on the analysis of phenotypic plasticity as well as Species Distribution Modeling methods allowed us to discuss the influence of environmental fluctuations on the success of this invasion.

Mots-clés: invasion biologique, génétique des populations, génétique quantitative, *Drosophila suzukii*, morphométrie géométrique, plasticité phénotypique, niche climatique