



HAL
open science

Invasion, démographie et évolution : le cas de l'hybridation

Gérald Bermond

► **To cite this version:**

Gérald Bermond. Invasion, démographie et évolution : le cas de l'hybridation. Sciences agricoles. Université Nice Sophia Antipolis, 2013. Français. NNT : 2013NICE4134 . tel-00959382

HAL Id: tel-00959382

<https://theses.hal.science/tel-00959382>

Submitted on 14 Mar 2014

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.

Université de Nice Sophia-Antipolis

UFR Sciences
École Doctorale des Sciences de la Vie et de la Santé

THÈSE

préparée au sein de *l'Institut Sophia-Agrobiotech*, UMR 1355

présentée publiquement à l'INRA de Sophia-Antipolis

pour obtenir le titre de **Docteur en Sciences** de l'Université de Nice Sophia-Antipolis

Discipline : Biologie

Mention : Biologie des interactions et écologie

Invasion, démographie et évolution : le cas de l'hybridation

par

Gérald BERMOND

Soutenue le 16 Décembre 2013

devant le jury composé de :

Tatiana Giraud	Rapportrice	Directrice de Recherche, CNRS, Université de Paris Sud
Lounès Chikhi	Rapporteur	Directeur de Recherche, CNRS, Université Paul Sabatier
Lorraine Bottin	Examinatrice	Maître de conférences, Université de Nice Sophia-Antipolis
David Renault	Examinateur	Maître de conférences, Université de Rennes
Thomas Guillemaud	Directeur de thèse	Directeur de Recherche, INRA, Sophia-Antipolis

RÉSUMÉ :

Cette thèse s'inscrit dans le contexte de la biologie de l'invasion de la chrysomèle des racines du maïs, *Diabrotica virgifera virgifera* (*Dvv*), en Europe. Les introductions multiples de cette peste sur le vieux continent à partir des années 1990 et en provenance des Etats-Unis ont conduit à la formation de plusieurs foyers envahissants dont deux principaux, différenciés génétiquement et situés en Italie du Nord-Ouest (Italie NO) et en Europe Centrale et du Sud-Est (Europe CSE). Suite à leur expansion respective, les deux foyers sont entrés en contact en 2008, en Italie du Nord, dans la région de la Vénétie. L'objectif principal de ces trois années de recherche était de détecter, d'étudier et de documenter un cas précis d'hybridation au cours d'une invasion biologique.

Dans un premier temps, j'ai cherché à montrer que la zone de contact détectée en Vénétie résultait bien en une zone hybride. Pour cela plusieurs types d'analyses de génétique des populations ont été réalisés à l'aide de marqueurs microsatellites. Les résultats obtenus révèlent la présence de plusieurs signes d'hybridation au sein d'échantillons prélevés en Vénétie et confirment l'existence d'une zone hybride en Vénétie suite au contact secondaire des deux foyers envahissants d'Italie NO et d'Europe CSE.

Dans un second temps, j'ai exploité la zone hybride précédemment détectée et utilisé la théorie des clines neutres et le déséquilibre de liaison pour estimer la dispersion (σ) de *Dvv* en Italie du Nord. En faisant l'hypothèse que la zone hybride est neutre, c'est-à-dire que les individus hybrides ne sont ni sélectionnés ni contre-sélectionnés, les analyses menées sur la diminution de la pente des clines observés aux loci microsatellites et sur la variation spatiale du déséquilibre de liaison à ces mêmes marqueurs m'ont permis d'obtenir deux estimations de σ aux alentours de 20 km/generation^{1/2}.

Enfin, j'ai cherché à confirmer l'absence d'impact de l'hybridation au cours de l'invasion européenne de *Dvv in natura*, en mesurant sur des génotypes hybrides et parentaux, des caractères phénotypiques positivement liés à la valeur sélective des individus (taux d'accouplement et survie à jeun). Les résultats indiquent que le génotype (hybride ou parental) n'a pas d'impact détectable sur les caractères mesurés et corroborent ceux d'une étude plus large, réalisée en laboratoire sur des caractères phénotypiques différents et plus nombreux. Globalement, les valeurs moyennes et les variances des caractères mesurés sur les populations hybrides ne dépassent jamais celles des populations parentales ou leur sont équivalentes. Ainsi, une sélection ou contre sélection des hybrides est à exclure dans ce cas précis d'hybridation et la dynamique de l'invasion de cette peste en Europe ne devrait être affectée en aucune façon.

D'un point de vue fondamental, les résultats obtenus au cours de cette thèse ont permis d'apporter un certain nombre d'informations sur le déroulement et l'évolution de l'invasion de l'Europe par *Dvv* et de documenter l'influence du phénomène d'hybridation intra-spécifique au cours d'une invasion biologique. Par ailleurs, les méthodes entreprises pour estimer la dispersion efficace de *Dvv* pourraient être facilement applicables à d'autres exemples d'invasions biologiques. D'un point de vue appliqué, certains résultats de cette thèse pourraient être principalement utilisés dans la gestion d'espèces nuisibles et envahissantes.

Mots clés : Introductions multiples, contact secondaire, hybridation, microsatellites, Approximate Bayesian Computation, clines, déséquilibre de liaison, dispersion, caractères d'histoire de vie

ABSTRACT :

This thesis deals with the invasion biology of the western corn rootworm, (WCR) in Europe. Multiple introductions of this pest in Europe from the United States and the 90's, led to the formation of two main invasive outbreaks, genetically differentiated and located in Northwest Italy (NW Italy) and in Central and South-Eastern Europe (CSE Europe). Both outbreaks converged towards each other and came into contact in 2008, in the region of Veneto (Northern Italy). The main objective of these three years of research was to detect, study and document a precise case of hybridization during a biological invasion.

First of all, I tried to show that the contact zone detected in Veneto results in a hybrid zone. For that, several types of population genetics analyses were performed using microsatellite markers. The results reveal the presence of several signs of hybridization in samples from Veneto and confirm the existence of a hybrid zone in Veneto due to secondary contact between the two invasive NW Italy and CSE Europe outbreaks.

In a second time, I exploited the previously detected hybrid zone and used the theory of neutral clines and the linkage disequilibrium to estimate dispersal (σ) of WCR in Northern Italy. Assuming that the hybrid zone is neutral, analyses conducted on the decrease of the slope of the observed allelic frequency clines at microsatellite loci and the spatial variation of linkage disequilibrium at these same markers allowed me to obtain two estimates of σ around 20 km/generation^{1/2}.

Finally I tried to confirm the absence of impact of hybridization in the European invasion of WCR *in natura*, by measuring on hybrid and parental genotypes, phenotypic traits positively associated to the fitness of individuals (the mating rate and fasting survival). The results indicate that the genotype (hybrid or parental) has no impact on the measured traits and corroborate those of a broader study conducted in the laboratory on different and more numerous phenotypic traits. Overall, the mean values and variances of the traits measured on hybrid populations never exceed those of the parental populations or are equivalent. Thus, a selection or a selection against hybrids is excluded in this particular case of hybridization and the invasion dynamics of WCR in Europe should not be affected in any way.

From a fundamental point of view, the results obtained in this thesis provided some information on the development and progress of the invasion of Europe by WCR and document the influence of the phenomenon of intraspecific hybridization in a biological invasion. Moreover, the methods undertaken to estimate the effective dispersal of WCR could easily be applied to other examples of biological invasion. On the applied point of view, some of the results of this thesis could be principally used in the management of pests and invasive species.

Key words : Multiples introductions, secondary contact, admixture, microsatellites, Approximate Bayesian Computation, clines, linkage disequilibrium, dispersal, life history traits

REMERCIEMENTS

Je remercie tout d'abord Thomas Guillemaud de m'avoir offert l'opportunité de travailler avec lui dès le Master 2 et durant ces 3 années de thèse. Les cadres scientifiques et matériels proposés m'ont permis d'apprendre énormément et de m'épanouir scientifiquement. Je tiens aussi à le remercier pour sa disponibilité et sa réactivité dans les échanges professionnels.

Je remercie tous les membres de mon comité de thèse : Nicolas Ris, Elodie Vercken, Benoit Facon, Virginie Ravigné, Arnaud Estoup, Yannis Michalakis, Stefan Toepfer et Etienne Klein pour leurs conseils avisés.

Merci aussi aux personnes avec lesquelles j'ai interagi ponctuellement mais sans lesquelles le gros du travail de thèse n'aurait pas été possible. Je pense particulièrement à Thomas Lenormand et Adrien Rieux.

Merci à Fanny Cavigliasso, pour la qualité du travail qu'elle a effectué au cours de son stage de Master 1, pour son sourire et sa bonne humeur me permettant ainsi de connaître une première expérience d'encadrement très agréable.

Merci aux membres de l'équipe « BPI » : Alex', Aurélie, Thibaut M., Eric, Isabelle L., Isabelle Y., Xavier, Gégé, Didier, Elo et Pascal, pour leur accueil et leur convivialité.

Un merci général aux membres du bureau « thésards » de BPI : Chloé, Thib', Soso, Margarita et Bastoune pour la bonne humeur partagée au cours de cette thèse au sein du bureau mais aussi pour leur soutien moral et professionnel. Un merci tout particulier à Sophie pour m'avoir accompagné sur le terrain l'année dernière et ne pas avoir hésité une seule seconde à venir s'empêtrer avec moi dans la boue des champs de maïs d'Italie et d'Europe Centrale au nom de la science !! La mission fut un succès....Merci aussi à Thib', alias « Mr Stats » pour ses coups de pouce ! Ce gars sait tout faire avec R....la classe !! ;)

En dehors du labo, je tiens à remercier les membres du groupe ROXO qui m'ont accompagné tout au long de cette thèse et m'ont offert de belles pauses musicales...Merci aussi aux amis/voisins Xavier, Bastien et Ali pour les agréables moments passés ensemble. Merci à mes deux compères ULTRAS, basse et ténor à leurs heures perdues....ils se reconnaîtront !

Merci à ma famille pour son soutien, son écoute et son intérêt : papa, maman, mamie, Julia, Romain, mes beaux-parents, Maureen et LAURA, mon âme sœur, qui fut de tous les combats durant cette thèse.

Cette soutenance marque pour moi la fin d'une période de travail de 5 ans à l'INRA de Sophia-Antipolis et je tiens donc à remercier toutes les personnes rencontrées, avec lesquelles j'ai évolué et qui se sentent impliquées : MERCI.

Je souhaite enfin remercier tous les membres du jury, Tatiana Giraud, Lounès Chikhi, Lorraine Bottin et David Renault d'avoir accepté mon invitation malgré des emplois du temps surchargés.

AVANT PROPOS

J'ai pris le parti de rédiger ce manuscrit suivant le format d'une « Thèse sur articles ». J'aime à croire que les résultats obtenus au cours de ces trois années de recherche permettent de raconter une histoire cohérente sur l'évolution de certains foyers envahissants européens de la chrysome des racines du maïs, *Diabrotica virgifera virgifera*.

Chacun des objectifs particuliers de ma thèse a fait l'objet d'une étude et a vocation à être valorisé sous la forme d'un article. Après avoir présenté dans une large introduction, le contexte scientifique global et les différents concepts théoriques abordés au cours de cette thèse, j'ai donc décidé de consacrer un chapitre par objectif particulier visé. Un résumé en français précèdera chacun des trois premiers articles publiés ou en préparation et permettra au lecteur de cerner directement la nature des questions posées, des objectifs visés, des méthodes entreprises pour y parvenir et des principaux résultats obtenus.

Le dernier chapitre concerne une étude non prévue initialement dans le cadre de ma thèse et déjà réalisée par des collaborateurs du laboratoire CABI en Suisse avant mon arrivée au sein de l'équipe. Mon investissement récent dans le traitement des analyses génétiques et statistiques et la proximité des questions posées dans cette étude avec celles évoquées dans le **Chapitre III** (Absence d'impact de l'hybridation *in natura* au cours de l'invasion Européenne de *Diabrotica virgifera virgifera*) m'ont permis de faire pleinement partie du projet. Néanmoins, comme cette étude n'était pas encore présentable sous la forme d'un article scientifique (la rédaction n'est pas assez avancée), j'ai décidé de l'intégrer dans un chapitre à part.

Une discussion générale permettra enfin de revenir sur l'ensemble des résultats obtenus et de les discuter de manière critique.

SOMMAIRE

I. – CONTEXTE GÉNÉRAL.....	9
I.1 – Les invasions biologiques	9
I.1.1 – Définitions et généralités	9
I.1.2 – Chronologie d'un processus complexe.....	14
I.1.3 – Une réussite non-systématique	16
I.1.4 – Pourquoi étudie-t-on les invasions biologiques ?	18
I.1.5 – Importance des routes d'invasion.....	19
I.2 – L'hybridation.....	22
I.2.1 – Concept et généralités	22
I.2.2 – Conséquences évolutives de l'hybridation.....	26
I.2.3 – Etudes de la balance migration/sélection.....	29
I.3 – L'hybridation intra-spécifique au cours des invasions biologiques	30
I.3.1 – Définition et origine.....	30
I.3.2 – Stimulus des invasions biologiques	31
I.3.3 – Remise en cause.....	32
I.3.4 – Estimation de la dispersion dans le cas d'une diffusion pure.....	33
II. – LE MODÈLE BIOLOGIQUE : LA CHRYSOMÈLE DES RACINES DU MAÏS, <i>DIABROTICA VIRGIFERA VIRGIFERA</i>	35
II.1 – Généralités sur la biologie de la chrysomèle des racines du maïs.....	35
II.2 – Aire native et invasion.....	38
II.3 – Méthodes de lutte et résistance.....	40
III. – OBJECTIFS DE LA THÈSE	44
IV. – CHAPITRE I : LA VÉNÉTIE, UNE ZONE HYBRIDE ENTRE DEUX FOYERS ENVAHISSANTS EUROPÉENS DE <i>DIABROTICA VIRGIFERA VIRGIFERA</i>	46
V. – CHAPITRE II : ÉTUDE DE LA ZONE HYBRIDE DE VÉNÉTIE ET ESTIMATION DE LA DISPERSION DE <i>DIABROTICA VIRGIFERA VIRGIFERA</i>	68
VI. – CHAPITRE III : ABSENCE D'IMPACT DE L'HYBRIDATION <i>IN NATURA</i> AU COURS DE L'INVASION EUROPÉENNE DE <i>DIABROTICA VIRGIFERA VIRGIFERA</i>	110
VII. – CHAPITRE IV : IMPACT DE L'HYBRIDATION ENTRE POPULATIONS ENVAHISSANTES DE <i>DIABROTICA VIRGIFERA VIRGIFERA</i> EN EUROPE SUR DES CARACTÈRES GÉNÉTIQUES ET PHÉNOTYPIQUES MESURÉS AU LABORATOIRE	152
VIII. – DISCUSSION GÉNÉRALE	180
VIII.1 – L'Italie du Nord, une zone de contact riche en informations	180
VIII.2 – Contact secondaire, hybridation et estimation de la dispersion.....	183
VIII.3 – Aucun effet majeur de l'hybridation au cours de l'invasion européenne de <i>Diabrotica virgifera virgifera</i>	188
VIII.4 – Synthèse	194
VIII.5 – La zone hybride temporaire comme opportunité pour estimer la dispersion	197
RÉFÉRENCES BIBLIOGRAPHIQUES	200

Liste des Figures

Figure 1 : Nombre cumulé d'espèces introduites et établies A) dans la baie de San Fransisco (tiré de Cohen, Carlton, 1998) et B) en République Tchèque au cours du temps.....	12
Figure 2 : Représentation schématique du processus séquentiel d'invasion biologique (adaptée à partir Duncan <i>et al.</i> , 2003).....	15
Figure 3 : Les zones hybrides.....	24
Figure 4 : Conséquences évolutives de l'hybridation.....	26
Figure 5 : La chrysomèle des racines du maïs, <i>Diabrotica virgifera virgifera</i> et ses dégâts associés dans les champs de maïs.....	36
Figure 6 : Cycle de vie de <i>Diabrotica virgifera virgifera</i>	37
Figure 7 : Aire native et invasion de <i>Diabrotica virgifera virgifera</i> sur le continent américain.....	38
Figure 8 : Routes d'invasion de <i>Diabrotica virgifera virgifera</i> en Europe.....	39
Figure 9 : Routes d'invasion et évolution de l'aire de répartition de <i>Diabrotica virgifera virgifera</i> en Europe, entre 2004 et 2008.....	40
Figure 10 : Aire de répartition de <i>Diabrotica virgifera virgifera</i> avant (2005) et après (2010) le contact secondaire entre les foyers d'Italie du Nord-Ouest et d'Europe Centrale et du Sud-Est.....	47
Figure 11 : Localisation des échantillons de <i>Diabrotica virgifera virgifera</i> en Europe : en Italie du Nord-Ouest, dans la zone de contact (Vénétie), en Italie du Nord-Est et en Europe Centrale et du Sud-Est.....	49
Figure 12 : Localisation des échantillons de <i>Diabrotica virgifera virgifera</i> collectés en 2012 en Italie du Nord-Ouest, dans la zone hybride, en Italie du Nord-Est et en Europe Centrale et du Sud-Est.....	70
Figure 13 : Localisation des deux principaux foyers européens envahissants de <i>Diabrotica virgifera virgifera</i> (<i>Dvv</i>) en Italie du Nord-Ouest et en Europe Centrale et du Sud-Est et localisation des échantillons de <i>Dvv</i> collectés en 2012 dans la zone hybride.....	113
Figure 14 : Localisation des deux principaux foyers européens envahissants de <i>Diabrotica virgifera virgifera</i> (<i>Dvv</i>) en Italie du Nord-Ouest et en Europe Centrale et du Sud-Est et localisation des échantillons de <i>Dvv</i> collectés en 2006 en Italie NO et Europe CSE.....	155
Figure 15 : Design expérimental de l'étude menée au sein du Chapitre IV.....	157

Liste des Tableaux (relatifs au CHAPITRE IV)

Tableau 1 : Nombre de familles reconstruites par croisement et par bloc.....	164
Tableau 2 : Statistiques résumant la variation génétique au sein de chaque croisement de <i>Diabrotica virgifera virgifera</i>	165
Tableau 3 : Tests de comparaison de la variabilité génétique entre échantillons quand l'effet famille est pris ou non en considération.....	166
Tableau 4 : Statistiques résumant la variation génétique au sein du jeu de données ré-échantillonnées de <i>Diabrotica virgifera virgifera</i>	168
Tableau 5 : Résultats obtenus à partir du modèle linéaire (généralisé) mixte générique pour chacun des caractères phénotypiques mesurés.....	170

I. – CONTEXTE GÉNÉRAL

I.1 – Les invasions biologiques

I.1.1 – Définitions et généralités

Le déplacement des espèces à la surface de la terre est un processus naturel qui dépend des exigences écologiques et de l'histoire évolutive de chaque taxon. Au cours du temps, la pression exercée par l'évolution — au travers des forces évolutives telles que la dérive génétique et la sélection — sur les espèces qui se déplacent dans l'espace a permis leur diversification et généré la richesse spécifique que l'on connaît aujourd'hui à l'échelle du globe. Notre perception des différents milieux terrestres et des frontières spatio-temporelles nous a permis de différencier les espèces natives¹ des non-natives² d'une région ou d'une localité donnée (Pysek, Richardson, 2006). Depuis environ deux siècles, le déplacement des espèces hors de leur aire native, leur introduction et évolution dans de nouveaux milieux n'ont cessé d'augmenter, principalement du fait de l'essor de certaines activités anthropiques. Ces phénomènes sont décrits sous le terme d'« invasions biologiques » ; mais que décrit-il exactement ?

a) Un terme à plusieurs sens...

Le sens général du terme « invasions biologiques » décrit un processus de colonisation d'une zone géographique par une espèce dite envahissante et initialement absente de la zone concernée. Dans la littérature, la définition d'une invasion biologique varie en fonction du

¹ Taxon indigène, endémique ou lié à une invasion préhistorique, d'après Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology* 77, 1653-1655.

² Invasions historiques incluant à la fois les expansions naturelles d'aire de répartition et les introductions liées à l'homme, d'après *Ibid.*

domaine considéré, de l'importance que l'on attribue à son impact et de la dimension géographique de l'aire envahie.

Pour certains auteurs et organisations telles que l'Union Internationale pour la Conservation de la Nature et des ressources naturelles (UICN), les invasions biologiques concernent uniquement les espèces qui provoquent des dégâts au sein des écosystèmes ou qui impactent l'économie et la santé humaine (IUCN SSC Invasive Species Specialist Group (ISSG), 2009). Dans ce cas, la définition du terme « invasion biologique » est intimement liée aux secteurs d'activités anthropiques (l'économie, l'agriculture et la santé humaine) et centrée principalement sur la nuisibilité que peut générer l'espèce introduite.

Certains autres auteurs pensent au contraire que les notions d'« invasions biologiques » et d'« espèces envahissantes » doivent être indépendantes des nuisances causées et considérées uniquement en raison de leur expansion géographique et démographique (Colautti, MacIsaac, 2004; Richardson *et al.*, 2000). Deux définitions correspondent à cette vision neutre des invasions biologiques. La première, initialement proposée par Elton (1958) et plus récemment reprise par Mack (2000), stipule qu'une invasion biologique se définit comme l'introduction, l'établissement et la prolifération d'une espèce dans un nouvel environnement. La seconde proposée par Williamson (1996) a un sens plus large puisqu'une invasion biologique correspond simplement au déplacement d'un organisme au-delà de son aire de répartition précédente, sans notion de déconnexion géographique par rapport à son aire native. Cette dernière définition n'est pas complètement satisfaisante car nous considérons que les processus liés à la prolifération et l'expansion d'un taxon dans son aire d'origine sont différents de ceux permettant la colonisation d'un nouveau milieu et parce que la délimitation précise des bordures d'une aire de répartition — par nature instables (*e.g.* Ricklefs, Bermingham, 2002) — reste difficile.

Dans la suite de ce manuscrit le terme d' « invasions biologiques » se rapportera donc à la définition plus restrictive d'Elton (1958) et de Mack (2000) qui nous paraît objective et neutre vis-à-vis des impacts potentiels et qui introduit déjà l'idée qu'une invasion biologique est un processus constitué de plusieurs étapes successives (voir **section I.1.2**).

b) Dynamique des invasions biologiques

Nous l'avons déjà évoqué, le nombre d'invasions biologiques a littéralement explosé au cours de ces deux derniers siècles principalement sous l'influence de l'homme. Depuis la révolution industrielle, les activités anthropiques (l'agriculture, l'aquaculture, le tourisme, le commerce...) et les moyens de transport n'ont cessé de se développer, se multiplier et se moderniser, provoquant par la même occasion un déplacement anormal des espèces hors de leur aire native (Lockwood *et al.*, 2007; Tatem, Hay, 2007). Cohen and Carlton (1998) ont par exemple montré que le nombre de taxons envahissants n'avait cessé d'augmenter dans la baie de San Francisco depuis 1850 et que la moitié de ces taxons était apparue à partir de 1960 (Figure 1A). En Europe, Pysek *et al.* (2003) ont également montré que le nombre cumulé d'espèces non-natives augmentait constamment depuis la fin du 18^e siècle en République Tchèque et s'était particulièrement accéléré à la fin du 20^e siècle (Figure 1B).

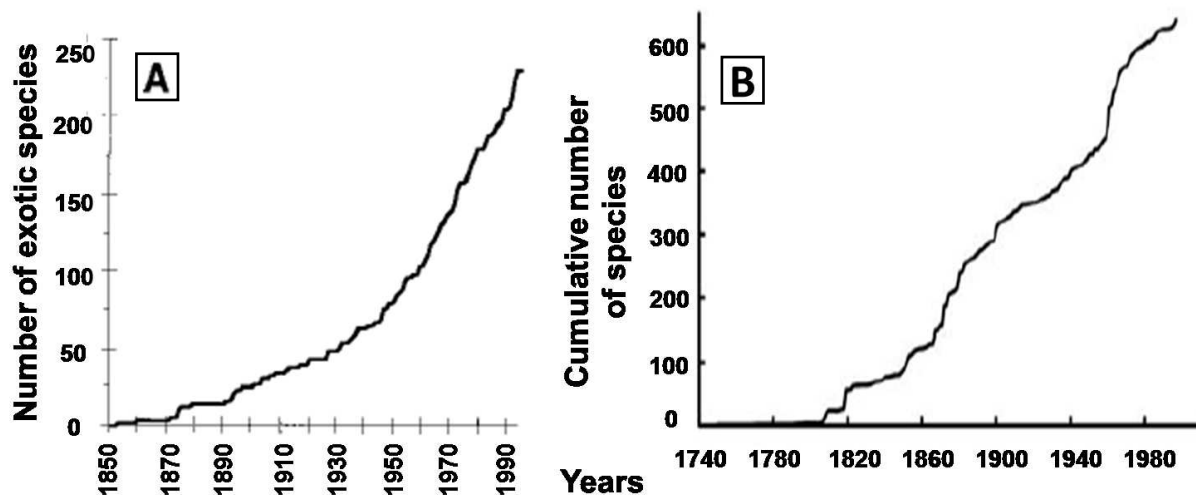


Figure 1 : Nombre cumulé d'espèces introduites et établies A) dans la baie de San Fransisco (tiré de Cohen, Carlton, 1998) et B) en République Tchèque au cours du temps (tiré de Pysek *et al.*, 2003).

c) Impact des invasions biologiques

La prolifération du nombre d'invasions biologiques à l'échelle mondiale n'est pas sans conséquence.

Tout d'abord, les invasions biologiques peuvent s'avérer dramatiques pour la biodiversité et l'équilibre des écosystèmes (Courchamp *et al.*, 2003; McKinney, Lockwood, 1999; Olden *et al.*, 2004). Elles constituent actuellement la seconde cause d'extinction d'espèces après la destruction et la fragmentation des habitats (Vitousek *et al.*, 1997). Certains auteurs les tiennent même pour responsables du phénomène d'homogénéisation biotique à l'échelle mondiale, défini comme le remplacement graduel de nombreuses espèces indigènes par un nombre restreint d'espèces envahissantes (McKinney, Lockwood, 1999; Olden *et al.*, 2004).

Les invasions biologiques s'avèrent également catastrophiques pour un certain nombre d'activités économiques majeures telles que l'agriculture, la pêche, la sylviculture, le tourisme et les loisirs (Born *et al.*, 2005). Le coût des dommages engendrés par plus de 120 000 espèces envahissantes, répertoriées dans six pays (États-Unis, Royaume-Uni, Australie,

Afrique du Sud, Inde et Brésil), a été estimé à plus de 314 milliards US\$ par an (Pimentel *et al.*, 2001). L'extrapolation de ces résultats à l'ensemble de la planète aboutirait à 480 000 espèces envahissantes, dont 20 à 30 % engendreraient 1 400 milliards US\$ de coûts par an (Pimentel *et al.*, 2001).

Enfin, la santé humaine peut aussi être fortement impactée par les invasions biologiques. En effet, certaines espèces envahissantes (rongeurs, insectes, oiseaux) sont vectrices de nombreuses maladies humaines. L'introduction du rat noir, *Rattus rattus*, en Europe a par exemple entraîné de larges épidémies de peste bubonique au 17^e siècle (Monecke *et al.*, 2009). Plus récemment, il a été montré que des microorganismes marins potentiellement pathogènes pouvaient être transportés sur de très longues distances *via* les ballasts des bateaux (Ruiz *et al.*, 2000).

Nous venons de dresser un tableau assez négatif de l'impact des invasions biologiques mais nous tenons tout de même à faire remarquer que ces dernières peuvent aussi avoir un effet positif. En effet, bon nombre de plantes cultivées et d'animaux d'élevage sont issus de processus envahissants et permettent aujourd'hui de fournir des ressources alimentaires à l'ensemble de la planète et de générer des bénéfices économiques (Pimentel *et al.*, 2001; Pimentel *et al.*, 1997). Certaines populations envahissantes peuvent également profiter à la pêche commerciale. L'introduction de la perche du Nil au sein du lac Victoria en est un exemple évident (Ogutu-Ohwayo, 2001). Enfin, certaines invasions biologiques contrôlées par l'homme peuvent permettre de réduire les coûts engendrés par des ravageurs agricoles envahissants ; c'est notamment le cas de la lutte biologique classique³ (Eilenberg *et al.*, 2001; Pimentel *et al.*, 1997).

³ Introduction intentionnelle d'un ou plusieurs agents de lutte biologique exotiques (prédateurs ou parasites) pour un établissement permanent et le contrôle sur le long terme d'un ravageur.

d) Espèces ou populations envahissantes ?

Il est important de noter qu'au cours des invasions biologiques seules des populations s'introduisent, s'établissent et se propagent dans le nouvel environnement et non l'espèce entière (Colautti, MacIsaac, 2004). Même si nous avons tendance dans le langage usuel à employer le terme d'espèces envahissantes, il serait donc plus correct de parler de populations envahissantes. J'essaierai autant que faire se peut d'utiliser le terme de populations envahissantes dans la suite de ce manuscrit.

I.1.2 – Chronologie d'un processus complexe

Une invasion biologique est un processus qui se décompose en plusieurs étapes successives (Duncan *et al.*, 2003; Kolar, Lodge, 2001; Leung *et al.*, 2002; Sakai *et al.*, 2001; Williamson, Fitter, 1996). Suivant les auteurs, le nombre d'étapes varie mais plusieurs d'entre eux s'accordent pour en considérer quatre principales (Figure 2, Duncan *et al.*, 2003; Kolar, Lodge, 2001). Il y a tout d'abord l'étape d'émigration des organismes hors de leur aire native et vers une nouvelle localité. Le transport des organismes durant cette étape peut s'effectuer selon leurs propres capacités ou par l'intermédiaire d'un vecteur. Ensuite vient l'étape d'introduction (relâchement) d'individus fondateurs dans le nouvel habitat, éloigné et déconnecté de l'aire native. Cette étape souvent associée à un « goulot d'étranglement » est prépondérante pour la suite du processus (voir **section I.1.3.b**). L'étape suivante correspond à l'établissement de la population introduite au sein du nouvel environnement. Cette dernière se traduit par la survie des individus fondateurs, leur reproduction et la survie de leurs descendants sur plusieurs générations après l'étape d'introduction. Enfin, le processus s'achève par l'expansion géographique et démographique de la population introduite et établie au sein de l'aire envahie. La réussite de cette dernière étape détermine ou non le caractère envahissant des populations établies.

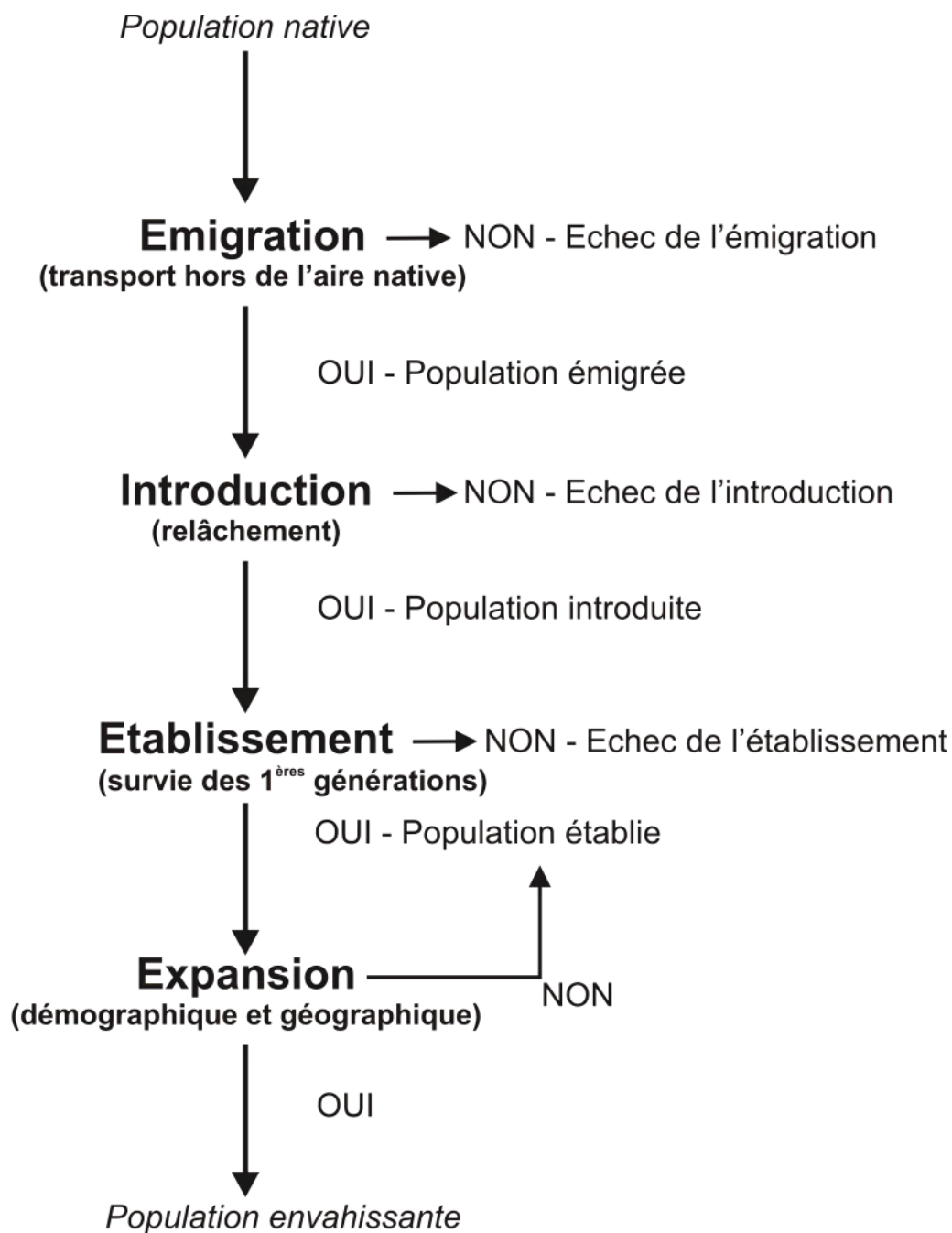


Figure 2 : Représentation schématique du processus séquentiel d'invasion biologique (adaptée à partir Duncan *et al.*, 2003). Les 4 étapes principales par lesquelles doivent passer les populations (en italique) avant de devenir envahissantes sont représentées en gras.

I.1.3 – Une réussite non-systématique

a) Théorie sur le taux de réussite des invasions biologiques

Malgré une augmentation exponentielle du nombre d'invasions biologiques au cours des deux derniers siècles (Mack *et al.*, 2000) provoquée, entre autres, par l'expansion des migrations humaines et du commerce, seule une minorité des populations qui émigrent hors de leurs aires natives parvient à devenir envahissante. Nous pouvons citer en exemple les travaux de Suarez *et al.* (2005) qui montrent que parmi les 232 espèces de fourmis introduites sur le territoire américain et interceptées dans des marchandises par l'USDA entre 1927 et 1985, seulement 12% se sont établies et 1% sont devenues envahissantes. Certains auteurs ont proposé des règles théoriques pour tenter de quantifier cette proportion. La fameuse règle des « dixièmes », établie par Williamson et Fitter (1996) considère un taux de réussite à chaque étape du processus d'invasion biologique de seulement 10%. Ainsi, seulement 10% des populations qui émigrent arriveraient à s'introduire dans le nouveau milieu, puis, parmi elles, seulement 10% parviendraient à s'établir et enfin 10% des populations établies deviendraient envahissantes. En suivant cette règle, une population émigrante sur mille parviendrait à devenir envahissante. La règle des dixièmes est souvent citée dans la littérature spécifique aux invasions biologiques mais ne devrait pas, selon ses auteurs, être considérée « trop sérieusement » (Williamson, 2006). Il ne faut pas envisager cette règle des « dixièmes » comme une règle immuable mais plutôt comme la traduction d'un fait réel : le succès des invasions biologiques est moins fréquent que leur échec. Par ailleurs, dans le cadre d'invasions biologiques contrôlées comme par exemple la lutte biologique, cette règle des « dixièmes » peut évoluer en règle des « tiers ». Williamson et Fitter (1996) ont estimé qu'environ 30% des espèces introduites en lutte biologique classique s'établissent et 30% des espèces établies agiront avec succès sur le ravageur. Pour résumer ce paragraphe, on peut dire que, quelle que soit l'invasion biologique considérée (fortuite ou contrôlée), une minorité des

espèces qui émigrent hors de leur aire native réussissent une à une les étapes successives (émigration, introduction, établissement et expansion) du processus. On peut s'interroger sur les facteurs qui influencent le succès ou l'échec de ces différentes étapes et *in fine* des invasions biologiques ?

b) Des facteurs limitant le succès de l'invasion biologique

D'après Fauvergue *et al.* (2012), les facteurs impactant le succès ou l'échec des invasions peuvent se classer en trois catégories distinctes.

La première catégorie est d'ordre environnemental. Les populations introduites dans de nouveaux habitats doivent souvent faire face à des conditions abiotiques et/ou biotiques qui défavorisent leur établissement en raison de leurs trop grandes différences avec celles de l'aire native. Il semble cependant que ces deux conditions n'agissent pas de la même manière vis-à-vis des populations envahissantes. Selon les auteurs, les conditions abiotiques peuvent agir comme de véritables remparts face à l'envahisseur et permettront ainsi de repousser l'invasion biologique. Par exemple chez les poissons, il a été montré que le facteur le plus important pour déterminer le succès de populations envahissantes au sein des cours d'eau et estuaires californiens était l'adéquation entre les individus introduits et les conditions hydrologiques du milieu (Moyle, Light, 1996). Les conditions biotiques (compétition, prédation), elles, ne permettent pas de repousser les populations envahissantes mais plutôt de contraindre leur expansion et leur impact au sein des communautés et des écosystèmes envahis. Cela a notamment été démontré chez les plantes (Levine *et al.*, 2004).

Les causes du succès ou de l'échec des invasions biologiques peuvent aussi être démographiques. Lors de l'étape d'introduction, seul un petit nombre d'individus de la population native est introduit dans le nouvel environnement et l'expression « goulot d'étranglement » prend alors tout son sens. Cette réduction d'effectif au sein de l'aire envahie

peut générer d'autres processus tels que la stochasticité démographique⁴ (Lande, 1993) et les effets Allee⁵ (1949) qui peuvent conduire à l'extinction des petites populations introduites.

Enfin, la dernière catégorie implique les causes génétiques. Les réductions d'effectif associées aux goulots d'étranglement lors de l'étape d'introduction se traduisent généralement par une diminution de la diversité génétique et entraînent une réduction du potentiel adaptatif des populations (Willi *et al.*, 2006). Les conséquences de cette réduction peuvent être importantes pour les populations introduites puisque le nouvel environnement représente de nouvelles pressions de sélection qui s'appliqueront alors sur un spectre limité de variabilité génétique. Les goulots d'étranglement peuvent aussi provoquer une hausse de la consanguinité dans les populations introduites (au bout de quelques générations du fait de la petite taille de la population initiale) et générer de la dépression de consanguinité⁶ voire la fixation d'allèles délétères. De tels processus augmenteront alors la probabilité d'extinction des populations introduites (Fauvergue *et al.*, 2012).

I.1.4 – Pourquoi étudie-t-on les invasions biologiques ?

Comme nous venons de le voir, les invasions biologiques se multiplient et représentent de réelles menaces pour l'homme (risques sanitaires), son environnement (érosion de la biodiversité) et la société qu'il a façonnée (coûts engendrés au secteur économique). Dans un premier temps la communauté scientifique a donc cherché à mieux les comprendre pour essayer de les éviter, les combattre ou les gérer au mieux.

⁴ Se réfère à la déviance entre le nombre d'individus observé et attendu à partir des valeurs moyennes de paramètres démographiques tels que la survie, la fécondité et ou le sex-ratio, d'après Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced populations, with special reference to biological control. *Evolutionary applications* **5**, 424-443.

⁵ Diminution de la valeur sélective (définie ici comme le nombre moyen des descendants laissés à la génération suivante par un individu) d'un individu ou du taux de croissance d'une population suite à une réduction d'effectif ou de densité.

⁶ Diminution de la valeur sélective des descendants de croisements entre apparentés. La dépression de consanguinité se produit chez des individus consanguins uniquement lorsque des mutations délétères récessives se retrouvent à l'état homozygote, d'après Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced populations, with special reference to biological control. *Evolutionary applications* **5**, 424-443.

Une seconde raison qui pousse les scientifiques à étudier les invasions biologiques est le fait qu'elles constituent de réelles expérimentations en grandeur nature et permettent de mieux comprendre certains concepts en écologie et biologie évolutive tels que les processus d'adaptation, de spéciation, de colonisation, d'expansion spatiale mais aussi des phénomènes évolutifs plus rapides comme des évènements d'hybridation intra-spécifique (voir **section I.3**) par exemple (Phillips *et al.*, 2010; Reznick, Ghalambor, 2001; Sax *et al.*, 2007; Verhoeven *et al.*, 2011). Un autre avantage des invasions biologiques réside dans le fait qu'elles sont (pour une même espèce considérée) couramment répétées dans des localités géographiques différentes (*e.g.* Ciosi *et al.*, 2008; Lavergne, Molofsky, 2007; Lombaert *et al.*, 2010) et permettent alors de comparer les trajectoires évolutives suivies par les différentes populations introduites.

I.1.5 – Importance des routes d'invasion

Les routes d'invasion, se définissent par le parcours géographique des propagules entre l'aire native et l'aire envahie (Estoup, Guillemaud, 2010). Leur connaissance apporte de nombreuses informations historiques sur le processus envahissant et permet également de fournir des éléments sur l'origine et la mise en place de la composition génétique des populations envahissantes. Différentes hypothèses concernant les facteurs environnementaux et évolutifs sous-jacents aux invasions biologiques peuvent être testées à partir de ces connaissances (Estoup, Guillemaud, 2010). En comparant les populations sources et introduites, on peut par exemple tenter de déterminer si les divergences phénotypiques observées entre l'aire native et l'aire envahie résultent d'évènements purement aléatoires et démographiques ou bien d'évènements sélectifs (Keller, Taylor, 2008). La reconstruction des routes d'invasion apparaît donc comme primordiale pour identifier la nature et mesurer

l'étendue de ces changements évolutifs concernant la composition génétique des populations introduites ou les changements évolutifs qui ont pu se produire au cours de l'invasion.

Étudier les routes d'introduction de *Diabrotica virgifera virgifera* en Europe n'était pas le but de ma thèse et par ailleurs, cela a déjà été effectué (Ciosi *et al.*, 2008; Miller *et al.*, 2005) ; cependant j'ai utilisé dans mes analyses certaines des méthodes qui permettent de les retracer. Je présenterai donc très brièvement ici les deux types de méthodes qui existent pour reconstruire les routes d'invasions (Estoup, Guillemaud, 2010):

a) Les méthodes directes

Ces méthodes reposent essentiellement sur la récupération et l'analyse de données géographiques et historiques telles que les dates et les sites de première observation des populations introduites et leur proximité avec des lieux relatifs au transport (ports, gares ferroviaires, aéroports...). Même si ce genre de méthodes a déjà permis de retracer les routes d'invasions de certaines populations envahissantes (Suarez *et al.*, 2001; Tatem *et al.*, 2006), elles restent très limitées. La première raison à cela est que, trop souvent, les informations obtenues pour une invasion donnée sont incomplètes ou absentes. Par ailleurs, la méthode en elle-même ne permet pas de prendre en considération des scénarios complexes d'invasion comme par exemple les introductions multiples qui semblent être plus communes qu'il n'y paraît (Bossdorf *et al.*, 2005; Roman, Darling, 2007).

b) Les méthodes indirectes

Ces méthodes font appel à la génétique des populations et sont basées sur l'utilisation de patrons génétiques intra et inter-populationnels, observés à l'aide de marqueurs moléculaires. Les différents patrons génétiques permettront d'inférer des relations de parenté entre populations et *in fine* de reconstruire les routes d'invasion. Certaines méthodes se servent de statistiques résumées de la diversité génétique telles que les F_{ST} (Weir, Cockerham, 1984) ou des vraisemblances d'assignation (Rannala, Mountain, 1997) pour retracer les routes

d'invasion (Ciosi *et al.*, 2008; Genton *et al.*, 2005). D'autres telle que celle implémentée dans le logiciel STRUCTURE (Pritchard *et al.*, 2000) et développées plus récemment, regroupent les individus ou les populations sur la base de modèles et de calculs de vraisemblance et permettent aussi d'inférer les populations sources des populations introduites (e.g. Darling *et al.*, 2008; Marrs *et al.*, 2008; Rollins *et al.*, 2009). De manière générale, toutes ces méthodes indirectes présentent trois limites principales : (i) elles ne permettent pas de prendre en considération la stochasticité démographique et génétique (la dérive) qui peuvent s'avérer particulièrement importantes lors d'une invasion biologique ; (ii) elles se basent sur un échantillonnage considéré comme exhaustif or ça peut ne pas être le cas ; (iii) elles ne permettent pas de calculer statistiquement le degré confiance relatif attribué à chacun des résultats obtenus.

Une dernière méthode peut enfin être présentée. Il s'agit de l'Approximate Bayesian Computation (ABC Beaumont *et al.*, 2002). L'ABC est une approche bayésienne dans laquelle les probabilités *a posteriori* de modèles évolutifs (ou scénarios historiques) et/ou les distributions de paramètres démographiques selon un modèle donné sont déterminées en mesurant la similarité entre un jeu de données observées (les données qui viennent du terrain par exemple) et un large jeu de données simulées. Les deux jeux de données sont résumés par des statistiques de variabilité génétique (nombre moyen d'allèles par locus, hétérozygotie par population...) (Estoup, Guillemaud, 2010). [Pour plus de détails sur cette méthode, je conseille au lecteur de se rapporter aux articles suivants : (Bertorelle *et al.*, 2010; Csillery *et al.*, 2010)] Cette méthode a de nombreux domaines d'application et s'est récemment montrée très utile pour traiter de larges jeux de données en biologie évolutive, en biologie des populations et en épidémiologie (Estoup, Guillemaud, 2010). Par ailleurs, l'ABC présente plusieurs avantages qui la rendent très puissante et précieuse dans le domaine de la biologie de l'invasion. Elle permet tout d'abord d'associer des données historiques (temps de

divergence entre populations...) et biologiques (nombre d'individus fondateurs...) au sein des analyses. Deuxièmement, elle permet de prendre en considération l'existence de goulots d'étranglement, de populations non échantillonnées et même d'évènements d'hybridations au sein des modèles évolutifs que l'on peut élaborer et comparer (*e.g.* Bermond *et al.*, 2012; Miller *et al.*, 2005). Enfin, elle permet de quantifier la probabilité relative d'un scénario d'introduction particulier et d'évaluer les erreurs de type I (de première espèce) et II (de deuxième espèce) des résultats obtenus. Les principales critiques que l'on peut émettre à l'encontre de cette méthode concernent le choix du nombre et la nature des statistiques résumées à utiliser, son cadre bayésien impliquant une définition *a priori* de la distribution des paramètres impliqués et enfin un choix restreint de modèles (Templeton, 2009; Templeton, 2010).

Du fait des limites de chacune des méthodes exposées ci-dessus, une manière efficace de reconstruire les routes d'invasion serait de combiner les méthodes directes et indirectes afin de corroborer les résultats obtenus. Une fois les routes d'invasion retracées, on peut alors se consacrer à l'étude des changements évolutifs qui peuvent intervenir au cours de l'invasion comme par exemple des évènements d'hybridation...

I.2 – L'hybridation

I.2.1 – Concept et généralités

a) Définitions

Parmi les nombreuses définitions du mot « hybridation » que l'on peut trouver dans la littérature (Harrison, 1993), nous avons choisi celle d'Harrison (1990) qui stipule que l'hybridation correspond à un croisement entre individus appartenant à deux populations ou groupes de populations, distinguables sur la base d'un ou plusieurs caractères héréditaires. Par

cette définition Harrison subdivise le processus d'hybridation en deux. En effet, il inclut aussi bien les croisements entre individus appartenant à des espèces différentes (hybridation interspécifique) que les contacts entre populations conspécifiques (hybridation intra-spécifique, voir **section I.3**).

Les conséquences évolutives de ces deux types d'hybridation étant communes, nous développerons cette partie en considérant l'hybridation au sens large puis nous nous intéresserons plus particulièrement dans la **section I.3** à l'hybridation intra-spécifique, compte tenu du contexte de cette thèse (voir **section I.4**).

b) Les zones hybrides

L'hybridation se matérialise au niveau géographique par une zone hybride, c'est-à-dire une zone dans laquelle des individus divergents se rencontrent, se reproduisent et donnent naissance à des individus hybrides (Barton, Hewitt, 1989; Harrison, 1990). D'un point de vue théorique, ces zones hybrides se définissent souvent par un cline ou un ensemble de clines, c'est-à-dire une ou plusieurs variations graduelles de valeurs mesurées sur différents caractères (phénotypiques, génétiques, écologiques, comportementaux...) le long d'un axe géographique et entre deux populations parapatriques (Figure 3A, Huxley, 1938).

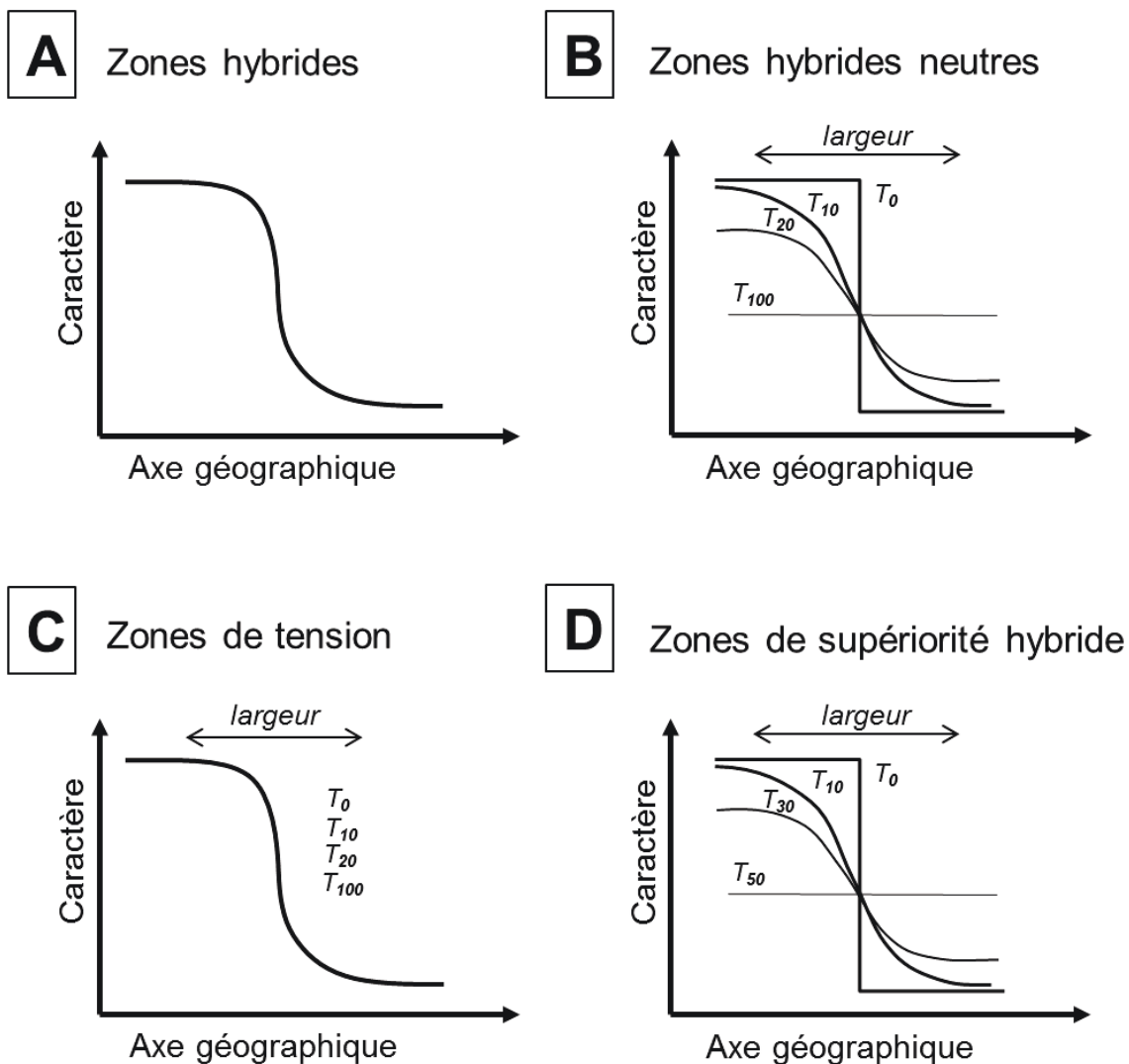


Figure 3 : Les zones hybrides. A : Représentation théorique des zones hybrides par le biais de cline ou variation graduelle et monotone d'un ou plusieurs caractère(s) en fonction d'un axe géographique. B : Représentation théorique des zones hybrides neutres, caractérisées par l'absence de différence significative de valeur sélective entre les génotypes parentaux et hybrides. Ce type de zone s'agrandit (la largeur du cline augmente) en fonction du temps et de la dispersion des individus. C : Représentation théorique des zones de tension, caractéristiques du désavantage aux hétérozygotes (contre-sélection endogène des hybrides). Ce type de zone est considéré comme stable au cours du temps et ne s'agrandit pas. D : Représentation théorique des zones de supériorité hybride. Le profil de ce type de zones ressemble à celui des zones hybrides neutres mais on peut s'attendre dans ce cas là à ce que la largeur des clines observés augmente plus rapidement. L'évolution de ces graphiques théoriques au cours du temps est représentée par les indices de temps suivants : T_0 , T_{10} , T_{20} , T_{30} , T_{50} et T_{100} et leurs courbes associées.

Ces clines peuvent être générés soit de manière primaire — les populations évoluent et accumulent des différences dans une aire de répartition continue — soit de manière secondaire — les populations ont évolué et accumulé des différences dans des

environnements isolés puis entrent en contact dans une même localité géographique (Hewitt, 1988). Bien qu'une multitude de types de zones hybrides aient été décrites dans la littérature, on peut en retenir trois principales (Hewitt, 1988) :

- les zones hybrides neutres (Figure 3B) : zones transitoires générées par le croisement de deux populations divergentes après un contact secondaire, dans lesquelles aucune différence significative de valeur sélective n'est détectée entre génotypes parentaux et hybrides et ayant tendance à s'agrandir (la largeur du cline augmente) en fonction du temps et de la dispersion des individus.

- les zones de tension (Figure 3C) : ces zones caractérisent le désavantage aux hétérozygotes ou la contre-sélection endogène (indépendante de l'environnement) des hybrides (nous verrons les mécanismes sous-jacents dans la **section I.2.2.a**). Considérées comme étroites et stables (elles ne s'agrandissent pas) — un équilibre se crée entre la dispersion des individus occasionnant l'apparition d'hybrides et la contre-sélection de ces derniers — ces zones peuvent se former n'importe où et ont tendance à se déplacer vers les régions les moins densément peuplées ou comportant des barrières à la dispersion (Harrison, 1993).

- les zones de supériorité hybride (Figure 3D) : la recombinaison entre individus divergents génère chez les hybrides des combinaisons génétiques avantageuses d'un point de vue sélectif et se traduisent par une augmentation de leur valeur sélective (nous verrons aussi en détail les mécanismes sous-jacents dans la **section I.2.2.b**). Peu de cas sont répertoriés dans la littérature en raison des difficultés liées à la détection de la propagation dans l'espace de telles combinaisons.

I.2.2 – Conséquences évolutives de l'hybridation

a) La dépression hybride

Dans le passé, l'hybridation a longtemps été considérée comme une impasse évolutive (Mayr, 1963) où les hybrides souffraient d'une diminution des niveaux de viabilité et/ou fertilité (*e.g.* Barton, Hewitt, 1981; Coyne, 1989) à cause de mécanismes génétiques (voir ci-dessous) induisant de la dépression hybride — *i.e.* une valeur sélective plus faible chez les hybrides qu'au sein des populations parentales — (Figure 4A, Turelli, Orr, 2000).

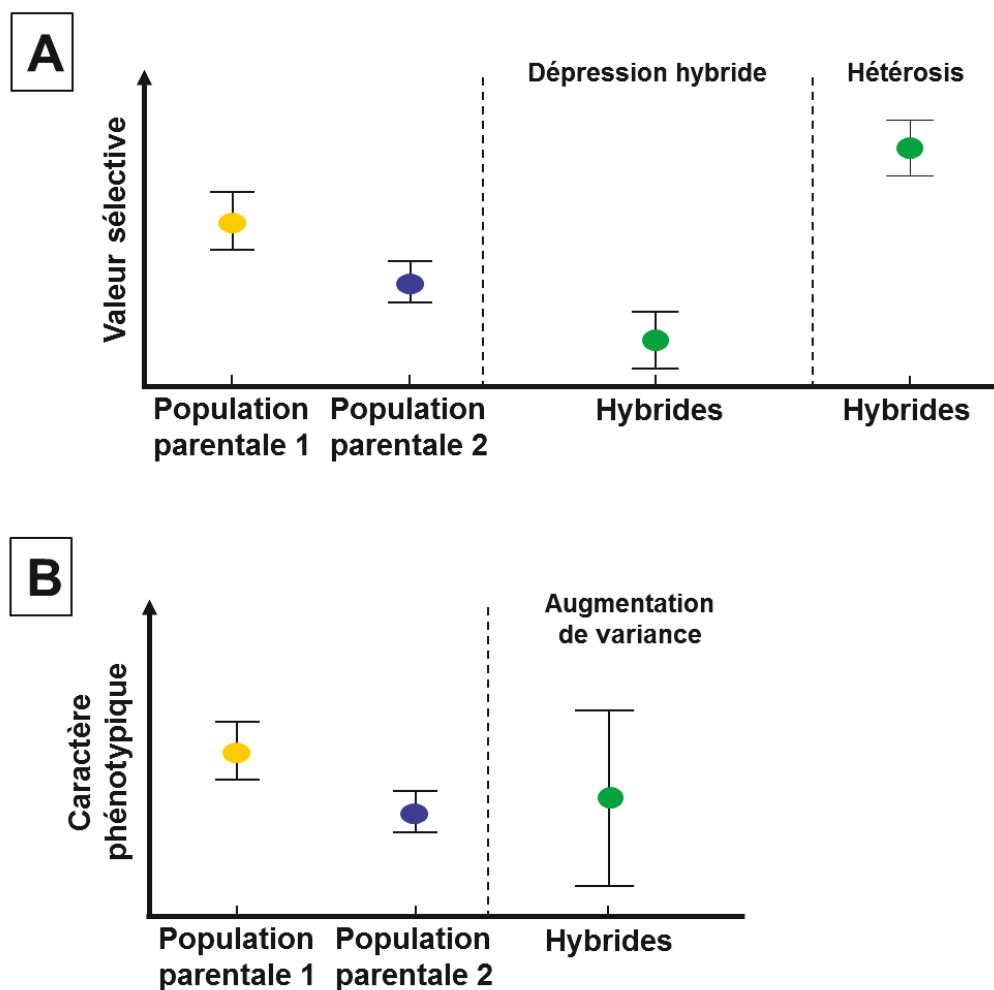


Figure 4 : Conséquences évolutives de l'hybridation. A : Conséquences de l'hybridation sur la valeur sélective moyenne des populations. L'hybridation provoque soit une baisse de la valeur sélective chez les populations hybrides par rapport aux populations parentales (dépression hybride), soit une hausse (vigueur hybride ou hétérosis). B : Conséquence de l'hybridation sur la variance d'un caractère phénotypique. La recombinaison entre deux populations divergentes (ici les populations parentales) peut également entraîner une augmentation de la variance d'un ou plusieurs caractère(s) phénotypique(s) chez les populations hybrides.

Plusieurs causes génétiques de cette dépression hybride peuvent être évoquées :

- la sous-dominance : mécanisme se traduisant par le désavantage des hétérozygotes en raison de la dominance de gènes ayant un impact négatif sur la valeur sélective des individus (Schierup, Christiansen, 1996).

- la rupture de complexes de gènes co-adaptés : la recombinaison provoque une dilution des relations épistatiques positives entre gènes établies par le phénomène d'adaptation locale au sein de chaque population parentale (Tallmon *et al.*, 2004).

- les incompatibilités de type Dobzhansky-Muller (Turelli, Orr, 2000) : les changements génétiques (mutations par exemple) qui apparaissent dans des fonds génétiques différents (des populations isolées) ne sont pas nécessairement compatibles entre eux (Fitzpatrick, 2008). En interagissant, ces changements génétiques indépendants induisent (chez les hybrides) une baisse de fertilité ou de viabilité voire de l'isolement reproducteur postzygotique.

Bien que la dépression hybride puisse apparaître chez les hybrides de première génération (F_1), elle s'observe plus fréquemment chez les hybrides de deuxième génération (F_2) du fait de la recombinaison (Edmands, 1999).

b) La vigueur hybride

En dépit de cette conséquence évolutive négative, de nombreuses autres études ont conduit à considérer l'hybridation comme un processus créatif largement répandu parmi les espèces végétales ou animales (Arnold, 1992; Grant, Grant, 1994; Harrison, 1993; Wang *et al.*, 1997). L'effet contraire de la dépression hybride a notamment été décrit comme l'effet « hétérosis » ou « vigueur hybride » (Figure 4A, Dobzhansky, Gowen, 1952; Shull, 1914) et se définit comme une augmentation de la valeur sélective des hybrides de première génération

(F₁) par rapport aux individus parentaux. Différents mécanismes génétiques peuvent conduire à l'hétérosis (Burke, Arnold, 2001; Lippman, Zamir, 2007) :

- la surperdominance (phénomène monolocus), c'est-à-dire l'avantage des individus hétérozygotes sur les individus homozygotes accumulés sur plusieurs loci.

- la création de nouvelles relations épistatiques positives se traduisant par de nouvelles interactions entre allèles situés à des locus différents et issus de chacun des génotypes parentaux. Lors de l'hybridation, ces allèles sont associés et leur interaction peut conférer aux descendants une valeur supérieure du caractère.

- l'accumulation des relations de dominance de « bons allèles » à de nombreux locus chez les individus hybrides. L'idée est que, dans chaque population parentale, des « bons allèles » sont en forte fréquence à certains loci du fait de la sélection naturelle. Dans la population hybride, on retrouvera les « bons allèles » qui avaient été sélectionnés dans l'une ou l'autre des populations parentales. Le nombre de « bons allèles » est donc mécaniquement plus grand dans la population hybride que dans chacune des populations parentales.

- le masquage des mutations délétères récessives présentes au sein de chacune des populations parentales (voir **section I.3.2**).

c) Innovation génétique et adaptation

Un autre avantage de l'hybridation réside dans le fait que la recombinaison entre populations divergentes peut générer une augmentation de la variabilité génétique *via* la création de nouveaux génotypes et augmenter le spectre phénotypique sur lequel s'opèrera la sélection naturelle (Figure 4B, Facon *et al.*, 2005; Orians, 2000). Cette augmentation de la variabilité génétique pourra ainsi promouvoir l'adaptation des populations hybrides à de nouveaux environnements (Rieseberg *et al.*, 2003).

Par ailleurs, la recombinaison entre populations divergentes peut également favoriser l'apparition de phénotypes extrêmes (*i.e.* qui dépasse de manière significative la valeur seuil d'une mesure faite sur un caractère quantitatif) chez les hybrides, par rapport aux phénotypes parentaux. Ces phénotypes extrêmes, dits « transgressifs », semblent constituer un mécanisme majeur par lequel de nouvelles adaptations apparaissent au sein des lignées hybrides (Rieseberg *et al.*, 1999).

I.2.3 – Etudes de la balance migration/sélection

Les zones hybrides représentent un intérêt particulier pour les chercheurs en biologie évolutive car elles constituent de véritables laboratoires naturels dans lesquels les interactions génétiques et écologiques entre populations peuvent être examinées. Leur étude permet d'obtenir des informations nécessaires à une meilleure compréhension des processus évolutifs tels que la spéciation, la sélection et l'adaptation (Barton, Hewitt, 1985; Harrison, 1993; Hewitt, 1988).

Par ailleurs le développement de la théorie des clines en zone hybride au cours des années 1970-1980 et l'utilisation de marqueurs génétiques moléculaires ont permis d'étudier et estimer chez plusieurs espèces la balance dispersion/sélection. Cette balance reflète la stabilité des zones de tension et s'exprime par l'observation de clines, maintenus par l'effet conjoint de la dispersion (homogénéisant) et de la sélection (diversifiant : les hybrides sont contre-sélectionnés). Des modèles analytiques ont été développés pour prédire la forme du cline (une fonction de sa largeur) en fonction des paramètres de dispersion, σ , (défini comme l'écart-type de la distribution des distances parents-enfants) et l'intensité de sélection, s . Dans le cas où la sélection agit sur un cline par effet de sous-dominance — les hétérozygotes ont une moins bonne valeur sélective que leur parents homozygotes — l'équilibre du cline s'exprime comme ceci :

$$\sigma = \sqrt{\frac{s}{8}}w \quad \{1\}$$

avec w la largeur du cline.

Le problème est que pour estimer σ , l'intensité de la sélection, s , doit être connue. Au cours des années 1980-1990 cependant, les travaux de N.H. Barton sur les clines multilocus à l'équilibre dispersion/sélection ont permis de contourner cette difficulté en estimant σ directement à partir du déséquilibre de liaison (Barton, 1982; Barton, Gale, 1993). Barton a exploité à travers cette technique le fait que la dispersion des populations divergentes génère un déséquilibre de liaison entre loci, maximal au centre de la zone hybride et maintenu par la sélection. Ainsi avec un taux de recombinaison r entre deux loci en déséquilibre de liaison, σ s'obtient ainsi :

$$\sigma^2 = D_{max}rw_1w_2 \quad \{2\}$$

avec D_{max} le déséquilibre de liaison maximal entre deux loci et w_i la largeur du cline au locus i .

L'intensité de la sélection s , peut être déduite à partir de l'équation {1} et de l'estimation de σ obtenue en {2}. I.3 – L'hybridation intra-spécifique au cours des invasions biologiques

I.3.1 – Définition et origine

Le concept d'hybridation intra-spécifique (ou « admixture » en anglais) se définit en biologie évolutive et biologie de l'invasion comme un croisement entre individus de la même espèce appartenant à des populations différenciées génétiquement (Keller, Taylor, 2010; Verhoeven *et al.*, 2011). Ce phénomène peut se produire suite à l'expansion géographique de certaines populations précédemment isolées (*e.g.* Haanes *et al.*, 2011) mais également suite à un processus envahissant impliquant des introductions multiples dans une même localité

(Genton *et al.*, 2005; Kolbe *et al.*, 2004) ou dans des localités géographiques déconnectées (Lavergne, Molofsky, 2007; Lucek *et al.*, 2010; Rieux *et al.*, 2013).

I.3.2 – Stimulus des invasions biologiques

De par les conséquences bénéfiques qu'il peut engendrer au sein des populations introduites, ce phénomène d'hybridation intra-spécifique a été qualifié par certains auteurs de stimulus des invasions biologiques, permettant l'expansion d'une espèce dans un nouvel environnement et *in fine* de favoriser son succès envahissant (Drake, 2006; Ellstrand, Schierenbeck, 2000).

D'après Keller et Taylor (2010) la valeur sélective des populations introduites peut être affectée par trois processus différents, non exclusifs et pouvant agir simultanément. Tout d'abord le croisement entre populations différenciées génétiquement peut résulter en une augmentation de la variance génétique permettant ainsi de contrebalancer la réduction de potentiel adaptatif des populations introduites lors de l'étape d'introduction et de faire face à de nouvelles pressions de sélection imposées par la nouvelle aire envahie (*e.g.* Kolbe *et al.*, 2008; Lavergne, Molofsky, 2007). Deuxièmement, la recombinaison entre populations d'origine distincte peut créer de nouvelles combinaisons génétiques occasionnant de nouveaux phénotypes au sein des individus hybrides qui peuvent être favorisés par la sélection dans le nouvel environnement (*e.g.* Facon *et al.*, 2008). Enfin, une augmentation de la valeur moyenne de caractères phénotypiques positivement corrélés à la valeur sélective des individus (vigueur hybride ou hétérosis, voir **section I.2.2.b**) peut être trouvée chez les génotypes hybrides principalement à cause du masquage du fardeau génétique (Drake, 2006; Moody, Les, 2002). Les mutations délétères récessives accumulées dans chacune des populations introduites se retrouvent à l'état hétérozygote chez les hybrides et ne s'expriment plus.

L'effet attendu de ces processus divers au cours du temps est cependant différent (Keller, Taylor, 2010). L'augmentation du potentiel adaptatif des populations envahissantes *via* l'hybridation intra-spécifique ne serait pas due aux effets individuels (non immédiats) mais plutôt à l'effet conjoint de la sélection et de la recombinaison qui, avec le temps, permettrait d'augmenter la valeur sélective moyenne de la population. De la même manière, la création de phénotypes favorables est un phénomène dynamique qui dépend de la recombinaison au cours des générations. Inversement, les bénéfices de l'hétérosis devraient avoir une incidence immédiate et maximale en première génération (chez les hybrides F₁) et s'estomper au fil des générations sous l'influence des croisements aléatoires (entre génotypes hybrides et parentaux) et par conséquent de la diminution du nombre d'individus hétérozygotes (Drake, 2006; Lynch, 1991).

I.3.3 – Remise en cause

Bien que ce type d'hybridation soit souvent considéré comme un effet stimulant des invasions biologiques (Drake, 2006; Ellstrand, Schierenbeck, 2000), certains auteurs estiment que cet effet est souvent surestimé. Les résultats obtenus en 2012 par Benvenuto *et al.* (2012) dans le cadre d'un programme de lutte biologique contre la mouche de l'olive, *Bactrocera oleae*, ne montrent aucun signe d'hétérosis chez l'auxiliaire de lutte biologique *Psytalia lounsburyi* (guêpe parasitoïde) à la suite de croisements entre des populations parentales affichant pourtant des différences marquées en termes de stratégies d'histoire de vie. Encore plus récemment, Chapple *et al.* (2013) ont montré que le lézard *Lompropholis delicata* avait été introduit à plusieurs reprises dans des régions de l'océan Pacifique (Hawaii, Nouvelle Zélande et les îles Lord Howe) à partir de l'Australie, en présence ou absence du phénomène d'hybridation intra-spécifique. De plus, nous avons vu plus haut que le risque existe pour que

des hybridations mènent à l'effet inverse d'un effet positif (la dépression hybride) (*e.g.* Edmands, 1999; Marshall, Spalton, 2000).

I.3.4 – Estimation de la dispersion dans le cas d'une diffusion pure

Les invasions biologiques sont des expérimentations naturelles de dynamique des populations (n'étant pas à l'équilibre) et dans lesquelles la dispersion peut être estimée directement à partir de situations transitoires.

Des modèles théoriques adéquats peuvent en effet être utilisés pour estimer la dispersion directement à partir de la description de la structure génétique spatiale neutre de la zone de contact entre deux populations en expansion et sa dynamique temporelle (Endler, 1977). Dans la littérature cependant, les invasions biologiques ont rarement été utilisées pour estimer la dispersion. Seulement une étude semble avoir utilisé cette approche (Rieux *et al.*, 2013) bien que ce contexte biologique soit courant et que la méthodologie soit relativement simple à implémenter. Les invasions biologiques sont fréquemment initiées par des introductions multiples dans les localités géographiquement déconnectées (Bossdorf *et al.*, 2005; Ciosi *et al.*, 2008; Dlugosch, Parker, 2008; Facon *et al.*, 2003 ; Kolbe *et al.*, 2004; Lockwood *et al.*, 2005; Miller *et al.*, 2005; Roman, Darling, 2007). L'effet conjugué des goulots d'étranglement associés aux introductions et/ou de la dérive lors de la phase d'établissement peut provoquer une importante différenciation génétique entre les populations introduites (*e.g.* Ciosi *et al.*, 2008). L'expansion spatiale de ces populations peut par la suite conduire à des contacts secondaires et éventuellement former des clines transitoires de fréquences alléliques (Rieux *et al.*, 2013). Dans le cas d'une zone hybride neutre —*i.e.* dire une zone dans laquelle les individus dispersent sans que leur valeur sélective soit impactée par l'hybridation — les fréquences alléliques (mesurées sur des marqueurs neutres) ont tendance à s'homogénéiser au travers de la zone de contact au cours du temps.

Sous l'hypothèse d'une migration des individus suivant un mécanisme de diffusion — *i.e.*, une migration aléatoire des individus, isotrope et s'effectuant selon un processus local — J.A. Endler a établi mathématiquement la relation entre la largeur du cline (w et correspondant à l'inverse de la pente maximale du cline mesuré), le temps T compté en génération depuis le contact initial entre les deux populations et le paramètre de dispersion σ (Endler, 1977) :

$$\sigma = \frac{w}{1.68\sqrt{T}} \quad \{3\}$$

Notons que la relation mathématique initiale d'Endler est inexacte et a été corrigée récemment pour donner la formule suivante (Gay *et al.*, 2008; Rieux *et al.*, 2013) :

$$\sigma = \frac{w}{\sqrt{2\pi T}} \quad \{4\}$$

Cette relation mathématique indique que la largeur du cline augmente proportionnellement à la dispersion des individus et à la racine du temps depuis le contact.

Par ailleurs, la dispersion peut également être estimée en utilisant l'approche du déséquilibre de liaison développée par Barton et Gale (Barton, 1982; Barton, Gale, 1993) et présentée en **section I.2.3**. Deux différences avec l'utilisation précédente du déséquilibre de liaison (voir l'équation {2}) peuvent être néanmoins soulignées. Premièrement dans le cas d'une zone hybride neutre, le déséquilibre de liaison n'est pas maintenu par la sélection et aura tendance à disparaître au cours de temps en fonction de la recombinaison et de la dispersion. Par ailleurs, toujours dans le cas d'une zone hybride neutre, l'ensemble des clines observés aux différents marqueurs utilisés sont coïncidents (même centre) et concordants (même pente ou même largeur). En conséquence une seule largeur est estimée et la relation {2} devient :

$$\sigma^2 = D_{max}rw^2 \quad \{5\}$$

II. – LE MODÈLE BIOLOGIQUE : LA CHRYDOMÈLE DES RACINES DU MAÏS, *DIABROTICA VIRGIFERA* *VIRGIFERA*

II.1 – Généralités sur la biologie de la chrydomèle des racines du maïs

Pour répondre à l'objectif global de cette thèse, nous avons choisi d'utiliser un modèle biologique offrant des situations propices à l'étude des différents concepts que nous avons présentés en introduction : il s'agit de la chrydomèle des racines du maïs, *Diabrotica virgifera virgifera* (Dv), Figure 5) LeConte (Coleoptera : Chrysomelidae).

Dv constitue l'un des plus gros ravageurs du maïs cultivé, *Zea mays* L., aux Etats-Unis (Levine, Oloumi-Sadeghi, 1991) et une peste majeure en Europe (Sappington *et al.*, 2006). Les dégâts qui lui sont associés (en termes de pertes agricoles ou d'investissement pour tenter de la contenir ou de l'éradiquer) dépassent le milliard de dollars par an à l'échelle mondiale (Spencer *et al.*, 2009; Wessler, Fall, 2010). Ces dégâts se traduisent par de fortes baisses de rendement à l'échelle de la plante et des champs et sont principalement causés par les larves. En se nourrissant exclusivement du réseau racinaire du maïs, ces dernières réduisent les apports en eau et nutriments nécessaires au bon développement de la plante et déstabilisent son ancrage dans le sol (Levine *et al.*, 2002). Sous l'effet du vent ou de la pluie, les plants de maïs ainsi attaqués sont plus sujets à la verse et il devient alors difficile de les récolter. Par ailleurs, les pertes de rendement peuvent aussi être occasionnées par les adultes

qui en consommant certaines parties de la plante (le pollen, les soies, les grains) perturbent la pollinisation du maïs au sein du champ (Levine, Oloumi-Sadeghi, 1991) ou rendent tout simplement la plante impropre à la consommation.

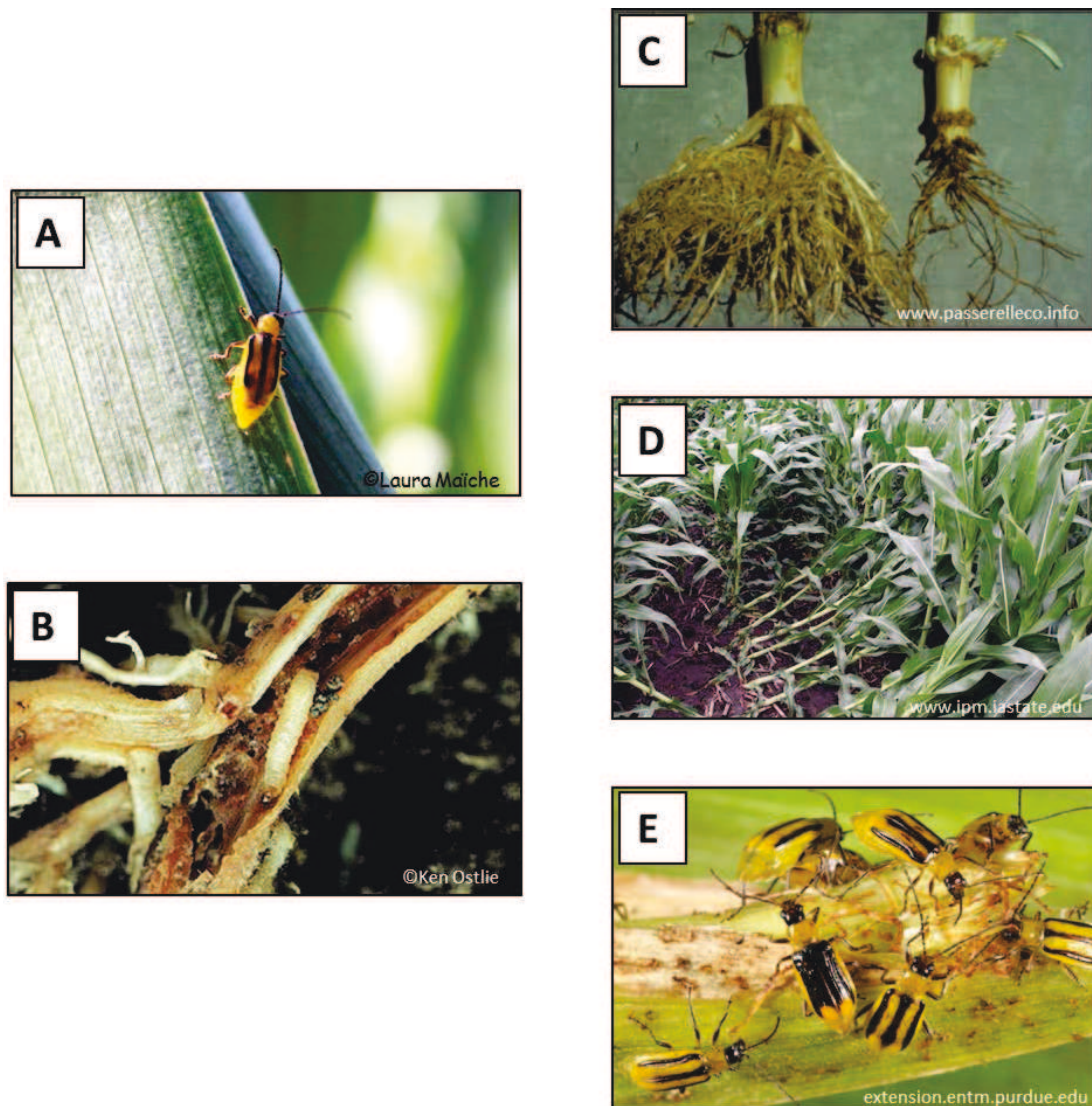


Figure 5 : La chrysomèle des racines du maïs, *Diabrotica virgifera virgifera* (*Dvv*) et ses dégâts associés dans les champs de maïs. A : Adulte de *Dvv*. B : Larve de *Dvv* à la recherche de nourriture dans le réseau racinaire du maïs. C : Réseau racinaire du maïs sain et attaqué par *Dvv*, respectivement de gauche à droite. D : Affaiblissement des plants de maïs (verse) provoqué par la consommation de leur racines par les larves de *Dvv*. E : Adultes de *Dvv* en train de se nourrir de plusieurs parties du plant de maïs (soies, grains, pollen, feuille).

Le cycle de vie de *Dvv* (Figure 6) est décrit par Branson et Krysan comme étant univoltin (1981). Il commence en été, par la ponte d'œufs blancs et ovales (de moins d'un mm

de longueur) dans le sol. Ces derniers hivernent alors jusqu'à leur éclosion au printemps. Les larves migrent ensuite dans le sol vers les racines des jeunes plants de maïs et s'en nourrissent jusqu'à ce que leur développement soit complet. Elles s'éloignent ensuite de leur source de nourriture pour construire entre la fin du printemps et le début de l'été, une cellule de terre dans laquelle elles réaliseront leur transformation en nymphe de couleur blanche. L'émergence des adultes a lieu de fin mai à début août selon les régions et les conditions climatiques. Les adultes meurent à l'automne lors de l'apparition des premiers froids, après avoir pondu.

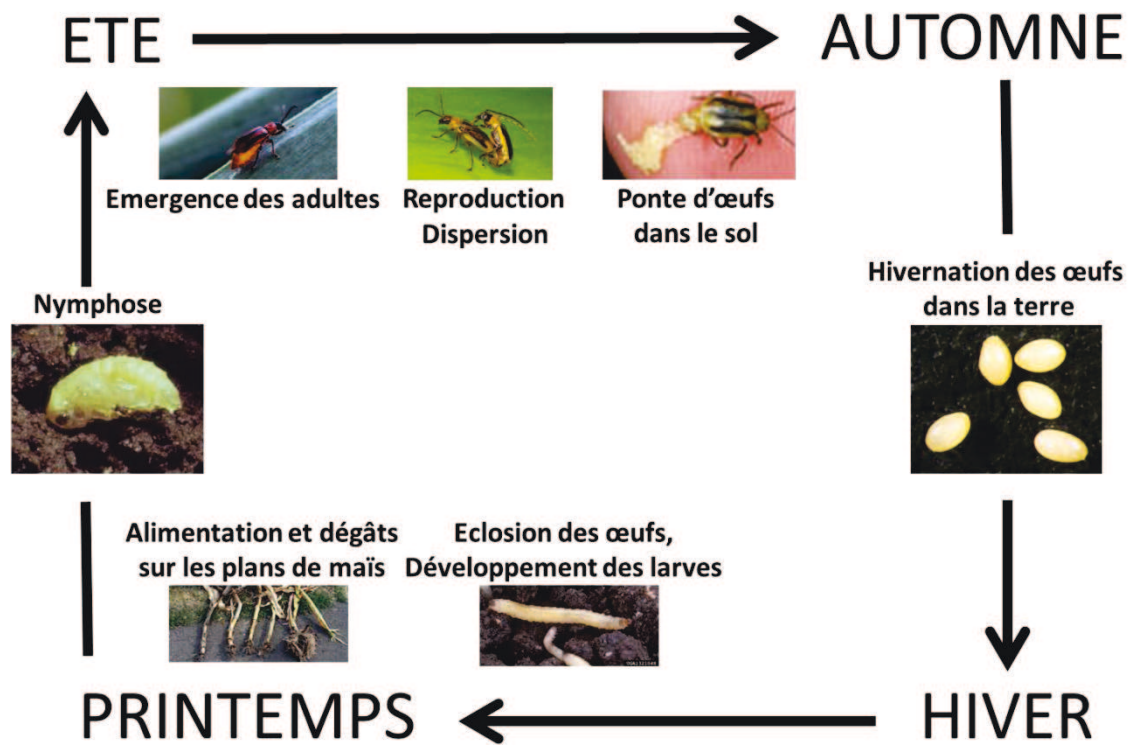


Figure 6 : Cycle de vie de *Diabrotica virgifera virgifera*.

II.2 – Aire native et invasion

La chrysomèle des racines du maïs serait originaire d'Amérique centrale (Branson, Krysan, 1981; Smith, 1966). Elle a probablement évolué avec le maïs au Mexique et atteint (conjointement avec ce dernier) ce qui est aujourd'hui le Sud-Ouest des Etats-Unis il y a environ 3 000 ans (Figure 7A, d'après Krysan, Smith, 1987). Aux Etats-Unis, *Dvv* a été collectée pour la première fois au milieu du 19^e siècle à l'ouest de l'état du Kansas par Leconte. Ce n'est que cinquantaine d'années plus tard, en 1909, que *Dvv* fut décrite comme une peste (Gillette, 1912). Durant la seconde moitié du 20^e siècle, *Dvv* a rapidement étendu son aire de répartition au travers des régions céréalières du « MidWest » américain pour finalement atteindre la côte Nord-Est des USA à la fin des années 1980 (Figure 7B, revue dans Gray *et al.*, 2009).

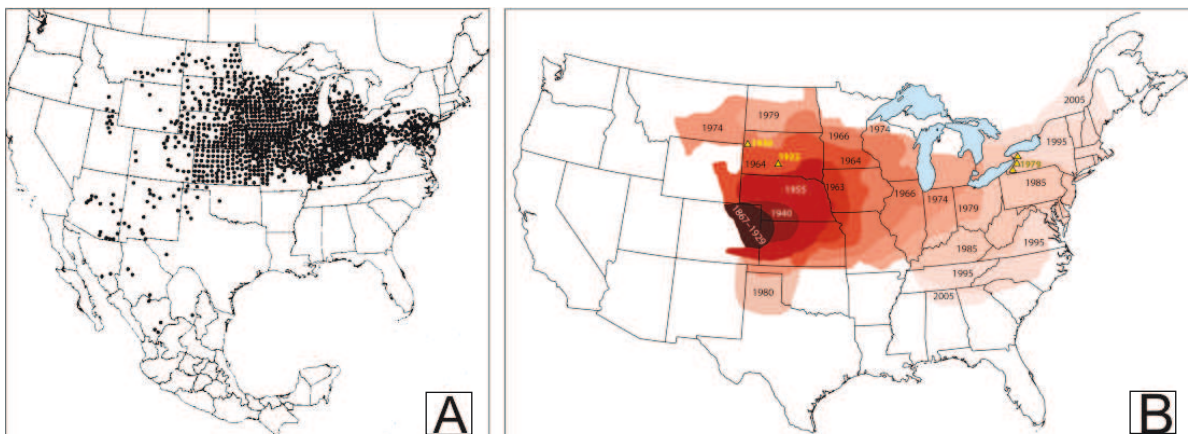


Figure 7 : Aire native et invasion de *Diabrotica virgifera virgifera* (*Dvv*) sur le continent américain. A : Aire de répartition de *Dvv* en Amérique du Nord (issu de Ciosi, 2009). B : Reconstruction de l'expansion géographique de la chrysomèle des racines du maïs, *Dvv* en Amérique du Nord entre 1867 et 2005 (issu de Gray *et al.*, 2009). Les triangles et dates en jaune indiquent des observations de *Dvv* en amont de l'année indiquée par la distribution d'établissement.

A la fin du 20^e siècle et au début des années 2000, *Dvv* a été introduite en Europe de manière répétée en provenance d'Amérique du Nord (Figure 8 et 9A, Ciosi *et al.*, 2008; Miller *et al.*, 2005). Elle a été observée pour la première fois en Europe, près de l'aéroport de Belgrade en Serbie, en 1992 (Sivcev *et al.*, 1994). Un réseau international de surveillance a

depuis été mis en place pour suivre son aire de répartition en Europe et son expansion annuelle (Kiss *et al.*, 2005). Deux types de zones infestées furent alors identifiés : (i) les zones d'expansion géographiques continues en Europe Centrale et du Sud-Est (Europe CSE), en Italie du Nord-Ouest (Italie NO) et en Bavière au Sud de l'Allemagne, correspondant à des « foyers envahissants » et (ii) plusieurs petits foyers déconnectés qui n'ont pas persisté dans le temps (foyers éteints) ou qui se sont établis mais n'ont pas connu d'expansion géographique comme celui d'Ile de France (Ciosi *et al.*, 2008; Miller *et al.*, 2005).

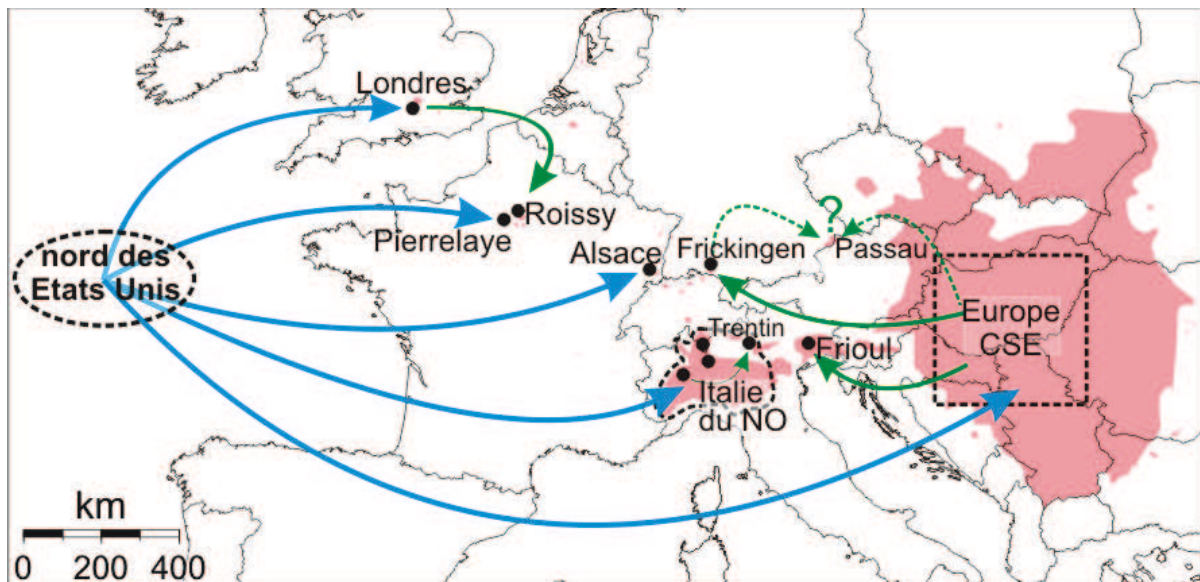


Figure 8 : Routes d'invasion de *Diabrotica virgifera virgifera* en Europe. Les introductions inter et intra-contininentales sont représentées par les flèches bleues et vertes, respectivement. Les incertitudes relatives aux routes d'introduction de la population de Passau sont représentées par des flèches en pointillés et un point d'interrogation. Chaque délimitation en pointillés noirs représente une zone où la structure génétique est faible et qui a été traitée comme une population. Italie NO : Italie du Nord-Ouest, Europe CSE : Europe Centrale et du Sud-Est. D'après (Ciosi, 2009; Ciosi *et al.*, 2008).

Le foyer le plus important de *Dvv* en Europe est celui d'Europe CSE et s'étend actuellement sur 16 pays, de l'Autriche à l'Ukraine et du Sud la Pologne au Nord de la Grèce. Ce foyer s'est par ailleurs étendu vers l'Ouest en Slovénie puis en Italie, englobant ainsi le foyer d'Italie NE. Parallèlement, le foyer envahissant d'Italie NO a progressé vers l'Est. Depuis 2008, ces deux grands foyers envahissants, différenciés génétiquement (Ciosi *et al.*,

2008), semblent être entrés en contact dans la région de la Vénétie (Figure 9A, Edwards, Kiss, 2008; Edwards, Kiss, 2009). La répartition géographique actuelle de *Dvv* entre ces foyers est désormais continue en Italie du Nord (Figure 9B Edwards, Kiss, 2010).

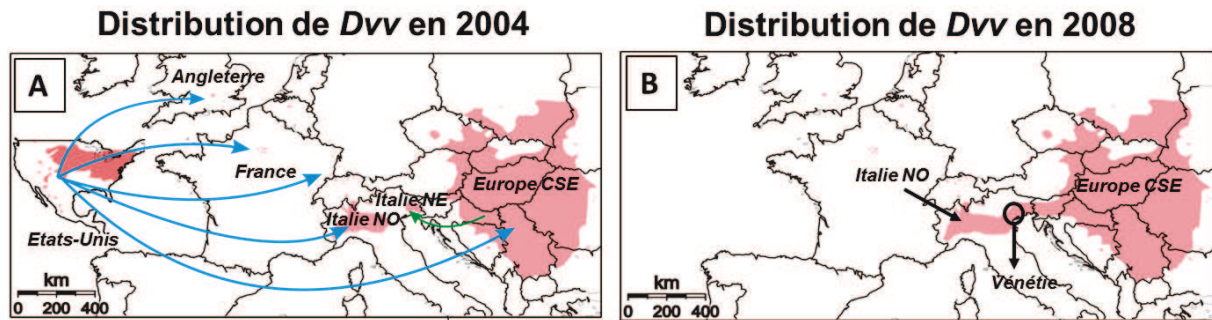


Figure 9 : Routes d'invasion et évolution de l'aire de répartition de *Diabrotica virgifera virgifera* (*Dvv*) en Europe, entre 2004 et 2008. A : Fondation des différents foyers européens et aire de répartition de *Dvv* en 2004 (modifiée à partir de Ciosi *et al.* (2008) et Edwards and Kiss (2005)). B : Aire répartition de *Dvv* en 2008, désormais continue entre les foyers d'Italie du Nord-Ouest (Italie NO), d'Italie du Nord-Est (Italie NE) et d'Europe Centrale et du Sud-Est (Europe CSE) suite à leur contact en Vénétie (modifiée à partir de Edwards et Kiss (2009)). Le cercle noir représente la zone de contact localisée en Italie du Nord dans la région de la Vénétie.

II.3 – Méthodes de lutte et résistance

A l'échelle mondiale il existe actuellement trois principaux modes de lutte contre la chrysomèle des racines du maïs : la rotation des cultures, les traitements insecticides et enfin l'utilisation du maïs transgénique. Après avoir permis de contrôler efficacement les populations de *Dvv* pendant plusieurs années, chacune de ces méthodes a cependant été mise en échec par les capacités d'adaptation du ravageur. Je décris brièvement ici chacune de ces stratégies de lutte.

La rotation des cultures : le principe de cette méthode de lutte consiste à alterner les cultures d'une année sur l'autre (au sein d'une même parcelle) afin de casser le cycle biologique du ravageur. Les œufs pondus en été par les femelles — exclusivement dans les champs de maïs — éclosent l'année suivante dans un champ d'une autre culture (généralement du soja aux USA) dont les racines ne sont pas la source de nourriture des larves

de *Dv*v. Par ailleurs, le champ de soja n'étant pas un site de ponte attractif pour les femelles, aucun œuf n'y sera pondu au cours de l'été et les larves de *Dv*v seront donc absentes du champ de maïs au cours de l'année suivante. Aux USA, cette méthode de lutte historique — mise en place dès 1910 jusqu'au milieu des années 1990 — a permis de contrôler efficacement les populations de *Dv*v pendant près d'un siècle avant d'être sérieusement remise en question à partir des années 1980 (Gray *et al.*, 2009; Levine, OloumiSadeghi, 1996; Levine *et al.*, 2002; Spencer *et al.*, 1997). Cette réduction d'efficacité de la rotation des cultures est probablement due à une modification du comportement de ponte des femelles, qui les pousse à pondre dans des cultures autres que le maïs (Levine *et al.*, 2002; Spencer *et al.*, 1997). En Europe et à ce jour, il n'existe pas de phénomène de contournement de la rotation des cultures chez cette espèce.

Les traitements insecticides : la lutte chimique contre *Dv*v constitue une approche complémentaire de la rotation des cultures. Elle s'applique à même le sol contre les larves (au moment des semis par exemple) et au niveau aérien contre les adultes (épandages réalisés en ULM ou par avion). Cette lutte chimique n'a cessé d'évoluer depuis les années 1940, sous l'effet conjoint des mécanismes de résistance du ravageur et de la prise de conscience des risques phytosanitaires qu'elle engendre. Les insecticides utilisés contre *Dv*v sont regroupés en différentes classes du fait de leur propre spécificité mais perturbent tous l'influx nerveux de l'insecte. Ils provoquent la paralysie (plus ou moins immédiate selon l'insecticide) puis la mort ou la mort instantanée (choc neurotoxique) de l'individu. Aux USA, la classe des organochlorés a d'abord été recommandée au début des années 1940 (Hill *et al.*, 1948) pour contrôler les populations de *Dv*v avant d'être abandonnée dans les années 1960, suite aux premiers signes de résistance développés par le ravageur. Les insecticides organophosphorés et carbamates ont donc été introduits dans les années 1970 pour contrebalancer les problèmes liés à l'émergence de la résistance aux organochlorés et en raison de leur plus faible

persistance dans l'environnement. Néanmoins plusieurs cas de résistance à ces insecticides ont été détectés à partir des années 1990 (Meinke *et al.*, 1998; Wright *et al.*, 1996). En Europe, les organochlorés ont rarement été appliqués mais les organophosphorés, les carbamates, les pyréthriinoïdes et les néonicotinoïdes (aussi présents aux USA) sont actuellement utilisés de manière conjointe (Sivcev *et al.*, 2012). Il a récemment été montré que plusieurs populations d'Europe centrale étaient résistantes aux organochlorés et sensibles aux organophosphorés (Ciosi *et al.*, 2009).

Le maïs transgénique : cette technique consiste à transférer au génome de la plante hôte un ou plusieurs gènes issus de la bactérie *Bacillus thuringiensis* et codants pour une famille de toxines destinées à *Dv* (CRYBb1, CRY34/24Ab1...). Une fois intégrés, les gènes seront exprimés de manière constitutive et la plante produira donc de manière autonome les différentes entomotoxines. Ces dernières seront alors ingérées par les larves de *Dv* au niveau des racines de la plante. L'ingestion de ces toxines perturbe le système digestif de l'insecte et stoppe sa prise alimentaire. En conséquence les larves finissent par mourir. Le réseau racinaire ainsi « protégé » est plus conséquent chez les plantes transgéniques, diminuant ainsi leur risque de verse. A partir de 2003 et compte tenu de l'inefficacité des méthodes de lutte conventionnelles (insecticides et rotation des cultures), les agriculteurs américains ont rapidement adopté le maïs transgénique *Bt* exprimant CRY3Bb1 (du nom de la bactérie *Bacillus thuringiensis* et commercialisé par Monsanto) comme moyen de lutte contre *Dv* (James, 2009). Moins d'une dizaine d'années après sa commercialisation, des chercheurs de l'université de l'Iowa ont constaté que des mécanismes de résistance à l'entomotoxine CRY3Bb1 avaient évolué chez *Dv* (Gassmann *et al.*, 2011). D'après les auteurs de cette étude, deux facteurs contribuant chacun au développement de la résistance peuvent expliquer ce constat d'échec : le gène conférant la résistance à CRY3Bb1 est non-récessif et la proportion de zone refuges — zone dans laquelle du maïs non-Bt est planté et des individus

sensibles sont présents — à l'intérieur des champs de maïs est insuffisante dans le « MidWest » américain. La combinaison de ces deux facteurs n'a donc pas permis dans ce cas une dilution optimale du caractère résistant à CRY3Bb1 chez *DvV*. Aucun maïs *Bt* exprimant des toxines destinées à *DvV* n'est actuellement cultivé en Europe.

III. – OBJECTIFS DE LA THÈSE

De manière globale mon projet de thèse s'inscrit dans le cadre des invasions biologiques et a été réalisé au sein d'une équipe qui, par des approches génétiques, démographiques et évolutives, cherche à expliquer les raisons du succès ou de l'échec de ces invasions biologiques.

Le modèle biologique sur lequel j'ai travaillé durant ces trois années de thèse était déjà étudié au laboratoire avant mon arrivée et les études réalisées ont permis de retracer pour la première fois ses routes d'invasion (Ciosi *et al.*, 2008; Miller *et al.*, 2005). Cet effort descriptif réalisé au sein de l'équipe par différents travaux de thèse et post-doctorat est primordial et constitue une première étape en vue de comprendre les mécanismes complexes mis en jeu lors d'une invasion biologique. Ces travaux ont abouti, entre autres, à une description précise de la mise en place des principaux foyers de *Dvv* en Europe.

Ma thèse s'inscrit dans la continuité de ces travaux et son objectif global consistait à détecter, étudier et documenter un cas précis d'hybridation au cours d'une invasion biologique. [Ici et dans la suite de ce manuscrit, le terme hybridation se rapportera uniquement à l'hybridation intra-spécifique.] Cet objectif global s'est décliné au cours de ma thèse en trois objectifs particuliers. Dans un premier temps, j'ai cherché à prouver l'existence d'un événement d'hybridation en Italie du Nord (région de la Vénétie) entre les deux principaux foyers envahissants de *Dvv* en Europe. Ensuite, j'ai exploité cette zone hybride pour la caractériser et estimer un paramètre clé des invasions biologiques : la dispersion. Enfin, j'ai étudié l'impact de cet exemple d'hybridation sur la valeur sélective des individus pour en déterminer les éventuelles conséquences évolutives et dynamiques sur l'invasion européenne de *Dvv*.

La situation que j'ai étudiée est assez rare et en ce sens, mérite d'être soulignée : dans la nature, les cas d'hybridation s'étudient souvent *a posteriori* (Barton, Hewitt, 1985; Hewitt, 1988) et il est donc peu courant de pouvoir en suivre une pratiquement en temps réel. Par ailleurs la proximité entre la zone géographique considérée et mon équipe d'accueil a permis un échantillonnage régulier tout au long de la thèse et représentait donc un cadre d'étude quasi-parfait !

IV. – CHAPITRE I : LA VÉNÉTIE, UNE ZONE HYBRIDE ENTRE DEUX FOYERS ENVAHISSANTS EUROPÉENS DE *DIABROTICA* *VIRGIFERA VIRGIFERA*

Ce premier chapitre repose sur une étude de génétique des populations réalisée à l'aide de marqueurs microsatellites et dans laquelle nous cherchons à détecter un événement d'hybridation au cours de l'invasion européenne de *Dvv*. Cette étude se rapporte à l'**article n°1**.

Article n°1 : **Bermond, G.**, Ciosi, M., Lombaert, E., Blin, A., Boriani, M., Furlan, L., Toepfer S.; Guillemaud, T. (2012). Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe. *PloS one*, 7(11), e50129.

Contexte, objectif visé et questions posées

Comme je l'ai déjà indiqué dans la **section III.2** de l'introduction, *Dvv* a été introduite de manière multiple et indépendante en Europe en provenance des USA (Ciosi *et al.*, 2008; Miller *et al.*, 2005). Ces introductions multiples à partir des années 1990 ont conduit à la formation de plusieurs foyers envahissants européens dont deux principaux, fortement différenciés génétiquement (Ciosi *et al.*, 2008) : le foyer d'Italie du Nord-Ouest (Italie NO) et

celui d'Europe Centrale et du Sud-Est (Europe CSE). La convergence géographique de ces deux foyers a entraîné à partir de 2008, une zone de contact localisée près de la région de la Vénétie en Italie du Nord (Figure 10, Edwards, Kiss, 2008; Edwards, Kiss, 2009). L'objectif de cette première étude est de démontrer que cette zone de contact est une zone hybride, c'est-à-dire une zone dans laquelle des individus d'origine différente se croisent, s'accouplent et produisent des descendants hybrides. Les questions sous-jacentes sont les suivantes :

- 1) Est-ce que la zone de contact localisée en Italie du Nord entre les foyers d'Italie NO et d'Europe CSE est une zone hybride ?
- 2) Si la présence d'individus hybrides est détectée en Vénétie, assiste-t-on à une augmentation des niveaux de variabilité génétique neutre comme attendu lors d'évènements d'hybridation entre populations différenciées génétiquement ? Assiste-t-on à une restauration de la variabilité génétique de la population source ?
- 3) Quelle est l'origine des différents échantillons vénitiens ?
- 4) Peut-on caractériser cette éventuelle zone hybride et en définir des paramètres géographiques (centre, largeur)?

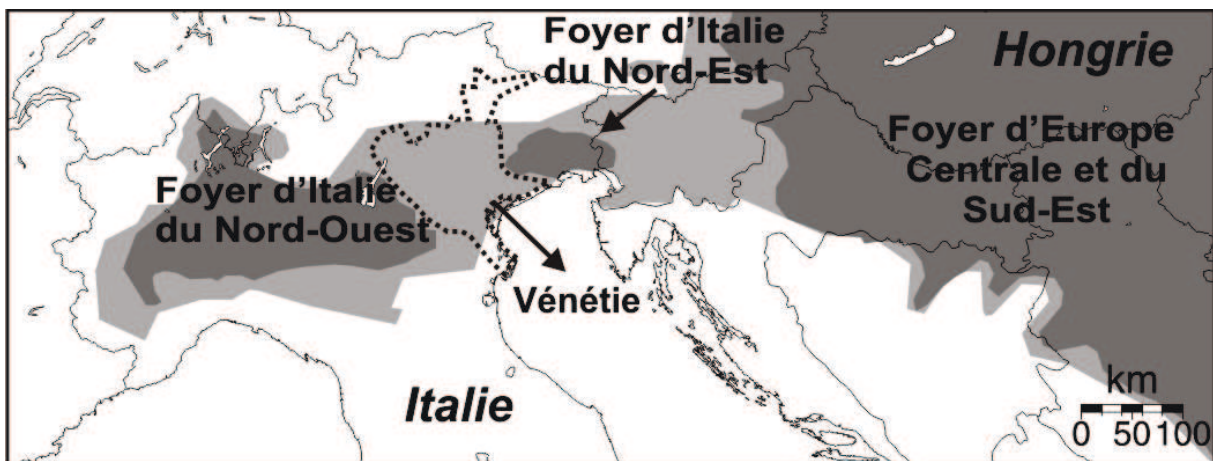


Figure 10 : Aire de répartition de *Diabrotica virgifera virgifera* (Dvv) avant (2005) et après (2010) le contact secondaire entre les foyers d'Italie du Nord-Ouest et d'Europe Centrale et du Sud-Est. Les aires de répartition de Dvv avant (2005) et après (2010) le contact sont respectivement représentées en gris foncé et gris clair. La délimitation en pointillés représente la surface administrative de la région de Vénétie, où s'est produit le contact.

Matériel utilisé et Méthodes employées

Le matériel biologique que nous avons utilisé dans cette étude se composait globalement de plusieurs échantillons collectés sur un transect Ouest-Est, au travers des différents foyers parentaux (Italie NO et Europe CSE) et de la zone de contact (Vénétie) (Figure 11). Les individus appartenant à chacun de ces échantillons ont dans un premier temps été génotypés à 13 loci microsatellites. Les génotypes et fréquences alléliques à ces mêmes loci ont ensuite été analysés. Pour tenter de répondre à toutes les questions précédemment évoquées et atteindre *in fine* l'objectif visé, différentes approches de génétique des populations ont été abordées :

- par une approche classique de génétique des populations, nous avons tout d'abord évalué les niveaux de variabilité génétique neutre des différentes populations échantillonnées. Pour cela, l'hétérozygotie moyenne attendue et le nombre moyen d'allèles par locus ont été mesurés et comparés entre types de populations (Italie NO, Europe CSE et Vénétie).

- nous avons ensuite mené une analyse bayésienne de la structure des populations des échantillons de *Dvv* afin de déterminer le nombre de clusters génétiquement homogènes sur l'ensemble du transect considéré. Dans la mesure où ce nombre correspond à ce que l'on attend dans le cas d'un contact secondaire (deux clusters génétiquement homogènes, Ouest et Est, correspondant respectivement aux foyers parentaux Italie NO et Europe CSE) l'estimation d'un taux d'hybridation individuel est envisagée.

- en utilisant une approche ABC (pour Approximate Bayesian Computation voir **section I.1.5.b**), nous avons par la suite comparé différents scénarios historiques (mettant en jeu ou non des évènements d'hybridation) permettant d'expliquer l'origine des différents échantillons situés en Vénétie.

- enfin nous avons testé si la distance géographique relative à un point de référence (défini ici de manière arbitraire comme étant l'échantillon situé le plus à l'Ouest) avait un effet sur la répartition des fréquences alléliques le long du transect.

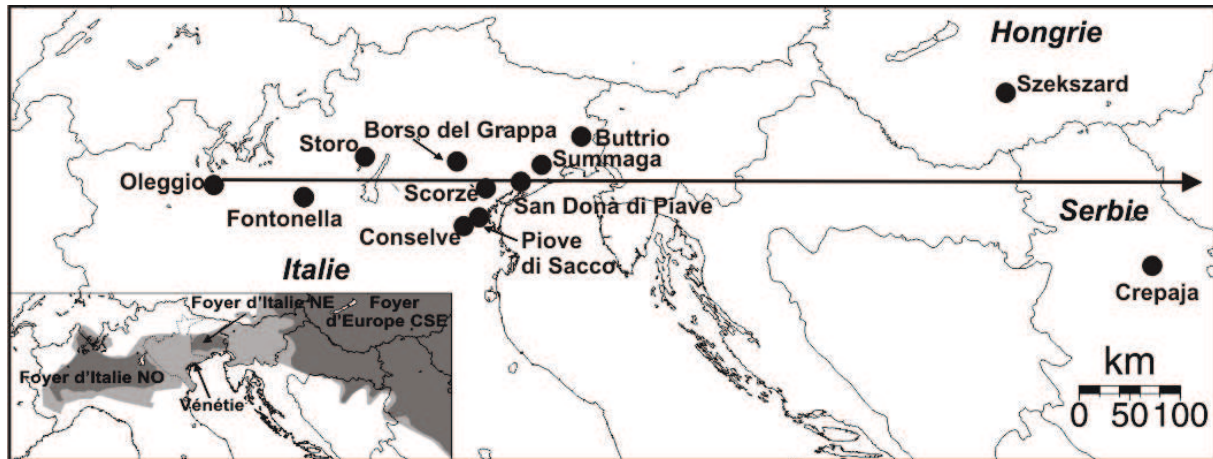


Figure 11 : Localisation des échantillons (cercles noirs) de *Diabrotica virgifera virgifera* (*Dvv*) en Europe : en Italie du Nord-Ouest (Italie NO), dans la zone de contact (Vénétie), en Italie du Nord-Est (Italie NE) et en Europe Centrale et du Sud-Est. La flèche noire indique l'axe géographique utilisé pour les analyses spatiales. Sur la petite carte, les aires de répartition de *Dvv* avant (2005) et après (2010) le contact entre les populations des foyers Ouest et Est sont représentées en gris foncé et gris clair, respectivement. La délimitation en pointillés représente la surface administrative de la région de Vénétie, où s'est produit le contact.

Résultats

Les résultats que nous avons obtenus grâce à ces différentes approches montrent que (i) la variation génétique est plus importante dans la zone de contact que dans chacun des foyers parentaux et que le croisement des deux foyers envahissants a permis de restaurer en moyenne 8% de la variabilité génétique de la population source ; (ii) l'analyse bayésienne de la structure des populations a défini deux clusters génétiquement homogènes comme attendu et révèle la présence d'individus hybrides dans certains échantillons vénitiens ; (iii) des scénarios évolutifs mettant en jeu des événements d'hybridation ont été retenus par la méthode ABC, pour décrire l'origine de la majorité des échantillons vénitiens (Borso del Grappa, Scorzè, San Donà di Piave et Summaga) alors qu'une origine simple (le foyer d'Italie NO) a été retenue pour les échantillons situés dans la partie occidentale de la Vénétie (Conselve et Pieve di Sacco) ; (iv) des clines (défini dans la **section I.2.1.a**) de fréquences

alléliques ont été observés aux loci microsattellites considérés et ont permis une première caractérisation géographique de la zone hybride : un centre localisé aux alentours de 300 km du point de référence (ici Oleggio) et une largeur d'environ 100 km.

Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe

Gérald Bermond^{1,2,3*}, Marc Ciosi^{4,5}, Eric Lombaert^{1,2,3}, Aurélie Blin^{1,2,3}, Marco Boriani⁶, Lorenzo Furlan⁷, Stefan Toepfer⁸, Thomas Guillemaud^{1,2,3}

1 INRA, UMR 1355 Institute Sophia Agrobiotech, Equipe "Biologie des Populations Introduites", Sophia Antipolis, France, **2** Université de Nice Sophia Antipolis, UMR Institute Sophia Agrobiotech, Equipe "Biologie des Populations Introduites", Sophia Antipolis, France, **3** CNRS, UMR 7254 Institute Sophia Agrobiotech, Equipe "Biologie des Populations Introduites", Sophia Antipolis, France, **4** Institute III, University of Glasgow, Glasgow Biomedical Research Centre, Glasgow, United Kingdom, **5** MBB Unit, International Centre of Insect Physiology and Ecology, Nairobi, Kenya, **6** Servizio fitosanitario regionale, Regione Lombardia, Milano, Italy, **7** Veneto Agricoltura, Legnaro, Italy, **8** CAB International, c/o Plant Protection Directorate, Hodmezovasarhely, Hungary

Abstract

The western corn rootworm, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae), is one of the most destructive pests of corn in North America and is currently invading Europe. The two major invasive outbreaks of rootworm in Europe have occurred, in North-West Italy and in Central and South-Eastern Europe. These two outbreaks originated from independent introductions from North America. Secondary contact probably occurred in North Italy between these two outbreaks, in 2008. We used 13 microsatellite markers to conduct a population genetics study, to demonstrate that this geographic contact resulted in a zone of admixture in the Italian region of Veneto. We show that i) genetic variation is greater in the contact zone than in the parental outbreaks; ii) several signs of admixture were detected in some Venetian samples, in a Bayesian analysis of the population structure and in an approximate Bayesian computation analysis of historical scenarios and, finally, iii) allelic frequency clines were observed at microsatellite loci. The contact between the invasive outbreaks in North-West Italy and Central and South-Eastern Europe resulted in a zone of admixture, with particular characteristics. The evolutionary implications of the existence of a zone of admixture in Northern Italy and their possible impact on the invasion success of the western corn rootworm are discussed.

Citation: Bermond G, Ciosi M, Lombaert E, Blin A, Boriani M, et al. (2012) Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe. PLoS ONE 7(11): e50129. doi:10.1371/journal.pone.0050129

Editor: Norman Johnson, University of Massachusetts, United States of America

Received: June 8, 2012; **Accepted:** October 16, 2012; **Published:** November 26, 2012

Copyright: © 2012 Bermond 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.

Funding: This work was supported by grants from the Agence Nationale de la Recherche (ANR-06-BDIV-008-01 and ANR-09-BLAN-0145-01) and from the Agropolis Foundation (RTRA-Montpellier, BIOFIS project number 1001-001). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: gerald.bermond@sophia.inra.fr

Introduction

In the early stages of biological invasions, the genetic diversity of populations may be reduced by bottlenecks [1,2]. This may hinder adaptation to new environmental conditions. However, multiple introductions into the same area followed by admixture may increase genetic variability, offsetting the effect of genetic bottlenecks associated with each introduction [3,4]. If the various sources are genetically differentiated, this process results in the conversion of interpopulation genetic variation into intrapopulation genetic variation (as reported for the Cuban lizard in Florida, [5]). Conversely, intrapopulation variation may be converted into interpopulation variation if multiple introductions from a single source occur in geographically disconnected areas (e.g., [6]). Secondary contact and admixture between such independently introduced populations can eventually lead to restoration of the genetic variation found in the source population.

Admixture may thus have a positive impact on invasion, through the generation of novel genotypes [3], an increase in intrapopulation genetic variation [5] and heterosis [7–9], in which admixed individual fitness exceeds the fitness of the parental

populations. Finally, deleterious mutations may be purged from introduced populations during bottlenecks [10,11]. This may lead to an unusual phenomenon in which multiple introductions followed by admixture may result not only in the restoration of genetic variation, but also in the fitness of the population exceeding that of the source population.

The invasion of Europe by the western corn rootworm (WCR, *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae)) has involved multiple introductions into several disconnected geographic areas from a single source population in the northern USA [6,12]. The WCR is native to Central America [13,14], probably evolved with corn in Mexico and reached the South Western USA about 3000 years ago, together with its host plant, corn [15]. During the second half of the 20th century, the WCR has rapidly expanded its range across areas of corn production in the American Mid West (see [16] for a review), finally reaching the north-eastern coast of the USA in the late 1980s (see [17] for a review). In the late 20th century and the early 2000s, WCR was introduced into Europe repeatedly from the Northern USA [6,12]. It first observed near Belgrade Airport in Serbia, in 1992. An international monitoring

network has since monitored the distribution and of the WCR and its annual expansion in Europe [18]. There are two types of infested area: i) geographic areas of continuous WCR range expansion in Central and South-Eastern Europe (CSE Europe), North-West Italy (NW Italy) and Bavaria in Southern Germany, corresponding to “invasive” outbreaks and ii) many small disconnected outbreaks that have not persisted over time or expanded geographically, such as area of the North-East Italian (NE Italy) outbreak, which originated from CSE Europe [6,12,19]. During the multiple introductions of WCR in Europe from North America [12] strong bottlenecks lead to the formation of several genetically differentiated outbreaks [6]. For example Ciosi et al. [6] found a F_{ST} of 0.25 between CSE Europe and NW Italy. The largest outbreak, in CSE Europe, currently covers 16 countries extending from Austria to Ukraine and Southern Poland to Northern Greece. There has also been a westward expansion into eastern parts of Italy. In parallel, the NW Italian invasive outbreak has progressed eastwards. Since 2008, the NW and NE Italian and CSE European outbreak populations have been in close geographic proximity or even in contact, as in the Italian region of Veneto [19–21]. The current geographic distribution of WCR is now continuous in Northern Italy [21,22] (Figure 1). Because the various outbreaks were strongly genetically differentiated, there is probably currently an admixture Northern Italy between the NW Italian, NE Italian and CSE European invasive outbreaks. This situation provides us with an ideal opportunity to follow an evolutionary process in real time *in natura*. The aim of this study was to identify, document and characterize the putative admixture in Northern Italy, following contact between the WCR outbreak populations of NW and NE Italy and that of CSE Europe.

Materials and Methods

Sample Collection

No permission is required to collect samples of this species. WCR has no value and is a pest species whose populations are controlled by insecticide treatments or plant rotation wherever

they occur. Samples were collected on a west-east transect crossing the outbreak areas of NW Italy, NE Italy and CSE Europe (Figure 1 and Table 1). Adult beetles were caught with sticky traps at the Storo site (NW Italian outbreak) or by hand (net or funnel bound with a muslin bag) at all other sites. The WCR samples used in this study were collected during summer, from June to August, at 12 sites in Italy, Hungary and Serbia (see details in Table 1). At each of the study sites, all the individuals collected were obtained from a single corn field. We distinguished between two types of populations: the parental populations, corresponding to the invasive outbreaks of NW Italy, NE Italy and CSE Europe and the potentially admixed populations in the contact zone between the parental invasive outbreak populations.

Additional samples (Figure 1, Table 1) collected in 2010 in the outbreaks of NW Italy, NE Italy and CSE Europe and in 2003 in Northern USA (the native area) were only used for the temporal analysis (see Results section, *Temporal and spatial analyses*). For the Northern USA sample, DNA was extracted from individuals with the BioRad Aqua Pure isolation kit (BioRad, Hercules, CA) according to the manufacturer’s instructions, whereas, for other samples, DNA was extracted with the DNeasy tissue kit (Qiagen, Hilden, Germany), as explained above.

DNA Extraction and Microsatellite Analyses

All WCR samples were stored in 90–96% ethanol until DNA extraction. Template material for the polymerase chain reaction (PCR) amplification of microsatellite loci was obtained with two different protocols. For the NE Italian sample, the ‘salting out’ protocol of Sunnucks and Hales [23] was used for the rapid extraction of DNA from the head of each individual. For the other samples, we extracted DNA from the thorax or half the body, cut lengthwise, with the DNeasy® tissue kit (Qiagen, Hilden, Germany), according to the manufacturer’s instructions, with an elution volume of 100 μ l. Individuals were washed at least three times in 0.065% NaCl before extraction, to remove ethanol from the tissues.

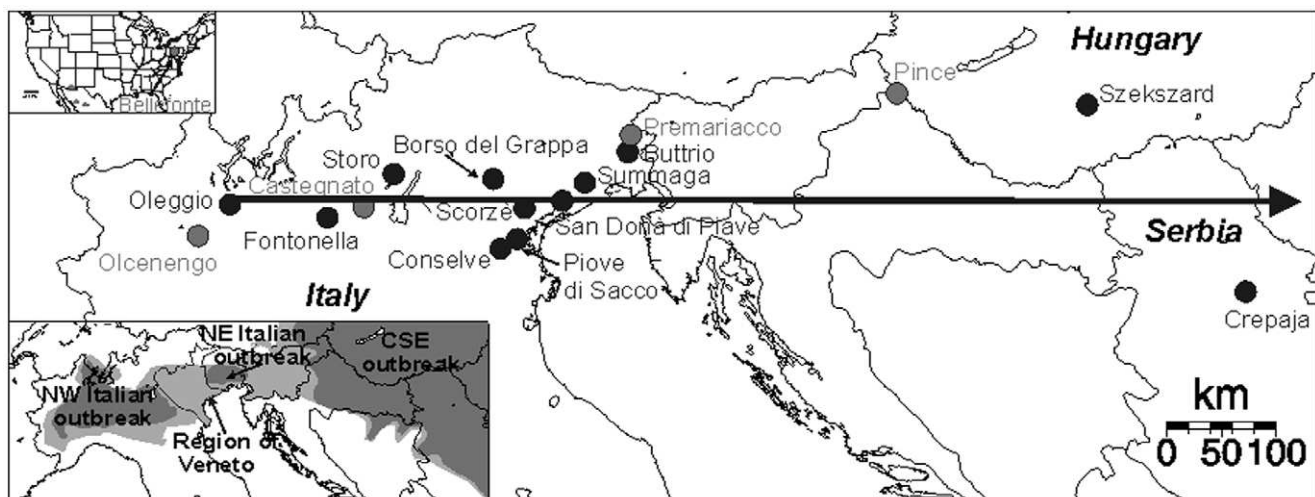


Figure 1. Location of the western corn rootworm (WCR) sample sites in Europe: in North Western (NW) Italy, in the contact zone (Veneto), in North Eastern (NE) Italy, in Central South Eastern (CSE) Europe and in the United States (Pennsylvania). The black circles represent the samples used in the spatial analysis and the black arrow indicates the geographic axis onto which these samples were projected (see Figures 3 and 4). The gray circles represent samples used in the temporal genetic differentiation analysis (collected in 2010 in Northern Italy and CSE Europe) and the sample representative of the native population (collected in 2003 in the United States). In the small map, the distribution areas of WCR before (2005) and after (2010) the secondary contact between the western and eastern outbreak populations are shown in dark gray and light gray, respectively. The dashed delimitation represents the surface of the administrative region of Veneto.
doi:10.1371/journal.pone.0050129.g001

Table 1. Description of the within-population genetic variation of western corn rootworm samples from Northern Italy and CSE Europe.

Population type	Sample	Outbreak	Origin (Region (Country))	Year of 1 st observation	N	Sampling date	Distance to Oleggio (km)	A		H _e	F _{IS}	p-HW
								DC	AR			
Parental	Oleggio	NW Italy	Piedmont (Italy)	2000	40	2007	0	3.77 (2.42)	2.70 (1.49)	0.41	0.04	0.40
	Olcenengo	NW Italy	Piedmont (Italy)	2000	30	2010	–	3.39 (2.66)	2.66 (1.60)	0.41	0.10	0.39
	Fontanella	NW Italy	Lombardia (Italy)	2000	33	2007	92	2.77 (1.79)	2.36 (1.32)	0.37	0.06	0.51
	Castegnato	NW Italy	Lombardia (Italy)	2000	30	2010	–	3.00 (1.92)	2.43 (1.32)	0.37	–0.06	0.56
	Storo	NW Italy	Trentino (Italy)	2000	90	2007	151	3.39 (1.85)	2.38 (1.24)	0.39	0.02	0.66
Potentially admixed	Borso del Grappa	–	Veneto (Italy)	2002	20	2009	240	3.15 (1.91)	2.85 (1.64)	0.41	0.04	0.66
	Conselve	–	Veneto (Italy)	2002	20	2009	247	3.00 (1.83)	2.60 (1.48)	0.39	–0.06	0.26
	Piove di Sacco	–	Veneto (Italy)	2002	11	2009	265	3.00 (2.04)	2.85 (1.79)	0.46	0.17	0.39
	Scorzè	–	Veneto (Italy)	2002	10	2009	278	3.15 (1.68)	3.08 (1.56)	0.53	0.00	1
	San Donà di Piave	–	Veneto (Italy)	2002	9	2009	305	3.39 (1.85)	3.37 (1.83)	0.54	0.00	0.33
	Summaga	–	Veneto (Italy)	2002	17	2009	327	3.23 (1.64)	2.94 (1.38)	0.48	0.04	0.46
Parental	Buttrio ^x	NE Italy	Friuli (Italy)	2003	27	2003	–	1.77 (0.73)	1.72 (0.64)	0.29	0.03	0.58
	Premariacco	NE Italy	Friuli (Italy)	2003	29	2010	–	2.70 (1.25)	2.22 (0.81)	0.34	0.08	0.15
	Pince	CSE Europe	Prekmurje (Slovenia)	2003	27	2010	–	2.77 (1.24)	2.63 (1.15)	0.44	0.08	0.06
	Szekszard	CSE Europe	Tolna (Hungary)	1992	39	2007	780	2.77 (1.24)	2.63 (1.11)	0.46	–0.05	0.98
	Crepaja	CSE Europe	Voivodine (Serbia)	1992	30	2007	940	2.92 (1.50)	2.69 (1.15)	0.46	0.00	0.59
Native	Bellefonte	Northern USA	Pennsylvania (USA)	1985	42	2003	–	6.10 (4.41)	4.23 (2.16)	0.61	0.05	0.20

Note: NW: North-West. NE: North-East. CSE: Central and South-Eastern. Buttrio^x: this sample was named Friuli in Ciosi et al. [6] and NE Italy in Miller et al. [12]. Distances to Oleggio are provided only for populations included in the spatial analysis. N: sample size. A: mean number of alleles per locus. A was determined by directcounts (DC) and allelic richness (AR) analysis. AR is based on the smallest sample size (N = 8 for several loci of the sample San Donà di Piave). Standard deviations between loci are shown in parentheses. H: mean expected heterozygosity [27]. p-HW: p-values for the exact test of deviation from HW equilibrium.
doi:10.1371/journal.pone.0050129.t001

We amplified 13 WCR microsatellite loci (including di and trinucleotides [6,24]) in three separate multiplex PCR performed in a PTC-225 MJ Research thermocycler. The first reaction amplified the DVV-D2, DVV-D4, DVV-D11, and DVV-T2 microsatellites, the second amplified DVV-D5, DVV-D8, DVV-D9 and DVV-ET1, and the third amplified DbA01, DbA05, DVV-T3, DVV-D12 and DbA07.

The thermal cycling conditions were the same for all three reactions and were as described by Miller *et al.* [25]. Forward primers were 5'-labeled with a fluorescent dye for detection of the PCR products on an Applied Biosystems 3130x1 Genetic Analyzer. Signal strength was rendered equivalent for different markers, by mixing labeled and unlabeled forward primers in the proportions (labeled:unlabeled) described by Miller *et al.* [25] for the first two sets of markers and in the following proportions for the third set: DbA01 2:1; DbA05 1:1; DVV-T3 1:0; DVV-D12 1:0 and DbA07 1:0. The primers used for DVV-D12 amplified were modified from those originally described by Kim *et al.* [26] (primers used here: F: 5'-GATTCTCAGTAATGGGAAACG-3'; R: 5'-CACACGCTTTCTCGTAATCTATC-3') This decreased the frequency of null allele detection at this locus to a negligible level. All three multiplex reactions were analyzed as described by Miller *et al.* [25]. All individuals were unambiguously assigned to a diploid multilocus genotype (two peaks per individual at the maximum).

Classical Statistical Genetics Analysis

Genetic variation within samples was evaluated by determining the mean number of alleles per locus (A) and mean expected heterozygosity (H_e) [27]. A and H_e were calculated with GENECLASS version 2.0.h [28]. We also calculated F_{IS} with GENEPOP ver. 4.0.1 [29,30]. We compared A values between population samples by estimating allelic richness (AR) based on the smallest sample size, by the rarefaction method [31] implemented in Fstat version 2.9.3 [32]. The various loci are the independent statistical units because they have their own coalescence story. Hence, the differences in AR and H_e between samples were assessed in a one-sided Wilcoxon signed rank test (with greater genetic variation in the supposed admixed populations in the contact zone between the parental invasive outbreaks as an alternative hypothesis) with locus as a repetition unit. The significance of the differences between the contact zone (Veneto) and the parental outbreak areas (NW Italy, NE Italy and CSE Europe) was then assessed by combining the probabilities obtained for each sample from Veneto by Fisher's method [33]. We also used the permutation procedure implemented in Fstat version 2.9.3 to test homogeneity of allelic richness and heterozygosity among samples [32]. We tested for deviation from Hardy-Weinberg equilibrium with the probability test approach implemented in GENEPOP version 4.0.1 [29,30]. Genetic variation between samples was assessed by calculating Weir and Cockerham's [34] pairwise F_{ST} and by assessing genic differentiation between pairs of samples [35] with GENEPOP version 4.0.1 [29,30]. When multiple tests were performed to test the same hypothesis, significance levels were lowered according to the Benjamini and Hochberg procedure [36].

Bayesian Analysis of the Population Structure of WCR

We inferred the genetic structure of the WCR in Northern Italy, using the Bayesian method implemented in STRUCTURE version 2.3.3 [37]. We performed 20 runs for each value of the number of clusters (K), defined as lying between 1 and the total number of sampling sites. Each run consisted of a burn-in of 5×10^4 iterations, followed by 10^5 iterations. We used the

admixture model of ancestry together with the correlated allele frequencies model [38], with and without the use of sampling location as prior information [39]. Default values were maintained for all other parameters. K was estimated with Evanno's ΔK statistic [40], which is based on the rate of change in log-likelihood between successive values of K and the variability of log-likelihood between different runs. The most likely run was then represented with DISTRICT version 1.0 ([41], see Figure 3).

Genetic differentiation was found between two samples collected from NE Italy in 2003 and 2010 (see details in Results, section *Temporal and spatial analyses* and Figure 1). We therefore removed the Buttrio (NE Italy) site from the Bayesian analysis and analyses of geographic population structure. However, as the NE Italian outbreak population may be a source of the Veneto populations, we considered the NE Italian sample from 2003 in the Approximate Bayesian Computation (ABC) analysis.

Geographic Analyses and Estimation of Admixture Rate

We tested the effect of geographic distance to Oleggio (the sampling site at the extreme west of the studied area, defined here as the geographic reference point) on microsatellite allelic frequencies. We defined a geographic axis passing through the various sampling sites and we then projected each sampled population orthogonally on this axis (Figure 1). The projected coordinates were then used to calculate the distance between each sample and the Oleggio sample. For each locus, the effect of this distance on allelic frequencies was assessed with a generalized linear model in SAS version 9.1.3 [42]. A multinomial distribution was chosen for the residual error and a cumulative logit function was used as the link. In cases of overdispersion an F test was carried out rather than a χ^2 test to evaluate the effect of distance [43].

The admixture rate was estimated (i) as described by Choisy *et al.* [44], with Oleggio and Crepaja as representative samples of the parental populations (these sites are the closest to the sites at which WCR was first observed in each parental outbreak), or (ii) with the coefficient of coancestry Q_c calculated with STRUCTURE version 2.3.3 [37,38].

ABC Analysis of Historical Scenarios of WCR Invasion in Northern Italy

For each sampling site within the contact zone (i.e. for each of the six Venetian target samples), we conducted six independent ABC analyses comparing various historical scenarios (Figure S1, Supporting information) differing in terms of the source of the population at the Venetian site concerned. Each parental outbreak population (CSE Europe, NE and NW Italy) was represented by the sample obtained from the site closest to the location at which the first observation for the outbreak concerned was reported (Crepaja, Buttrio and Oleggio for the CSE Europe, NE and NW Italy outbreaks, respectively). We thus considered a total of nine scenarios for each Venetian site, according to the origin of the population: (i) NW Italy outbreak (represented by Oleggio), (ii) NE Italy outbreak (represented by Buttrio), (iii) CSE Europe outbreak (represented by Crepaja), (iv, v, vi) all three possible scenarios of a single admixture between NW Italy, NE Italy and CSE Europe outbreak populations (vii, viii and ix) all three possible scenarios of a double admixture between NW Italy, NE Italy and CSE Europe outbreak populations. The history of the putative source populations was defined in accordance with published reports [6,12]. The NW Italy and CSE Europe populations were considered to have originated independently from North America, whereas the NE Italy population was considered to be derived from the CSE Europe population.

ABC analyses [45] were performed with DIYABC version 1.0.4.40 [46], with parameter values drawn from the prior distributions described hereafter and obtained by simulating 10^6 microsatellite datasets for each competing scenario.

Simulated and observed datasets were summarized with summary statistics, which were then used to calculate Euclidean distances between the simulated and observed datasets. We then estimated the posterior probabilities of the competing scenarios by polychotomous logistic regression [47] on the 1% of the simulated datasets closest to the observed dataset, after reducing the parameter space by a linear discriminant analysis (LDA) approach [48]. In cases of an overlap between the confidence intervals of the two largest posterior probabilities (each corresponding to a particular historical scenario), we repeated the ABC analyses with the two competing scenarios only.

We used the summary statistics describing genetic variation within and between populations generally used for approximate Bayesian computation analyses [47,49–51]. For each population and each population pair, we used the mean number of alleles per locus, mean heterozygosity, the mean ratio of the number of alleles to the range of allelic size, the F_{ST} between pairs of populations and mean individual assignment log-likelihoods of individuals from population i being assigned to population j and the maximum likelihood estimates for admixture proportions.

The prior distributions of the historical, demographic and mutational parameters used in the ABC analysis were as follows: $N_{s,i}$, the effective population size of the WCR American (USA) or European (CSE Europe, NW Italy, NE Italy and Veneto) source populations and $N_{g,G}$, the effective size of the ghost population (an unsampled population called the “ghost population” was included in the analysis for “double admixture” scenarios between the NE and NW Italy and CSE outbreaks; this “ghost population” is the result of a single admixture between two of the three putative source populations, depending on the scenario), were drawn from a uniform distribution bounded by 1000 and 20000 (Uniform[1000; 20000]); N_{Fi} and N_{FG} , the effective number of founders of the European populations, were drawn from a Uniform[1; 100] distribution; the bottleneck duration of population i , BD_i , was drawn from a Uniform[1;5] population. Each introduced population i ($i = 1, 2, 3, 4$ for Veneto (the contact zone in Northern Italy), NE Italy, NW Italy and CSE Europe, respectively) was founded by individuals originating from its source population t_i (and t_G for the ghost population) generations before the present (i.e. 2010, year of the study). As WCR is univoltine, t_i and t_G are also the number of years before the present. t_1 and t_G , t_2 , t_3 , and t_4 were drawn from between one and five generations (or years) before the date of the first observation (see Table 1) and were thus drawn from Uniform [9;13], [8;12], [11;15] and [19;23] distributions, respectively. In all double admixture scenarios, we fixed $t_1 < t_G$. We used a generalized stepwise mutation model (GSM [52]). A mean mutation rate across loci, μ , was first drawn from a Uniform [10^{-4} ; 5×10^{-3}] distribution, and single locus mutation rates, μ_k , were then drawn from gamma distributions with a mean of μ and a shape parameter of 2 (rate = $2/\mu$). For each locus, the coefficient P of the geometric distribution of repeat units by which a new mutant allele differs from its ancestor was drawn from an exponential distribution with a mean of 0.22.

We evaluated the ability of ABC to select the true scenario correctly, by analyzing test datasets simulated from known competing scenarios. For each scenario, one hundred such datasets were simulated with parameter values drawn from the same probability distributions as the priors. The posterior probabilities of each competing scenario were estimated for each simulated test dataset, with the same ABC procedure as described

above, and were used to calculate type I and II errors in the selection of scenarios. Type I error is the proportion of simulations in which the scenario considered is excluded but is actually the true one. Type II error is the proportion of simulations in which the scenario considered is selected but is not the true one. Small Type 2 errors provide good confidence in the results even if the Type 1 errors are large.

Results

Temporal and Spatial Analyses

The main body of the present study concerns spatial analyses of WCR genetic variation. When the samples were not collected within the same year as in our case, a procedure must guarantee that no temporal confounding effect occurs. Such a temporal effect was hence tested using samples collected within the same area at various dates (Lombardy, Piedmont, Friuli and CSE Europe). Locations showing a significant temporal effect were removed from the spatial analysis.

No temporal differentiation was found between samples collected in the NW Italy outbreak in Piedmont (Oleggio, 2007, and Olcenengo, 2010), in Lombardia (Fontanella, 2007, and Castegnato, 2010) and in the CSE Europe outbreak (Hungary (Tolna) and Serbia (Voivodine) in 2007 and Slovenia (Prekmurje) 2010; $p > 0.08$ for all pairwise comparisons considered). These locations were thus kept and the 2007 samples were arbitrarily chosen for subsequent spatial analyses. A significant temporal differentiation was found in Friuli (Buttrio, 2003, and Premariaco, 2010; $p = 0.001$). Friuli was thus not considered for the spatial analyses.

The Friuli sample of Buttrio was sampled in 2003, i.e. before the contact between the NW, NE Italian and CSE European outbreaks in 2008. Hence, it properly describes the north eastern Italian outbreak before any contact between the various outbreaks. Buttrio was therefore used in the ABC historical scenario comparison.

Genetic Variation within and between Populations

Overall, the European WCR populations displayed moderate polymorphism, with 5.31 (SD = 3.71) alleles per locus over all samples. Within samples, the mean number of alleles was between 1.77 (SD = 0.73) for Buttrio in NE Italy and 3.77 (SD = 2.43) in Oleggio in NW Italy (Table 1). Mean expected heterozygosity (H_e) was low to moderate and varied from 0.29 for Buttrio in NE Italy to 0.54 for San Donà di Piave in Veneto (Table 1). F_{IS} estimates were low and no significant deviation from Hardy-Weinberg equilibrium was observed (Table 1). Samples from Veneto contained a significantly larger number of alleles than parental samples (Fisher’s method for the combination of probabilities, $\chi^2 = 121.28$; df = 36; $p < 10^{-5}$; $\chi^2 = 66.31$; df = 12; $p < 10^{-5}$ and $\chi^2 = 61.24$; df = 24; $p < 10^{-5}$ for the comparisons “Veneto/NW Italy, Veneto/NE Italy and Veneto/CSE Europe”, respectively; Figure 2) and heterozygosity was significantly greater for samples from Veneto than for all other samples (Fisher’s method for the combination of probabilities, $\chi^2 = 110.97$; df = 36; $p < 10^{-5}$; $\chi^2 = 60.22$; df = 12; $p < 10^{-5}$ and $\chi^2 = 40.81$; df = 24; $p < 0.05$, for the comparisons “Veneto/NW Italy, Veneto/NE Italy and Veneto/CSE Europe”, respectively; Figure 2). Permutation tests performed with Fstat revealed a significantly larger allelic richness (one sided-test, $p = 4 \times 10^{-3}$) and a marginally significantly larger heterozygosity (one sided-test, $p = 0.062$) in Veneto compared to parental samples. Most pairwise comparisons of samples (92%) showed significant genetic differentiation (Table 2). Veneto samples generally displayed significant genetic differentiation from

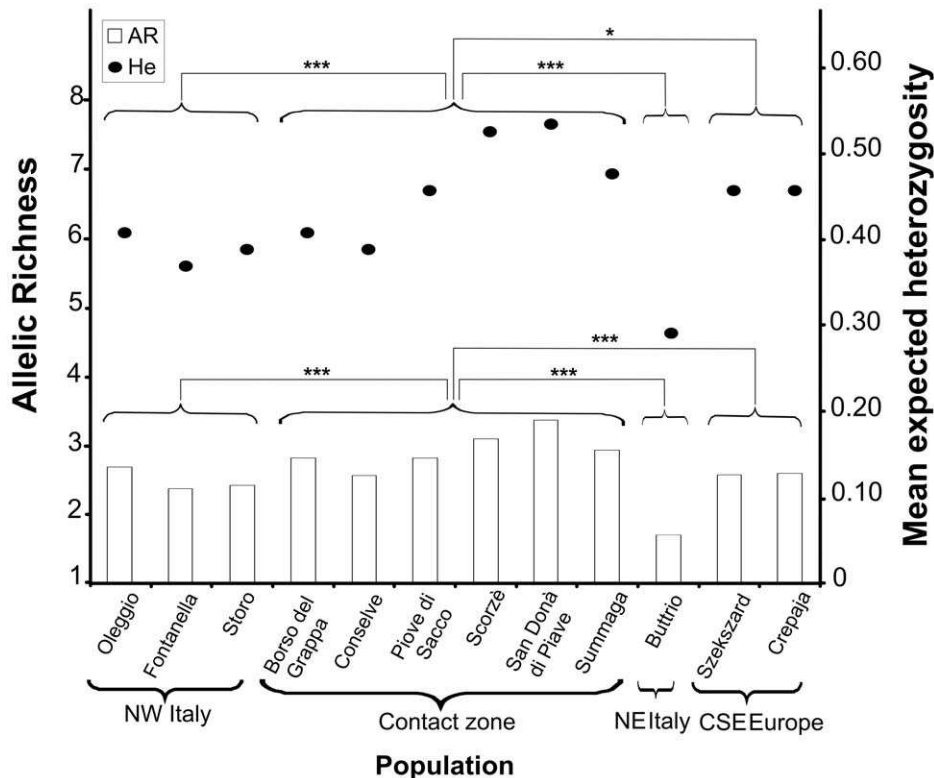


Figure 2. Comparison of genetic variation between potentially admixed western corn rootworm samples (Veneto) and their putative parental (NE and NW Italy and CSE Europe) populations. Gray diamonds indicate the mean expected heterozygosity (He) for the various samples on the right axis, whereas white bars indicate allelic richness (AR) on the left axis. Significant differences in mean expected heterozygosity and allelic richness between samples are indicated by asterisks (based Fisher's method for combining probabilities [33]). doi:10.1371/journal.pone.0050129.g002

i) the samples from the three putative parental outbreak populations and ii) from each other (Table S1).

Bayesian Analysis of the Structure of WCR Samples in Italy and CSE Europe

The number of genetic clusters (k) was estimated at $K=2$, whatever the model used. The analysis of coancestry coefficients (Q) (for the model with the use of sampling location as prior information) (Figure 3) indicated that (i) the vast majority of individuals from the west of the transect belonged to a single cluster (with $Q > 0.86$ for 161 of 163 individuals) (ii) individuals from the east belonged to another cluster (with $Q < 0.07$ for all individuals) and that (iii) a large number of individuals from Scorzè, San Dona di Piave, Borso del Grappa and Piove di Sacco are admixed between both clusters (with Q between 0.20 and 0.80). One individual from Summaga, a population belonging to the eastern genetic cluster, was assigned to the western genetic cluster ($Q = 0.97$).

Detection of Clines of Allelic Frequencies among WCR Samples in Italy and CSE Europe

A significant effect of geographic distance on allelic frequencies was found for six of the 13 loci studied ($p < 0.0033$ for each of these 6 loci; see Figure 4 for example). Distance had a significant effect over all loci (Fisher's method for the combination of probabilities, $\chi^2 = 113.62$, $df = 26$, $p < 10^{-5}$). The admixture rates obtained by the two different methods (the method of Choisy *et al.* [44] and the coancestry coefficient from STRUCTURE) were very similar and closely matched a sigmoid logit model ($R^2 = 0.90$ for both

methods). Figure 4 shows that the center of the cline is located about 310 km from Oleggio, in Veneto, and that this cline is about 100 km wide.

The effect of the geographic distance on the $F_{ST}/(1-F_{ST})$ between each sample and Oleggio, the sample located in the extreme west of the studied area was significant (Spearman rank-order correlation test, $p < 10^{-5}$, Figure 4). Genetic differentiations compared to Oleggio sample are very low for western samples (NW Italy outbreak) and substantial for eastern samples, with a F_{ST} of about 0.25 in CSE Europe outbreak. A sharp discontinuity of $F_{ST}/(1-F_{ST})$ is observed in the contact zone, i.e. in Veneto, where values are intermediate.

ABC Analysis of Historical Scenarios of WCR Invasion in Northern Italy

The ABC procedure was used to calculate the posterior probabilities of the evolutionary scenarios describing the origin of each population sampled in Veneto. According to the highest probabilities with non overlapping confidence intervals, four of the six Veneto samples probably originate from admixture events: samples from Scorzè and San Dona di Piave probably originate from admixture between the outbreak populations of NW Italy and CSE Europe (Table 3); samples from Borso del Grappa and Summaga probably result from double admixture between NW Italy, NE Italy and CSE Europe outbreak populations and samples from Conselve and Piove di Sacco have a simple origin in NW Italy.

Some of the largest posterior probabilities were only moderate (Table 3) and some type I errors were large (four

Table 2. Pairwise F_{ST} estimates [34] between Northern Italian and Central and South-Eastern European samples of the western corn rootworm.

Population	North-West Italy		Contact zone							North-East Italy	Central and South-Eastern Europe
	Oleggio	Fontanella	Storo	Borso del Grappa	Conselve	Piove di Sacco	Scorzè	San Donà di Piave	Summaga	Buttrio	Szekszard
Fontanella	0.00										
Storo	0.01	0.01									
Borso del Grappa	0.03	0.02	0.02								
Conselve	0.03	0.02	0.01	0.04							
Piove di Sacco	0.02	0.02	0.01	0.01	0.03						
Scorzè	0.11	0.12	0.12	0.06	0.11	0.06					
San Donà di Piave	0.08	0.10	0.10	0.06	0.09	0.06	0.03				
Summaga	0.24	0.27	0.28	0.20	0.29	0.20	0.08	0.08			
Buttrio	0.38	0.42	0.40	0.36	0.46	0.38	0.29	0.27	0.14		
Szekszard	0.25	0.27	0.28	0.19	0.29	0.20	0.08	0.11	0.02	0.12	
Crepaja	0.25	0.26	0.27	0.17	0.28	0.19	0.07	0.10	0.02	0.15	0.00

Note: Significant pairwise differentiation tests after correcting for the false-positive rate by the procedure of Benjamini and Hochberg [36] are shown in bold typeface. From left to right, samples are ordered from west to east, from Italy to Serbia.

doi:10.1371/journal.pone.0050129.t002

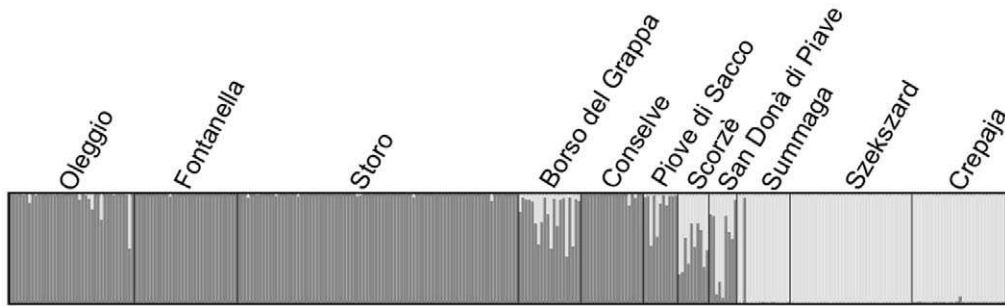


Figure 3. Population structure, based on 13 microsatellite loci from WCR, as estimated by genotypic clustering in STRUCTURE. Assignment of 319 individuals to $K=2$ genetically distinct clusters. Individuals are represented by a vertical bar, grouped by sampling location (the name of which is above the plot). The proportion of dark gray in each bar indicates the individual coefficient of coancestry (or admixture rate). doi:10.1371/journal.pone.0050129.g003

of the six values are between 0.43 and 0.67). However, as explained in the material and method, low type 2 errors (<0.09 , Table 3) suggest that we can have a high degree of confidence in the choice of scenario.

Discussion

Since the first observation of the CSE European outbreak of WCR near Belgrade in 1992, a monitoring network has followed the progression of this pest in Europe [17,18,22]. No contact between the NW Italian, the NE Italian and CSE European outbreak populations was detected until 2008, since

when the distribution of WCR in Northern Italy has been continuous, extending from NW Italy to CSE Europe [20,21]. Population genetic analysis of the WCR collected in the zone of secondary contact between the NW Italian and CSE outbreak populations revealed a high degree of genetic heterogeneity, determined principally by geography. Our results confirm show the existence of a zone of admixture in Northern Italy, with the occurrence of admixture between highly differentiated populations.

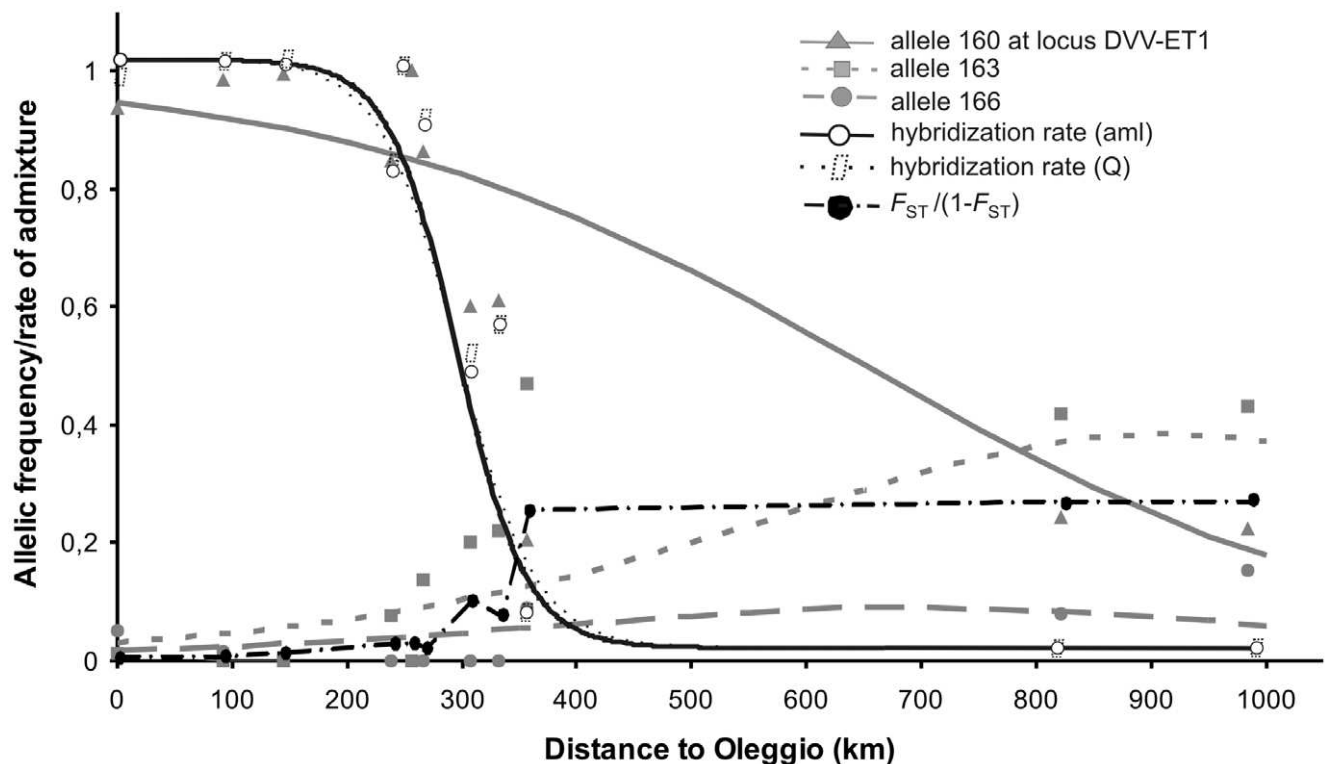


Figure 4. Observed and fitted values of allelic frequencies at microsatellite locus DVV-ET1, of rates of admixture and index of genetic differentiation ($F_{ST}/(1-F_{ST})$) relative to Oleggio, as a function of geographic distance to the Oleggio site in North-West Italy. Observed values are indicated by symbols whereas fitted values are indicated by lines. The fit for each allele corresponds to the generalized linear model described in the materials and methods. Three of the four alleles present at this locus are represented because the frequency of the 4th allele is simply the complement of those of the three others. Similar pattern were observed for 5 other loci. A sigmoid function was fitted to the admixture rates. For ($F_{ST}/(1-F_{ST})$), only observed values are shown. doi:10.1371/journal.pone.0050129.g004

Table 3. ABC analysis of the nine historical scenarios describing the history of each population sampled in the secondary contact zone (Veneto).

Target sample in the secondary contact zone	Historical scenarios									Errors	
	Without admixture			Single admixture			Double admixture			Type I	Type II (min-max)
	NW Italy	NE Italy	CSE Europe	NW Italy/NE Italy	NE Italy/CSE Europe	NW Italy/CSE Europe	(NW Italy/NE Italy)/CSE Europe	(CSE Europe/NE Italy)/NW Italy	(CSE Europe/NW Italy)/NE Italy		
Borso del Grappa	0.19 [0.18;0.21]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.17 [0.16;0.18]	0.0 [0.0;0.0]	0.31 [0.30;0.33] (0.37) [0.36;0.38]	0.02 [0.02;0.02]	0.29 [0.28;0.31] (0.63) [0.62;0.64]	0.01 [0.01;0.01]	0.66	0.08 (0.0–0.23)
Conselve	0.99 [0.98;0.99]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.01 [0.0;0.01]	0.0 [0.0;0.0]	0.10	0.04 (0.0–0.12)
Piove di Sacco	0.27 [0.21;0.32] (0.56) [0.55;0.58]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.18 [0.14;0.21]	0.0 [0.0;0.0]	0.22 [0.18;0.26]	0.0 [0.0;0.0]	0.33 [0.28;0.37] (0.44) [0.42;0.45]	0.0 [0.0;0.0]	0.06	0.05 (0.0–0.14)
Scorzè	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.07 [0.06;0.09]	0.0 [0.0;0.0]	0.54 [0.49;0.59]	0.10 [0.08;0.12]	0.26 [0.22;0.30]	0.03 [0.02;0.03]	0.58	0.09 (0.0–0.23)
San Donà di Piave	0.01 [0.0;0.01]	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.17 [0.14;0.20]	0.0 [0.0;0.0]	0.41 [0.37;0.46]	0.07 [0.06;0.09]	0.31 [0.26;0.34]	0.03 [0.02;0.04]	0.43	0.07 (0.02–0.30)
Summaga	0.0 [0.0;0.0]	0.0 [0.0;0.0]	0.03 [0.02;0.04]	0.02 [0.01;0.02]	0.01 [0.0;0.01]	0.13 [0.10;0.15]	0.61 [0.57;0.66]	0.11 [0.09;0.14]	0.09 [0.07;0.11]	0.67	0.06 (0.01–0.15)

Posterior probabilities are given, with their confidence intervals and associated type I and type II errors.

Note: 95% confidence intervals are indicated in brackets. Values in parentheses are the new posterior probabilities of scenarios re-analyzed in ABC due to the overlap between confidence intervals. The highest probability values are shown in bold typeface and indicate the best scenario. NW: North-West. NE: North-East.

CSE: Central and South-Eastern.

doi:10.1371/journal.pone.0050129.t003

A Contact and Zone of Admixture in Northern Italy, in the Veneto Region

The current contact zone between the different European outbreak populations (NW Italy, NE Italy and CSE Europe) is located in Veneto. This zone has been intensively monitored with pheromone traps since the first captures of WCR in the Venetian region in 1998, leading to the implementation of an eradication program. Since 2008, the geographic distribution of WCR has been continuous in Northern Italy [21,22] due to contact between the outbreak populations of NW and NE Italy and CSE Europe [19–21].

Our analysis of the population genetic structure of WCR in Northern Italy and CSE Europe showed that the WCR populations sampled in Veneto resulted mostly from contact and admixture between the NW Italian and CSE European outbreak populations. The admixture analysis, the Bayesian analysis of population genetic structure and the ABC analysis of scenario choice all indicate that the Veneto region is a zone of admixture, containing admixed individuals. The genomes of individuals sampled in Veneto could be attributed to both the genetically different clusters of NW Italy and CSE Europe, with various levels of admixture. Moreover, the rate of admixture between the NW Italian and CSE European outbreaks varied evenly across a gradient from west to east.

According to the ABC analysis, most samples from the western part of Veneto have a simple NW Italian origin, whereas all the other sample populations result from single or double admixture events between the three parental outbreak populations (NE and NW Italy and CSE Europe). Double admixture between individuals originating from distant populations (here, the three parental populations) accounted for the origin of the most westerly sample from Veneto (Borso del Grappa). This implies that individuals from the eastern outbreak populations (the NE Italy and CSE Europe outbreaks) migrated to the western part of Veneto, suggesting that long-distance dispersal occurs. Long-distance dispersal is also suggested by the detection of an individual genetically assigned to the western genetic cluster in Summaga, a population from the eastern genetic cluster. Recent studies have suggested that long-distance dispersal is common in WCR and that invasive populations of WCR are expanding through stratified dispersal [17,53].

The occurrence of a cline in the Veneto region in 2009, even with such a large width (about 100 km), is consistent with the short period of time between secondary contact (2008) and sampling (2009). The WCR has a considerable capacity for dispersal, as shown by the rapid rate of expansion of the CSE Europe outbreak population (60–100 km (60–100 km per year [54])). A long period between sampling and contact would therefore probably have led to the homogenization of microsatellite frequencies over space [55].

A thorough clinal analysis of WCR in Veneto is required, to estimate dispersal parameters (e.g., [56]). Under the effect of dispersal alone, allelic frequencies tend to become homogeneous over the zone of admixture, so the slope of the cline depends on the strength of dispersal and contact time (e.g., [55]). The speed at which the slope decreases over time is thus a direct function of the dispersal intensity [57]. Temporal analysis of the North Italian zone of admixture should therefore provide an estimate of the dispersal capacities of WCR, as recently reported for *Biston Betularia* by Saccheri *et al.* [58].

It is noteworthy that biological invasions may provide numerous opportunities to estimate dispersal parameters using such analysis of frequency cline. It is now admitted that multiple invasions are frequent [59,60]. This can lead to situations, as in WCR, in which

recently introduced populations display large neutral genetic differentiation and will eventually merge during their geographical expansion. This evolutionary scenario, that allows estimation of dispersal through clinal analysis, probably occurred in the case of the green crab *Carcinus maenas* in the north Eastern American coasts [60,61]. It also probably occurred in the plant pathogenic fungus *Mycosphaerella fijiensis* in Africa [62], and is currently occurring in the Asian ladybeetle *Harmonia axyridis* in North America [51,63]. Even though biological situations allowing such dispersal estimation are abundant, the literature provides no example of such studies to our knowledge. As discussed by Ciosi *et al.* [6] this lack may result from a simple difficulty: The rapid spatial spreading, the late observation and the late sampling of the invasive populations likely result in the observation of a single genetically homogenized population, with no observable frequency clines.

In cases of selection acting against hybrids, the clines may remain stable over time. The sampling of WCR in Northern Italy in consecutive years, followed by a temporal analysis of cline shape (slope and width), may thus provide information about the balance between dispersal and selection [64–66] in this pest species.

Evolutionary Implications of the Existence of a WCR Admixture Zone in Northern Italy

We found that genetic variation was greater in the zone of admixture than in the parental outbreak areas. Moreover, the genetic variation within the zone of admixture is approaching that of the Northern USA source population and thus displays substantial restoration of the genetic variation lost during the introduction and establishment of the invasive outbreaks. Indeed, the mean allelic richness and the mean heterozygosity of the zone of admixture represent 70% and 77% of that of the Northern USA source population, respectively, whereas the NW Italian, the NE Italian and CSE Europe outbreaks display a mean of 59, 41 and 63%, of the allelic richness and 64, 47 and 76% of the heterozygosity found in the Northern USA, respectively (see more details in the Figure S2). This study and previous population genetics studies of European populations of WCR provide information about the way in which this increase occurred, because it has been possible to observe (i) a loss of genetic variation during the founding of most of independently introduced populations [6,12], and (ii) admixture between the outbreak populations leading to a partial restoration of genetic variation (this study). An increase in neutral genetic variation was observed, but our results provide no information about differences in the phenotypic variability and fitness of WCR from the zone of admixture and from the parental areas. Recent invasion studies have reported example of a positive effect of admixture on the invasive capacity of animals such as the Asian ladybird *Harmonia axyridis* [67] and the freshwater snail *Melanoides tuberculata* [8], and of plants [3], such as *Silene vulgaris* [68]. In WCR, further quantitative genetics studies of life history traits (such as fertility, longevity and dispersal) are required to determine whether admixture has been and is currently an advantage for the invasion of Europe by this species.

Supporting Information

Figure S1 Graphical representation of the competing scenarios used for the ABC analyses on our European dataset. Scenarios (i), (ii) and (iii) represent the “simple origin” scenarios, scenarios (iv), (v) and (vi) are “single admixture” scenarios and (vii), (viii) and (ix) are “double admixture” scenarios. Historical and demographic parameters were identical for all

introduction models. Time 0 is the present and represents the year of the study (2010). The Veneto population was founded t_1 generations before the present, had an effective number of founders NF_{Veneto} with the population remaining at this size for BD_1 generations (bottleneck duration) and then reached a larger stable effective population size NS_{Veneto} . The putative source populations, the Central South-Eastern European (CSE Europe) and North-West (NW) Italian outbreak populations, diverged from the USA population t_4 and t_3 generations ago with an effective number of founders NF_{CSE} and $NF_{NWItaly}$, bottleneck durations BD_4 and BD_3 and stable effective population sizes NS_{CSE} and $NS_{NWItaly}$, respectively. The North-East (NE) Italian population was founded t_2 generations ago from the CSE European population, with an effective number of founders $NF_{NEItaly}$, a bottleneck duration BD_2 and an effective population size $NS_{NEItaly}$. When admixture occurs, the admixture rates ar and $1-ar$ are the genetic contribution of each of the source populations to the origin of the Veneto population. An unsampled population called the “Ghost population” was included into the analysis to allow “double admixture” scenarios (scenario (vii), (viii) and (ix)) between NE and NW Italy and CSE outbreak populations. This “Ghost population” is the result of a single admixture between two of the three putative source populations (the populations involved depend on the scenario), t_G generations ago ($t_G > t_1$ and $t_G < t_2$), with an effective number of founders $NF_{GhostPopulation}$, a bottleneck duration BD_G and an effective population size $NS_{GhostPopulation}$. The rates of admixture corresponding to the “Ghost population” are ar_G and $1-ar_G$. For all models, populations were assumed to be isolated from each other, with no exchange of migrants.

References

- Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, et al. (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics* 32: 305–332.
- Wares JP, Hughes AR, Grosberg RK (2005) Mechanisms that Drive Evolutionary Change: insights from species introductions and invasions. In: Sax DF, Stachowicz JJ, Gaines SD, editors. *Species Invasions: Insights into Ecology, Evolution and Biogeography*. Sunderland, MA USA: Sinauer Associates Inc. 229–257.
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* 97: 7043–7050.
- 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. *Current Biology* 18: 363–367.
- Kolbe JJ, Glor RE, Schettino LRG, Lara AC, Larson A, et al. (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* 431: 177–181.
- Ciosi M, Miller NJ, Kim KS, Giordano R, Estoup A, et al. (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology* 17: 3614–3627.
- Cheres MT, Miller JF, Crane JR, Knapp SJ (2000) Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. *Theoretical and Applied Genetics* 100: 889–894.
- Facon B, Jarne P, Pointier JP, David P (2005) Hybridization and invasiveness in the freshwater snail *Melanooides tuberculata*: hybrid vigour is more important than increase in genetic variance. *Journal of Evolutionary Biology* 18: 524–535.
- Reif JC, Melchinger AE, Xia XC, Warburton ML, Hoisington DA, et al. (2003) Genetic distance based on simple sequence repeats and heterosis in tropical maize populations. *Crop Science* 43: 1275–1282.
- Facon B, Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, et al. (2011) Inbreeding Depression Is Purged in the Invasive Insect *Harmonia axyridis*. *Current Biology* 21: 424–427.
- Pujol B, Zhou SR, Vilas JS, Pannell JR (2009) Reduced inbreeding depression after species range expansion. *Proceedings of the National Academy of Sciences of the United States of America* 106: 15379–15383.
- Miller N, Estoup A, Toepfer S, Bourguet D, Lapchin L, et al. (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* 310: 992.
- Branson TF, Krysan JL (1981) Feeding and oviposition behavior and life cycle strategies of *Diabrotica*: an evolutionary view with implications for pest management. *Environmental Entomology* 10: 826–831.

(TIF)

Figure S2 Percentage of the mean expected heterozygosity (He) and allelic richness (AR) of the Northern USA source population (Bellefonte, Pennsylvania) for each sample from the three putative parental outbreak areas (NW and NE Italy and CSE Europe) and for samples from Veneto. Dark gray bars indicate the mean expected heterozygosity (He) of the various samples, whereas light gray bars correspond to their allelic richness (AR).

(TIF)

Table S1 Mean pairwise F_{ST} comparisons between Northern Italian (NW Italy, Veneto and NE Italy) and Central and South-Eastern European (CSE Europe) samples of western corn rootworm (WCR).

(DOC)

Acknowledgments

We thank Isabelle Legoff and Alexandra Auguste-Maros for technical assistance.

Author Contributions

Conceived and designed the experiments: GB MC TG MB LF ST. Performed the experiments: GB AB. Analyzed the data: GB MC EL TG. Wrote the paper: GB MC EL AB MB LF ST TG. Performed the sampling: MB LF ST.

- Smith RF (1966) Distributional patterns of selected western North American insects: the distribution of diabroticites in western North America. *Bulletin of Entomological Society of America* 12: 108–110.
- Krysan JL, Smith RF (1987) Systematics of the virgifera species group of *Diabrotica* (Coleoptera: Chrysomelidae: Galerucinae). *Entomographia* 5: 375–484.
- Chiang HC (1973) Bionomics of the northern and western corn rootworms. *Annual Review of Entomology* 18: 47–72.
- Gray ME, Sappington TW, Miller NJ, Moeser J, Bohn MO (2009) Adaptation and Invasiveness of Western Corn Rootworm: Intensifying Research on a Worsening Pest. *Annual Review of Entomology* 54: 303–321.
- Kiss J, Edwards CR, Berger HK, Cate P, Cean M, et al. (2005) Monitoring of western corn rootworm (*Diabrotica virgifera virgifera* LeConte) in Europe 1992–2003. *Western Corn Rootworm: Ecology and Management*: 29–39.
- De Luigi V, Furlan L, Palmieri S, Vettorazzo M, Zanini G, et al. (2011) Results of WCR monitoring plans and evaluation of an eradication programme using GIS and Indicator Kriging. *Journal of Applied Entomology* 135: 38–46.
- Edwards CR, Kiss J (2008) New WCR 2007 General Spread Map for Europe. <http://extensionentm.purdue.edu/wcr/>.
- Edwards CR, Kiss J (2009) New WCR 2008 General Spread Map for Europe. *IWGO Newsletter* 29.
- Edwards CR, Kiss J (2010) New WCR 2009 General Spread Map for Europe. *IWGO Newsletter* 30.
- Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution* 13: 510–524.
- Kim KS, Stolz U, Miller NJ, Waits ER, Guillemaud T, et al. (2008) A core set of microsatellite markers for Western corn rootworm (Coleoptera: Chrysomelidae) population genetics studies. *Environ Entomol* 37: 293–300.
- Miller NJ, Ciosi M, Sappington TW, Ratcliffe ST, Spencer JL, et al. (2007) Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation tolerance. *Journal of Applied Entomology* 131: 378–385.
- Kim KS, Sappington TW (2005) Polymorphic microsatellite loci from the western corn rootworm (Insecta: Coleoptera: Chrysomelidae) and cross-amplification with other *Diabrotica* spp. *Molecular Ecology Notes* 5: 115–117.
- Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, et al. (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95: 536–539.
- Raymond M, Rousset F (1995) Genepop (version 1.2), a population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.

30. Rousset F (2008) GENEPOP' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.
31. Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* 12: 844–855.
32. Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995).
33. Sokal RR, Rolf FJ (1995) *Biometry. The Principles and Practice of Statistics in Biological Research*. New York: W.H. Freeman and Company.
34. Weir BS, Cockerham C (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
35. Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* 49: 1280–1283.
36. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 57: 289–300.
37. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
38. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
39. Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332.
40. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.
41. Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.
42. SAS Institute Inc. (2008) SAS. 9.1.3 ed. Cary, NS, USA. pp. Statistical Analysis.
43. Crawley M (1993) *GLIM for Ecologists*: Blackwell Science.
44. Choisy M, Franck P, Cornuet JM (2004) Estimating admixture proportions with microsatellites: comparison of methods based on simulated data. *Molecular Ecology* 13: 955–968.
45. Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162: 2025–2035.
46. Cornuet JM, Ravigne 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.
47. Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, et al. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24: 2713–2719.
48. Estoup A, Lombaert E, Marin JM, Guillemaud T, Pudlo P, et al. (2012) Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources* 12: 846–855.
49. Csillery K, Blum MGB, Gaggiotti OE, Francois O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* 25: 410–418.
50. 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.
51. Lombaert E, Guillemaud T, Cornuet JM, Malusa T, Facon B, et al. (2010) Bridgehead Effect in the Worldwide Invasion of the Biocontrol Harlequin Ladybird. *PLoS ONE* 5.
52. Estoup A, Jarne P, Cornuet JM (2002) Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11: 1591–1604.
53. Ciosi M, Miller NJ, Toepfer S, Estoup A, Guillemaud T (2011) Stratified dispersal and increasing genetic variation during the invasion of Central Europe by the western corn rootworm, *Diabrotica virgifera virgifera*. *Evolutionary Applications* 4: 54–70.
54. Baufeld P, Enzian S (2001) Simulations model for spreading scenarios of western corn rootworm (*Diabrotica virgifera virgifera*) in case of Germany. *IWGO Newsletter* 22: 14–15.
55. Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Monographs in population biology. New Jersey: Princeton University Press. 30–96.
56. Barton NH, Gale KS (1993) Genetic analysis of hybrid zone. In: Harrison RG, editor. *Hybrid zones and the evolutionary process*. New York: Oxford University Press. 13–45.
57. Gay L, Crochet PA, Bell DA, Lenormand T (2008) Comparing Clines on Molecular and Phenotypic Traits in Hybrid Zones: a Window on Tension Zone Models. *Evolution* 62: 2789–2806.
58. Saccheri IJ, Rousset F, Watts PC, Brakefield PM, Cook LM (2008) Selection and gene flow on a diminishing cline of melanic peppered moths. *Proceedings of the National Academy of Sciences of the United States of America* 105: 16212–16217.
59. Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, et al. (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* 144: 1–11.
60. Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* 22: 454–464.
61. Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society B-Biological Sciences* 273: 2453–2459.
62. Rieux A, Halkett F, de Bellaire LD, Zapater MF, Rousset F, et al. (2011) Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. *Molecular Ecology* 20: 1661–1674.
63. Lombaert E, Guillemaud T, Thomas CE, Handley LJJ, Li J, et al. (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*. *Molecular Ecology* 20: 4654–4670.
64. Barton NH, Hewitt GM (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics* 16: 113–148.
65. Mallet J, Barton N (1989) Inference From Clines Stabilized by Frequency-Dependent Selection. *Genetics* 122: 967–976.
66. Slatkin M (1973) Gene flow and selection in a cline. *Genetics* 75: 733–756.
67. Turgeon J, Tayeh A, Facon B, Lombaert E, De Clercq P, et al. (2011) Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *Journal of Evolutionary Biology* 24: 1044–1052.
68. Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology* 23: 1720–1731.

Supporting information

Table S1: Mean pairwise F_{ST} comparisons between Northern Italian (NW Italy, Veneto and NE Italy) and Central and South-Eastern European (CSE Europe) samples of western corn rootworm (WCR).

Population	NW Italy	Veneto	NE Italy	CSE Europe
NW Italy	0.01			
Veneto	0.09	0.09		
NE Italy	0.40	0.32	-	
CSE Eur	0.26	0.14	0.14	0

Note: The population of NW Italy is represented by 3 samples (Oleggio, Storo and Fontanella), the population of Veneto is represented by 6 samples (Borso del Grappa, Conselve, Piove di Sacco, Scorze, San Donà di Piave and Summaga), the population of NE Italy is represented by 1 sample (Buttrio) and the population of CSE Europe is represented by 2 samples (Szekszard and Crepaja).

Figure legends

Figure S1: Graphical representation of the competing scenarios used for the ABC analyses on our European dataset. Scenarios (i), (ii) and (iii) represent the “simple origin” scenarios, scenarios (iv), (v) and (vi) are “single admixture” scenarios and (vii), (viii) and (ix) are “double admixture” scenarios. Historical and demographic parameters were identical for all introduction models. Time 0 is the present and represents the year of the study (2010). The Veneto population was founded t_1 generations before the present, had an effective number of founders NF_{Veneto} with the population remaining at this size for BD_1 generations (bottleneck duration) and then reached a larger stable effective population size NS_{Veneto} . The putative source populations, the Central South-Eastern European (CSE Europe) and North-West (NW) Italian outbreak populations, diverged from the USA population t_4 and t_3 generations ago with an effective numbers of founders NF_{CSE} and $NF_{NWItaly}$, bottleneck durations BD_4 and BD_3 and stable effective population sizes NS_{CSE} and $NS_{NW Italy}$, respectively. The North-East (NE) Italian population was founded t_2 generations ago from the CSE European population, with an effective number of founders $NF_{NEItaly}$, a bottleneck duration BD_2 and an effective population size $NS_{NEItaly}$. When admixture occurs, the admixture rates ar and $1-ar$ are the genetic contribution of each of the source populations to the origin of the Veneto population. An unsampled population called the “Ghost population” was included into the analysis to allow “double admixture” scenarios (scenario (vii), (viii) and (ix)) between NE and NW Italy and CSE outbreak populations. This “Ghost population” is the result of a single admixture between two of the three putative source populations (the populations involved depend on the scenario), t_G generations ago ($t_G > t_1$ and $t_G < t_2$), with an effective number of founders $NF_{GhostPopulation}$, a bottleneck duration BD_G and an effective population size $NS_{GhostPopulation}$. The rates of admixture corresponding to the “Ghost population” are ar_G and $1-ar_G$. For all

models, populations were assumed to be isolated from each other, with no exchange of migrants.

Figure S2: Percentage of the mean expected heterozygosity (He) and allelic richness (AR) of the Northern USA source population (Bellefonte, Pennsylvania) for each sample from the three putative parental outbreak areas (NW and NE Italy and CSE Europe) and for samples from Veneto. Dark gray bars indicate the mean expected heterozygosity (He) of the various samples, whereas light gray bars correspond to their allelic richness (AR).

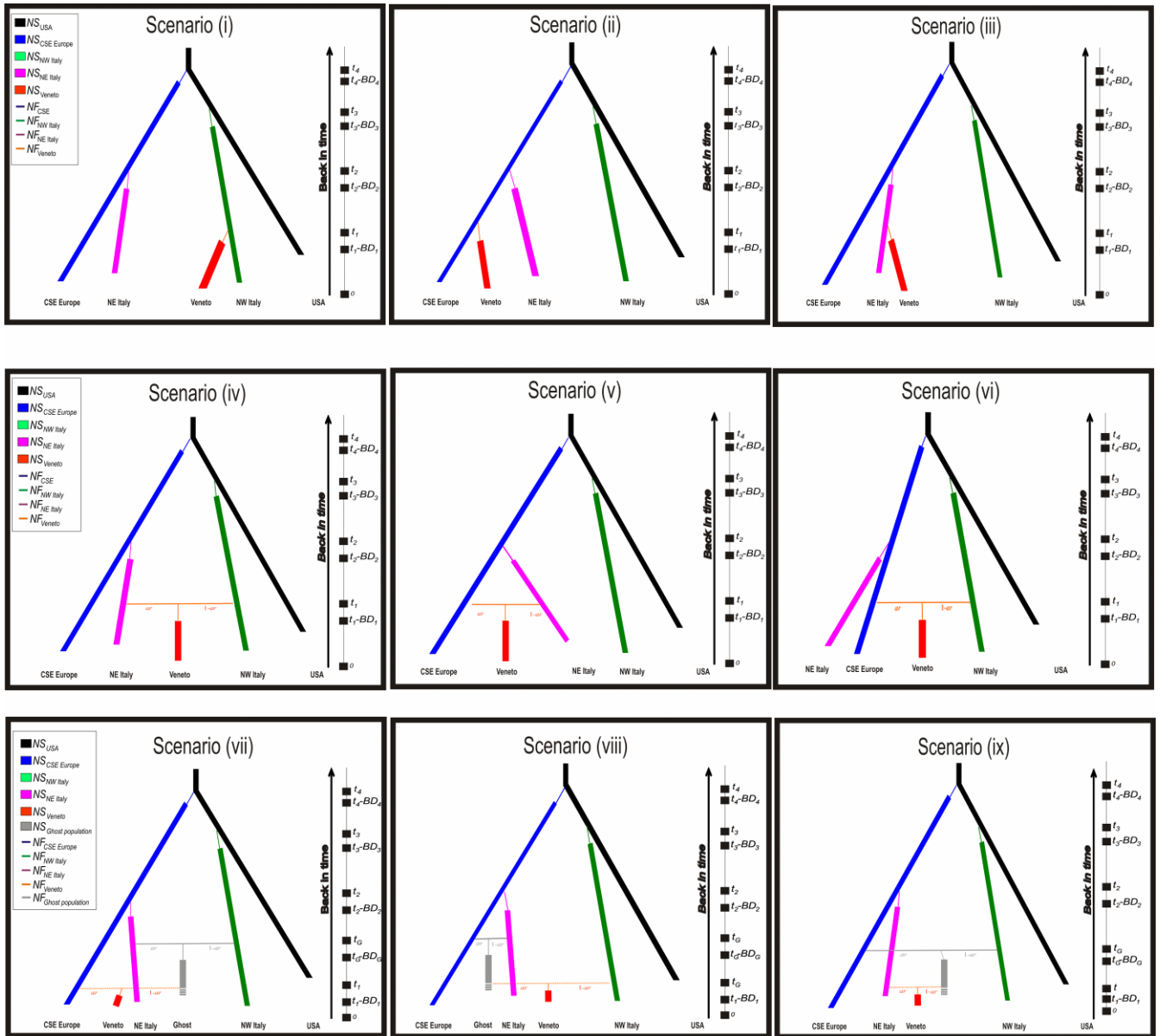


Figure S1

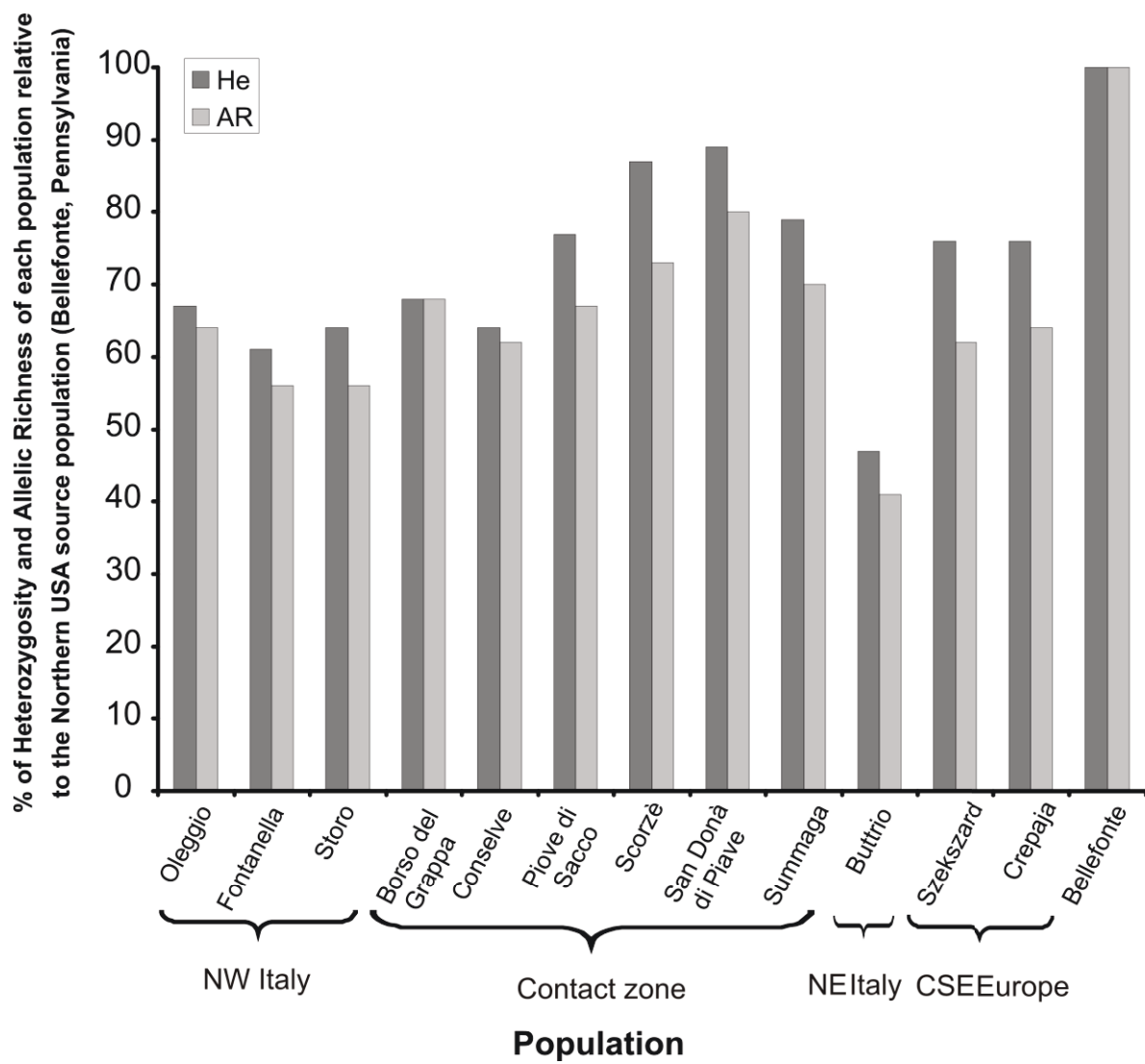


Figure S2

V. – CHAPITRE II : ÉTUDE DE LA ZONE HYBRIDE DE VÉNÉTIE ET ESTIMATION DE LA DISPERSION DE *DIABROTICA VIRGIFERA* *VIRGIFERA*

Ce deuxième chapitre correspond à une étude réalisée à l'aide de marqueurs microsatellites (les mêmes que précédemment) et dans laquelle nous utilisons la théorie des clines et le déséquilibre de liaison pour estimer la dispersion de *DvV* en Europe. Cette étude se rapporte à l'article n°2.

Article n°2 : **Bermond, G.**, Blin, A., Vercken, E., Ravigné, V., Rieux, A., Mallez, S., Morel-Journel, T., Guillemaud, T. (2013). Estimation of the dispersal of a major pest of maize by cline analysis of a temporary contact zone between two invasive outbreaks. *Molecular Ecology* **22**, 5368–5381.

Contexte, objectif visé et questions posées

Les zones hybrides sont particulièrement intéressantes d'un point de vue évolutif car elles nous renseignent sur les différences génétiques et les forces sélectives qui séparent les taxons/populations mis(es) en jeu. Leur étude peut nous donner de bonnes estimations de paramètres évolutifs majeurs tels que le flux de gènes ou l'intensité de la pression de sélection qui s'opèrent respectivement, entre et sur les taxons/populations impliqué(e)s (Barton, Hewitt, 1985; Hewitt, 1988). Le chapitre précédent nous a permis de mettre en évidence l'existence

d'une zone hybride dans la région de la Vénétie, en Italie du Nord, suite au contact secondaire de deux foyers envahissants majeurs de *Diabrotica virgifera virgifera* en Europe. Nous avons donc cherché à exploiter cette zone hybride vénitienne en utilisant la théorie des clines pour la documenter mais aussi pour estimer le paramètre clé de l'invasion européenne de *Dv* : sa dispersion. Cette dernière est définie ici comme l'écart-type de la distribution des distances parents-enfants et symbolisée par σ (voir **section I.2.3**). Les questions sous-jacentes sont les suivantes :

- 1) Quel est l'ordre de grandeur de la dispersion de *Dv* en Europe ?
- 2) Est-il comparable aux mesures déjà effectuées aux USA et en Europe centrale sur cette même espèce ?

Méthodes employées

Pour cette étude, 13 échantillons ont été collectés à l'été 2012 en Italie, Slovénie et Hongrie, le long d'un transect Ouest-Est traversant les deux foyers parentaux d'Italie NO et d'Europe CSE et la zone hybride (Figure 12). Notons que les échantillons utilisés ici sont différents de ceux du chapitre précédent. En moyenne, 35 individus ont été collectés par échantillon et génotypés à 12 marqueurs microsatellites (nous n'avons pas conservé le locus monomorphe *Dv*-D5 pour cette étude).

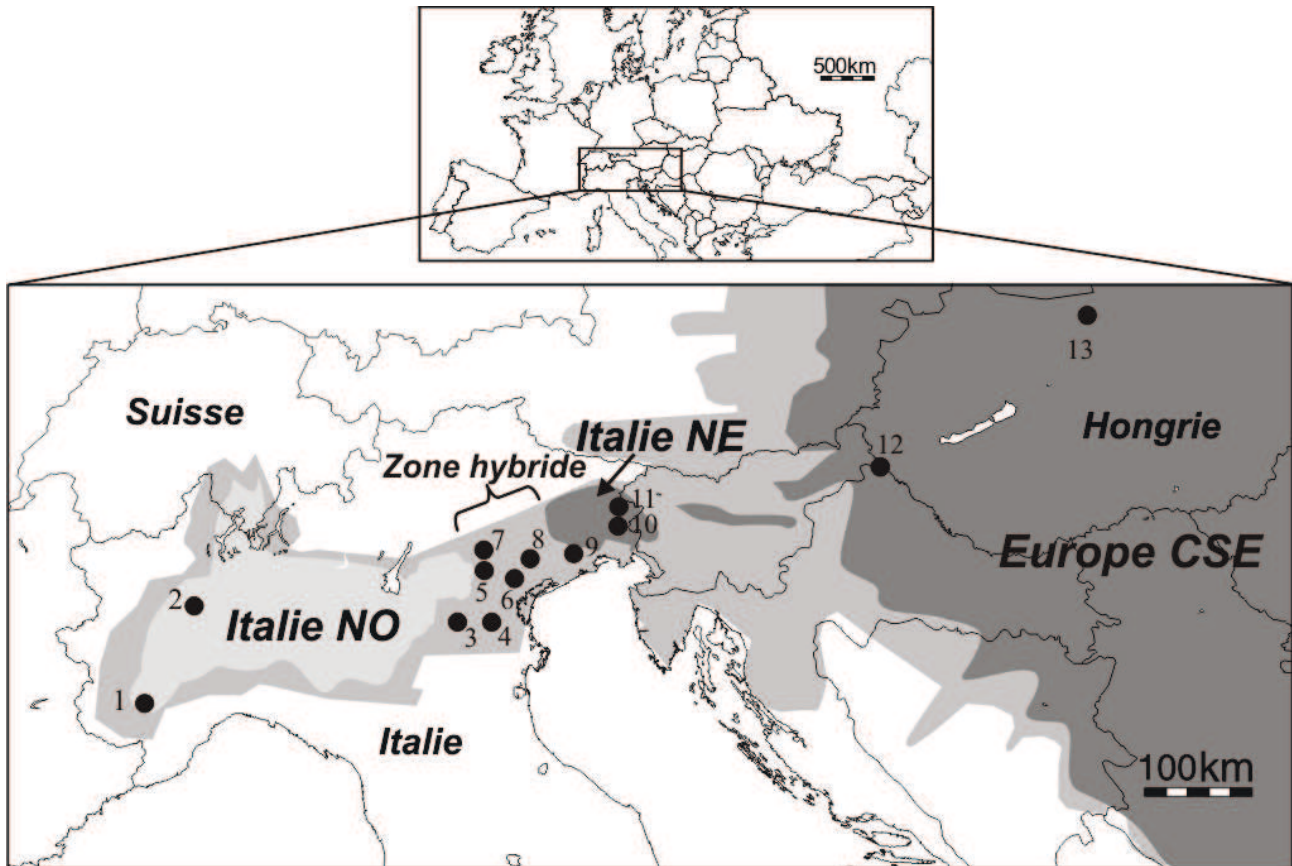


Figure 12 : Localisation des échantillons de *Diabrotica virgifera virgifera* (*Dvv*) collectés en 2012 en Italie du Nord-Ouest (Italie NO), dans la zone hybride, en Italie du Nord-Est (Italie NE) et en Europe Centrale et du Sud-Est (Europe CSE). Les nombres se réfèrent au nom des échantillons : 1 = Cuneo, 2 = Olcenengo, 3 = Vr24, 4 = Conselve, 5 = PD1, 6 = Scorzè, 7 = Borso del Grappa, 8 = Tv22, 9 = Summaga, 10 = Medeuza, 11 = Premariacco, 12 = Pince, 13 = Budapest. L'aire de répartition de *Dvv* avant le contact entre les populations des foyers Ouest et Est (en 2007) est représentée en gris clair (Italie NO) et gris foncé (Europe CSE). L'aire de répartition après le contact (2012) est représentée en gris intermédiaire.

Dans un premier temps les fréquences alléliques ont été analysées pour mettre en évidence la présence de clines sur l'ensemble des loci considérés.

Nous avons ensuite cherché à ajuster des fonctions mathématiques aux clines de fréquences alléliques observés en supposant une dispersion gaussienne des individus. Une Gaussienne cumulée a donc été ajustée à chacun des clines de fréquences alléliques observés et les différents paramètres du modèle (centres, largeurs des clines...) ont été estimés par maximum de vraisemblance. Plusieurs modèles ont été établis pour ajuster au mieux nos données : un modèle considéré comme le plus simple correspondant à un contact secondaire neutre et sans dérive apparente et des modèles plus complexes pouvant refléter l'existence de

sélection ou de dérive apparente. Ces modèles ont ensuite été comparés par maximum de vraisemblance et le meilleur d'entre eux a été sélectionné selon le critère QAICc⁷.

Etant donné que le modèle sélectionné ne comportait qu'une seule largeur pour l'ensemble des clines observés aux loci considérés (voir la section « **Results**, *cline analysis* » de l'**article n°2**), la théorie des clines neutres (voir **section I.3.4**, Endler, 1977) a donc pu être appliquée à nos données pour estimer σ . Lorsque deux populations génétiquement différenciées se rencontrent en l'absence de sélection sur le flux de gènes, des clines de fréquences alléliques se forment au niveau des marqueurs neutres et disparaissent progressivement sous l'effet de l'intensité de la dispersion entre les deux entités génétiques. Dans un contexte de contact secondaire neutre et lorsque la dispersion est correctement approximée par un noyau gaussien d'écart-type sigma, les fréquences alléliques sont solution

de l'équation de diffusion : $\frac{\partial f}{\partial t} = \sigma^2 \frac{\partial^2 f}{\partial x^2}$ (voir la partie «*Theory: allele frequencies after a*

secondary contact » du fichier Mathematica de l'**article n°2**, communication Personnelle de Virginie Ravigné). Le coefficient de diffusion ou l'écart-type de la distribution des distances parents-descendants (σ), peut alors être exprimé comme une fonction de w , la largeur du cline et T , le temps depuis le premier contact (Endler, 1977; Gay *et al.*, 2008; Rieux *et al.*, 2013) :

$\sigma = \frac{w}{\sqrt{2\pi T}}$. En remplaçant w et T par leurs valeurs estimées (respectivement par maximum de

vraisemblance et par Bermond *et al.* (2012), nous obtenons σ .

La dispersion entre populations différenciées génétiquement génère par ailleurs un déséquilibre de liaison (DL), qui atteint un pic au centre du cline et peut permettre d'obtenir une autre estimation de σ (voir **sections I.2.3 et I.3.4**, Barton, 1982; Barton, Gale, 1993). En effet, lorsque la zone hybride est neutre (ou soumise à une sélection faible) et sous

⁷ Ce critère de sélection est une version doublement corrigée de la valeur d'AIC. La première correction permet de prendre en compte la sur-dispersion des données (QAIC) alors que la seconde pénalise un nombre important de paramètres dans le modèle (QAICc).

l'hypothèse d'une migration par diffusion, σ peut s'exprimer comme une fonction de D_{max} , la valeur maximale du déséquilibre de liaison entre deux loci au centre du cline, de r , le taux de recombinaison entre ces deux loci et w_i la largeur du cline à chacun d'entre eux. Cette relation est la suivante : $\sigma^2 = D_{max} r w_1 w_2$. Sous l'hypothèse d'un contact secondaire neutre, la variation spatiale du déséquilibre de liaison peut être ajustée à une Gaussienne. Dans un premier temps nous avons estimé un déséquilibre de liaison moyen entre deux loci, sur l'ensemble des loci considérés et pour chaque échantillon grâce à la méthode de Barton (2000). L'ajustement gaussien de la variation spatiale du DL nous a ensuite permis d'estimer par maximum de vraisemblance sa valeur maximale (D_{max}). En supposant r égal à 0.5 et en utilisant la valeur w précédemment estimée (la même pour tous les loci, voir **section I.3.4**), σ a été estimé une seconde fois. Cette double estimation de σ nous a permis d'estimer T , le temps depuis le contact entre les deux foyers envahissants.

Enfin, nous avons cherché à évaluer l'effet du noyau de dispersion et de la densité de population sur l'estimation de σ . Nous nous sommes, pour cela, écartés des hypothèses de densité de populations égales et de dispersion gaussienne des individus en simulant des données génétiques selon des noyaux de dispersion connus et une certaine hétérogénéité dans les densités de populations.

Résultats

Les résultats obtenus montrent que (i) les marqueurs neutres utilisés dans cette étude ne révèlent aucune preuve de liaison avec des zones du génome sous sélection ; (ii) les différentes méthodes d'estimation de σ (théorie des clines neutres et déséquilibre de liaison) sont congruentes et fournissent des estimations de σ égales à environ 20 km/génération^{1/2} ; (iii) les estimations de σ sont cohérentes avec celles trouvées dans la littérature ; (iv) les simulations réalisées montrent que ces estimations sont robustes lorsqu'on s'écarte des

hypothèses de densité de populations égales et de dispersion gaussienne malgré un faible biais systématique.

**Estimation of the dispersal of a major pest of maize by cline analysis of a temporary
contact zone between two invasive outbreaks**

Gérald Bermond^{1,2,3}, Aurélie Blin^{1,2,3}, Elodie Vercken^{1,2,3}, Virginie Ravigné⁴, Adrien Rieux⁵,
Sophie Mallez^{1,2,3}, Thibaut Morel-Journel^{1,2,3} and Thomas Guillemaud^{1,2,3}

¹INRA, UMR 1355, F-06903 Sophia Antipolis, France

²Université de Nice Sophia Antipolis, UMR Institut Sophia Agrobiotech

³CNRS, UMR 7254

⁴CIRAD, UMR BGPI, Campus International de Baillarguet, F-34398 Montpellier, France

⁵University College London, London, WC1E 6BT, United Kingdom

Keywords: genetic clines, gene flow estimation, linkage disequilibrium, hybrid zone,

Diabrotica virgifera

Gerald Bermond, 400 Route des Chappes BP 167 - 06903 Sophia Antipolis cedex, France,

Fax: (33) 4 92 38 64 01, Gerald.Bermond@sophia.inra.fr

Running title: Estimation of *Diabrotica virgifera* dispersal.

ABSTRACT

Dispersal is a key factor in invasion, and in the persistence and evolution of species. Despite the importance of estimates of dispersal distance, dispersal measurement remains a real methodological challenge. In this study, we characterized dispersal by exploiting a specific case of biological invasion, in which multiple introductions in disconnected areas lead to secondary contact between two differentiated expanding outbreaks. By applying cline theory to this ecological setting, we estimated σ , the standard deviation of the parent-offspring distance distribution, of the western corn rootworm, *Diabrotica virgifera virgifera*, one of the most destructive pests of maize. This species is currently invading Europe, and the two largest invasive outbreaks, in Northern Italy and Central Europe, have recently formed a secondary contact zone in Northern Italy. We identified vanishing clines at 12 microsatellite loci throughout the contact zone. By analyzing both the rate of change of cline slope and the spatial variation of linkage disequilibrium at these markers, we obtained two σ estimates of about 20 km/generation^{1/2}. Simulations indicated that these estimates were robust to changes in dispersal kernels and differences in population density between the two outbreaks, despite a systematic weak bias. These estimates are consistent with the results of direct methods for measuring dispersal applied to the same species. We conclude that secondary contact resulting from multiple introductions is very useful for the inference of dispersal parameters and should be more widely used in other species.

INTRODUCTION

Biological invasions are of considerable significance in applied ecology because they often have major consequences for biodiversity, agriculture and human health (Lodge & Shradler-Frechette 2003; Mack *et al.* 2000; McKinney & Lockwood 1999; Olden *et al.* 2004; Pimentel *et al.* 2001; Ruiz *et al.* 2000). From a strictly scientific perspective, they may also be very useful for investigations of the various evolutionary processes underlying species distributions and abundances, community associations and relationships between native and invasive species (Sax *et al.* 2005, Lockwood, 2007). An understanding of biological invasions also appears to be a fundamental element of the fight to prevent population establishment in the early stages of invasion or to optimize the management of the invading species in the areas invaded.

We need to understand the demographic and evolutionary processes occurring in the invaded area in order to outline the mechanisms underlying biological invasions. The key factor in these processes is dispersal (Johnson & Gaines 1990) — i.e. the set of active or passive phenomena resulting in the establishment of a geographic distance between the breeding sites of individuals from two successive generations — defined here as effective dispersal. Dispersal is a major determinant of biological invasions (Lockwood *et al.* 2007), as it leads to expansion of the geographic range of the introduced species after its establishment outside its native area (Wilson *et al.* 2009), and because it determines the level of gene flow between introduced populations. The estimation of dispersal distances can thus be used to characterize the invasiveness of a species (Rehage & Sih 2004), to predict the expansion speed of invasive populations (Lockwood *et al.* 2007) and to evaluate the balance between migration and selection (Lenormand & Raymond 2000) controlling the distribution range of expanding populations (Kirkpatrick & Barton 1997).

Despite the importance of accurate dispersal distance estimates, it remains a real challenge to measure dispersal, due to the complexity of the methods involved and associated technical constraints (Broquet & Petit 2009; Rousset 2001). "Direct" methods, based on direct observations of the movements of individuals (Dobzhansky & Wright 1943), are logistically very demanding and do not generally provide estimates of effective dispersal. "Indirect" genetic methods (reviewed in Broquet & Petit 2009) are often based on assumptions, such as migration-drift equilibrium (Rousset 1997; Wright 1943), and thus reveal long-term dispersal patterns between populations (Peacock 1997; Slatkin 1987). In addition, the population genetics models on which these methods are based are dependent on highly stochastic processes (such as drift), resulting in a large variance for parameter estimates (Bossart & Pashley Prowell 1998).

Biological invasions are natural experiments of non-equilibrium population dynamics in which dispersal can be estimated directly from transient situations. Appropriate theoretical models can be used to estimate dispersal directly from (i) a knowledge of the expansion speed of expanding populations (e.g. Skellam's model (Skellam 1951) or Fisher's model (Fisher 1937)) or (ii) description of the neutral spatial genetic structure of the contact zone between two expanding populations (Barton 1982; Barton & Gale 1993) and its temporal dynamics (Endler 1977). Many studies have investigated the expansion speeds of invading populations (e.g. Arim *et al.* 2006; Phillips *et al.* 2006), and some have provided estimates of dispersal (e.g. Andow *et al.* 1990; Lubina & Simon 1988). By contrast, transient contact zones between invading populations have seldom been used to estimate dispersal. Only one study to date seems to have used this approach (Rieux *et al.* 2013), even though this biological context is common and the methodology is easy to implement. Biological invasions frequently begin with multiple independent introductions at different, remote locations (Bossdorf *et al.* 2005;

Ciosi *et al.* 2008; Dlugosch & Parker 2008; Facon *et al.* 2003; Kolbe *et al.* 2004; Lockwood *et al.* 2005; Miller *et al.* 2005; Roman & Darling 2007; Sax *et al.* 2005). In cases of bottlenecks during introduction and/or drift during establishment, considerable genetic differentiation between the introduced populations may occur (e.g. Ciosi *et al.* 2008). The spatial expansion of these populations may lead to secondary contact (e.g. Rieux *et al.* 2013) and, eventually, to the formation of transient clines of allelic frequencies (e.g. Bermond *et al.* 2012). In cases of recent introduction, the introduced populations may be isolated for too short a time for selection to act against their hybrids (Edmands 2002; Orr & Turelli 2001). In case of neutral diffusion, allelic frequencies therefore tend to become homogeneous throughout the secondary contact zone over time. Cline slope declines at a rate dependent on dispersal strength and time since contact (Endler 1977; Rieux *et al.* 2013). This rate can therefore be used to estimate dispersal. Furthermore, dispersal between differentiated populations generates linkage disequilibrium, which peaks in the center of the cline and can be used to obtain another estimate of dispersal (Barton 1982; Barton & Gale 1993). These methods are not completely independent because they both rely on the estimation of the cline slope. They are not dependent on migration-drift equilibrium, instead being based on the instantaneous impact of dispersal on the genetic structure of the transient contact zone between two expanding populations.

In this study, we estimated the dispersal of one of the major pests of maize, the western corn rootworm (Coleoptera: Chrysomelidae), *Diabrotica virgifera virgifera*, which is currently invading Europe. *D. v. virgifera* completes its life cycle in a year (Branson & Krysan 1981). Eggs are laid during the summer and they overwinter in the soil, hatching in late spring. They then pupate and the adults emerge in early summer. They mate, disperse and lay eggs during the summer (Branson & Krysan 1981; Spencer *et al.* 2009). *D. v. virgifera*

was introduced several times in Europe from the USA (Miller *et al.* 2005; Ciosi *et al.* 2008) and was detected for the first time in Europe near Belgrade Airport in 1992 (Sivcev *et al.* 1994). Population genetics analyses (Miller *et al.* 2005) showed that multiple introductions of *D. v. virgifera* were responsible for the European invasion and that the two principal invasive outbreaks in Europe, in North-West Italy (NW Italy, first observations in 2000) and Central and Southern Eastern Europe (CSE Europe, first observations in 1992) were founded independently. Bermond *et al.* (2012) have recently shown that these two outbreaks have converged, coming into contact in 2008, and that they now form a continuum in the Veneto region of Northern Italy. There is considerable genetic differentiation between these two invasive populations because of intense bottleneck during their introduction (Ciosi *et al.* 2008), so this area has become an admixture zone, with allelic frequency clines at microsatellite loci (Bermond *et al.* 2012). It therefore presents an ideal opportunity for estimation of the dispersal distances of *D. v. virgifera*.

We focused on the contact zone between the two largest invasive populations of *D. v. virgifera* in Europe to estimate the migration variance of this species, σ^2 , the variance of the parent-offspring distance distribution. We carried out a spatial analysis of linkage disequilibrium and a temporal analysis of frequency clines on the basis of individual multilocus genotypes for 12 microsatellite loci. We also performed simulations to evaluate the effect of departure from common hypotheses on the estimate obtained.

MATERIALS AND METHODS

Sample collection

Samples were collected in summer 2012, from July 4 to 22, at 13 sites (see details in Table 1) in Italy, Slovenia and Hungary, along a west-east transect crossing the outbreaks of

NW Italy and CSE Europe (Figure 1). The transect axis was chosen to match the direction of expansion of the outbreaks in this area (Figure 1). For each sampling site, between 23 and 41 individuals were collected from the same maize field, and GPS coordinates were recorded. A total of 466 adult beetles were caught by hand (funnel bound with a 50 ml tube filled with 96% alcohol) or with a mouth vacuum. All *D. v. virgifera* samples were stored at -20°C in 96% ethanol until DNA extraction.

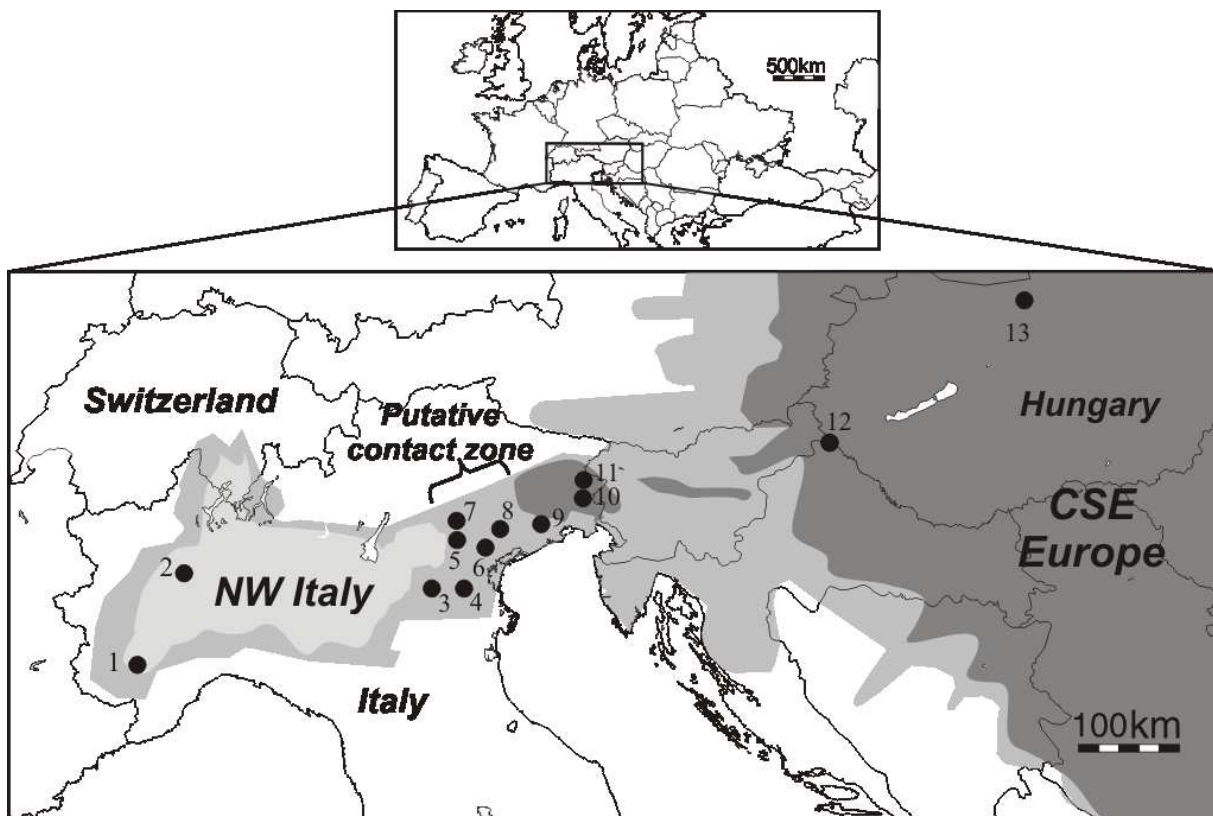


Figure 1: Location of the European *D. v. virgifera* samples collected in 2012 in North West (NW) Italy, in the contact zone and in Central South Eastern (CSE) Europe. Numbers refer to sample names (see in Table 1). The distribution areas of *D. v. virgifera* before contact between the western and eastern outbreak populations (in 2007) are shown in light gray (NW Italy) and dark gray (CSE Europe). The distribution area after contact (in 2012) is shown in medium gray.

Table1: Description of the *Diabrotica virgifera virgifera* samples from Northern Italy and Central and South-Eastern Europe. N is the sample size. x_i and y_i are the two-dimensional (2D) geographic coordinates of sample i obtained from its GPS coordinates and converted into km with orthodromic distances.

Sample name	Sample number	Region (Country)	Year of 1 st observation	N	x_i	y_i	GPS coordinates (Latitude; Longitude)
Cuneo	1	Piedmont (Italy)	2000	41	0	0	(44°28'39.10"N; 07°34'24.10"E)
Olcenengo	2	Piedmont (Italy)	2000	35	58	99	(45°21'59.50"N; 08°18'18.40"E)
Vr24	3	Veneto (Italy)	2002	40	302	85	(45°14'35.50"N; 11°22'40.70"E)
Conselve	4	Veneto (Italy)	2002	35	340	84	(45°14'40.60"N; 11°51'31.40"E)
PD1	5	Veneto (Italy)	2002	40	333	130	(45°38'40.50"N; 11°45'55.80"E)
Scorzè	6	Veneto (Italy)	2002	33	369	120	(45°33'90.29"N; 12°13'27.40"E)
Borso del Grappa	7	Veneto (Italy)	2002	40	334	148	(45°48'26.30"N; 11°46'51.80"E)
Tv22	8	Veneto (Italy)	2002	40	377	137	(45°42'43.90"N; 12°19'49.70"E)
Summaga	9	Veneto (Italy)	2002	24	424	144	(45°46'18.91"N; 12°55'01.60"E)
Medeuzza	10	Friuli (Italy)	2003	23	462	162	(45°55'58.10"N; 13°24'02.50"E)
Premariacco	11	Friuli (Italy)	2003	34	462	175	(46° 03'17.00"N; 13°24'00.00"E)
Pince	12	Prekmurje (Slovenia)	2003	41	710	227	(46°31'15.38"N; 16°31'23.70"E)
Budapest	13	Pest (Hungary)	1992	40	888	338	(47°31'12.90"N; 18°46'33.10"E)

DNA extraction and microsatellite analyses

Individuals were washed three times in 0.065% NaCl before extraction, to remove ethanol from the tissues. For all samples, we extracted DNA from the thorax or half body, cut lengthwise, with the DNeasy® tissue kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions, with an elution volume of 100 µl. Twelve microsatellite loci of *D. v. virgifera* (including di and trinucleotides) (details available in Bermond *et al.* 2012) were amplified in three separate multiplex PCR performed in a PTC-225 MJ Research thermocycler and were analyzed as described by Miller *et al.* (2007).

Cline analysis

Descriptive analysis and cline fitting

We first tested Hardy-Weinberg equilibrium for each sample, with the probability test approach implemented in GENEPOP version 4.0.1 (Raymond & Rousset 1995; Rousset 2008). We used the sequential Bonferroni procedure (Sokal & Rohlf 1995) to take multiple testing into account. For cline analysis, the alleles of each microsatellite were clustered into two groups identified in a multiple correspondence analysis (Bierne *et al.* 2003; Daguin *et al.* 2001), allowing maximal frequency variance along the cline. Assuming a scenario of neutral secondary contact, the frequency f_{ijk} of allele group k of locus j in sample i can be described by the scaled cumulative Gaussian distribution (a Mathematica notebook has been submitted to Dryad at doi:10.5061/dryad.10kd4 with full details):

$$f_{ij1}(\rho_{ij}) = h_{1j} + h_{pj}(1 - h_{1j}) \int_{-\infty}^{\rho_{ij}} \frac{1}{b_j \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{c_j - \rho_{ij}}{b_j})^2} \text{ and } f_{ij2}(\rho_{ij}) = 1 - f_{ij1}(\rho_{ij}) \quad \{\text{Eqn.1}\}$$

where ρ_{ij} is the geographic coordinate of sample i along the cline axis of locus j , c_j is the cline center, b_j is a cline slope parameter (and the standard deviation of the Gaussian), and h_{1j} , and $h_{1j} + h_{pj}(1 - h_{1j})$ are the frequencies of the first allele group at the two extremities of the cline of

locus j (see the Mathematica notebook). For each locus j , the angle θ_j between the parallel and the cline axis was used to calculate ρ_{ij} , the transformed geographic distance of each sample i along the cline axis of locus j , $\rho_{ij} = x_i \cos(\theta_j) + y_i \sin(\theta_j)$, where x_i and y_i are the two-dimensional (2D) geographic coordinates of sample i obtained from its GPS coordinates and converted into km with orthodromic distances (see Table 1).

Model selection and parameter estimation

Various models of scaled cumulative Gaussian distributions, differing in the number of parameters included, were fitted to the allelic group frequencies of the 12 microsatellites, making it possible to test hypotheses about changes in cline shape. The most complex model (model M₈) had five parameters (h_{1j} , h_{pj} , c_j , b_j and θ_j) per locus j and the simplest model (model M₁) had two parameters (h_{1j} , h_{pj}) per locus j plus three parameters (c , b and θ) common to all loci (all details of model configurations available in Table 2). Assuming a pure diffusion process, we would expect to select model M₁ with coincident, concordant and parallel clines (i.e. with identical parameters c , b and θ) for all loci. Assuming a diffusion process with selection acting on some genomic regions to which some of the microsatellites are linked or strong genetic drift, parameters c_j , b_j and θ_j might be expected to vary between loci, and we would expect to select a model other than model M₁. We thus compared the eight models, with various levels of constraint on parameters c_j , b_j and θ_j , using the information theory approach proposed by Burnham and Anderson (2010). The best model was chosen on the basis of the corrected QAIC_c criterion (Burnham and Anderson 2010). This modified Akaike criterion takes into account sample size and overdispersion of the data by correcting the deviance with the variance inflation factor \hat{c} (Lebreton *et al.* 1992).

Estimates of the scaled cumulative Gaussian distribution parameters of the chosen model were obtained by maximizing the log-likelihood of the observed frequencies:

$$\sum_i \sum_j \sum_k n_{ijk} \ln(f_{ijk}) \quad \{\text{Eqn.2}\},$$

with n_{ijk} the observed number and f_{ijk} the expected frequencies of individuals displaying allele group k at locus j in population i of {Eqn.1}. Maximum likelihood (ML) estimates of model parameters h_{1j} , h_{pj} , c_j , b_j and θ_j were obtained simultaneously with the FindMaximum function of Mathematica 8.0 (Wolfram Research 2010) (see the Mathematica notebook). At least 3 estimations were obtained with various starting values of the parameters to avoid local likelihood maxima.

Linkage disequilibrium analysis

Descriptive analysis of linkage disequilibrium

The second-order multilocus cumulant $\kappa_{0,2}$ defined by Barton (2000) was used as the average multilocus linkage disequilibrium, taking into account allelic frequencies and their variation. As pointed out by Gay *et al.* (2008), estimates of $\kappa_{0,2}$ are subject to drift and sampling errors. Moreover, no estimate of $\kappa_{0,2}$ was available for the precise center of the cline because there was no sampling site at this position. We therefore fitted a spatial function to $\kappa_{0,2}$. In secondary contact with pure diffusion and no selection, linkage disequilibrium — here assessed by measuring $\kappa_{0,2}$ — should be approximately proportional to a Gaussian distribution (see the Mathematica notebook). We therefore fitted it to a function of the form:

$$LD(\rho_i) = D_{\max} e^{-f(\rho_i - g)^2} \quad \{\text{Eqn.3}\},$$

where ρ_i is the geographic coordinate of sample i along the cline axis, D_{\max} is the maximum value of linkage disequilibrium at the center of the cline g , and f is a variance parameter of the Gaussian distribution. Note that in cases of pure diffusion, the clines of all loci have the same center and are parallel.

Maximum likelihood estimation of maximal linkage disequilibrium

Multilocus genotypic frequencies depend on allelic frequencies, heterozygote deficits or excesses and linkage disequilibrium. As no deviation from Hardy-Weinberg equilibrium was recorded for any locus, multilocus genotype frequencies depended on allelic frequencies and linkage disequilibrium only in this study. For each sample i , an estimate of $\kappa_{0,2i}$ can thus be obtained by maximizing the log-likelihood LL_i of the multilocus genotype frequencies (Barton 2000, see the Mathematica notebook for details). Note that locus Dvv-T3 was removed from this analysis because of its very low variation in space and for the sake of computational feasibility. In addition, the total log-likelihood $\sum_i LL_i$ in which each $\kappa_{0,2i}$ was replaced by $LD(\rho_i)$ (Eqn.3) was used to fit the spatial model of linkage disequilibrium. Maximum likelihood (ML) estimates of the spatial model parameters, including $Dmax$, were obtained simultaneously using a simulated annealing algorithm, adapted from the work of Szymura and Barton (1986) and Gay *et al.* (2008), and implemented in Mathematica 8.0 (Wolfram Research, Inc., 2010) using the code of Gregory (2013) (see the Mathematica notebook).

For analyses of both clines and the spatial distribution of linkage disequilibrium, the associated support limits of parameter estimates were defined as the range of values within two units of log-likelihood from the maximum (Edwards 1972).

Estimation of dispersal

Estimation of dispersal from cline width analysis

When two genetically differentiated populations meet in the absence of selection against gene flow, clines are formed at neutral markers and progressively vanish at a pace determined by the amount of dispersal between the two genetic entities. Specifically, when the dispersal kernel is well approximated by a diffusion equation (e.g. a Gaussian kernel), the decay of the cline after T generations since contact results in a cline width given by:

$$w = \sigma\sqrt{2\pi T} \quad \{\text{Eqn.4}\}$$

(Endler 1977; Gay *et al.* 2008), where σ is the diffusion coefficient (which, here, is also equal to the standard deviation of the normally distributed parent-offspring distances). If the widths of the clines are all the same, the point estimate of the parent-offspring variance distance and its support limits can thus be estimated by calculating the maximum log-likelihood of {Eqn.2}, in which we replaced f_{ijk} by the scaled cumulative Gaussian distribution of {Eqn. 1}, and parameters h_{1j} , h_{pj} , c_j , and θ_j by their point estimates. The maximum slope of the scaled cumulative Gaussian distribution is $\frac{1}{b\sqrt{2\pi}}$ (see the Mathematica notebook). Thus, in cases of a pure diffusion process, the width of the clines is the same for all neutral markers and is given by $w = b\sqrt{2\pi}$. We therefore replaced b with $w/\sqrt{2\pi} = \sigma\sqrt{T}$ in {Eqn.2}. Finally, historical observations suggest that these two invasive populations first came into contact four generations ago (Bermond *et al.* 2012). We thus used this value for T in {Eqn.2}. This method is referred to hereafter as Endler's method.

Estimation of dispersal from linkage disequilibrium analysis

Positive linkage disequilibrium is expected in cases of admixture between two genetically differentiated populations. According to Barton (1982; 1993), the maximum linkage disequilibrium between two loci occurs at the center of the contact zone and can be expressed as:

$$D_{\max} = \frac{\sigma^2}{r} \frac{1}{w_1 w_2} \quad \{\text{Eqn.5}\}$$

where w_i is the cline width at locus i , and r is the rate of recombination between the two loci. Even though Barton proposed this method for use with clines maintained by selection (e.g. Barton 1982), the method assumes an absence of selection at the center of the cline and is thus suitable for the neutral case presented here. If the widths of all the clines are the same

($w_i = b\sqrt{2\pi}$ for all microsatellites), the joint estimation of w and σ is obtained by maximizing the joint log-likelihood ($\sum_i LL_i + \sum_i \sum_j \sum_k n_{ijk} \ln(f_{ijk})$) of the data, in which each $\kappa_{0,2i}$ is replaced with $LD(\rho_i)$ ({Eqn.3}), and D_{\max} is replaced with $\frac{\sigma^2}{r} \left(\frac{1}{b\sqrt{2\pi}} \right)^2$. This method is referred to hereafter as the LD method.

The combination of {Eqn. 4} and {Eqn. 5} leads to:

$$T = \frac{1}{2\pi r D_{\max}} \quad \{\text{Eqn. 6}\}$$

This provides a way, independent of the estimation of σ , to estimate the time since contact T by maximizing the log-likelihood of the multilocus genotype frequencies $\sum_i LL_i$ in which each $\kappa_{0,2i}$ is replaced with $LD(\rho_i)$ ({Eqn.3}).

Simulations

The AIC-based comparison described above can be used to compare models with small or large numbers of parameters, when the parameters are common to or different for all loci, respectively. This method can be used to detect the effect of processes, e.g. selection, affecting the various loci differently. Demographic parameters, such as dispersal and population density, affect all neutral genetic markers equally and therefore cannot be evaluated by AIC-based model comparison. We thus simulated genetic data with a known dispersal kernel and heterogeneity of population density, to evaluate their effect on the estimation of dispersal and time since contact. Simulations used discrete time processes and included various dispersal kernels with various kurtoses and various population density differences across the contact zone. The variance of parent-offspring distance was then estimated from these simulated data with Endler's method and the LD method, and was

compared with to its known value. Time since contact was estimated by combining Endler's and LD methods and also compared to its known value.

Genotypic frequencies at two biallelic loci were simulated by a series of 700 demes connected by migration (Lenormand *et al.* 1998; Mallet *et al.* 1990) with reflecting boundaries. Reproduction was simulated by assuming Hardy-Weinberg equilibrium in each deme and independence between loci. In each generation, population densities were set to 1000 individuals, except for the simulations with varying densities. For those simulations, population densities were set to K_1 (varying between 200 and 1000) and K_2 (1000) individuals per deme for $x < z$, and $x > z$, respectively, with z the location of the initial contact between the two populations. Allelic frequencies at each locus were set to 1 for $x < z$, and to 0 $x > z$ just before contact. Diploid individuals dispersed randomly according to an isotropic migration model with various dispersal kernels. These kernels were mixtures of Gaussian distributions or exponential power distributions, including the Laplace kernel, differing in kurtosis and tail heaviness (Table S1). Parameter values of each dispersal kernel were chosen so that all distributions had the same variance, close to the estimated σ^2 for *D. v. virgifera*, but different kurtoses (γ ranging from -1 to 30, Table S1, Figure 5). For each set of parameters, five replicates of the dynamics of genotypic frequencies were simulated for four generations after contact.

For each parameter set, simulations were performed with R-2.15.1 (R Development Core Team 2012). The R script used for the simulations is available on Dryad at doi:10.5061/dryad.10kd4. As only two loci were simulated, it was not necessary to calculate an average multilocus linkage disequilibrium $\kappa_{0,2}$, but only the usual two-locus index of linkage disequilibrium. As individuals were sampled after dispersal, r was replaced by $r/(r + 1)$ in {Eqn.5}, as suggested by Barton & Gale (1993). We were not interested in evaluating

the effect of sampling variance on the precision of the estimations. Therefore 2 loci, 100 demes and 100 individuals were sampled per deme rather than the 12 loci and 13 samples of about 30 individuals in the real dataset. For this reason and for reasons of tractability, we provide only point estimates of σ and T , without support limits, in the simulations.

RESULTS

Cline analysis

We observed no significant deviation from Hardy-Weinberg equilibrium (not shown here) after sequential Bonferroni analysis (Sokal & Rohlf 1995), and western (Cuneo and Olcenengo) and eastern populations (Pince and Budapest) displayed significant genetic differentiation ($p < 10^{-5}$, mean value of $F_{ST} = 0.26$). All polymorphic loci except for DVV-T3 displayed a clinal pattern (Figure 2). According to the QAICc criterion, the model best fitting the allelic frequencies at the 12 microsatellite loci was the simplest model, M_1 , which accounted for 86% of the deviance (Table 2 and Figure 2). All frequency clines are thus coincident (they have the same center), concordant (they have the same width), and parallel (they are at the same angle to the parallel axis). These results are consistent with a pure diffusion model and thus allow estimation of the dispersal of *D. v. virgifera* by Endler's method and the LD method. According to the M_1 model, we obtained the following ML estimations: the center of the clines is located 293 km (with support limits of 265-321 km) from Cuneo, the width of the clines is 91.6 km (80.8-103.1 km) and the angle between the cline and the x axis is 1.08 radians (0.96-1.18 radians).

Table 2: Comparison of the quality-of-fit of the various clinal models of the spatial pattern of microsatellite frequencies for *D. v. virgifera* in the contact zone between the two main European invasive populations. *c* indicates the center of the cline, *s* the slope and θ the angle. QAICc is the modified Akaike criterion for sample size and for overdispersion (see the text for details). The best fit according to QAICc is shown in bold typeface.

Model	M₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈
Configuration	1 <i>c</i>, 1 <i>s</i>, 1 θ	12 <i>c</i> , 1 <i>s</i> , 1 θ	1 <i>c</i> , 12 <i>s</i> , 1 θ	1 <i>c</i> , 1 <i>s</i> , 12 θ	12 <i>c</i> , 1 <i>s</i> , 12 θ	1 <i>c</i> , 12 <i>s</i> , 12 θ	12 <i>c</i> , 12 <i>s</i> , 1 θ	12 <i>c</i> , 12 <i>s</i> , 12 θ
Maximum Log likelihood	-4395.17	-4359.67	-4360.01	-4376.41	-4348.71	-4338.21	-4339.00	-4337.20
Number of parameters	27	38	38	38	49	49	49	60
QAICc	1178.09	1205.24	1205.33	1209.46	1246.15	1243.50	1243.70	1304.4
% of Deviance	86	87	87	87	88	88	88	88

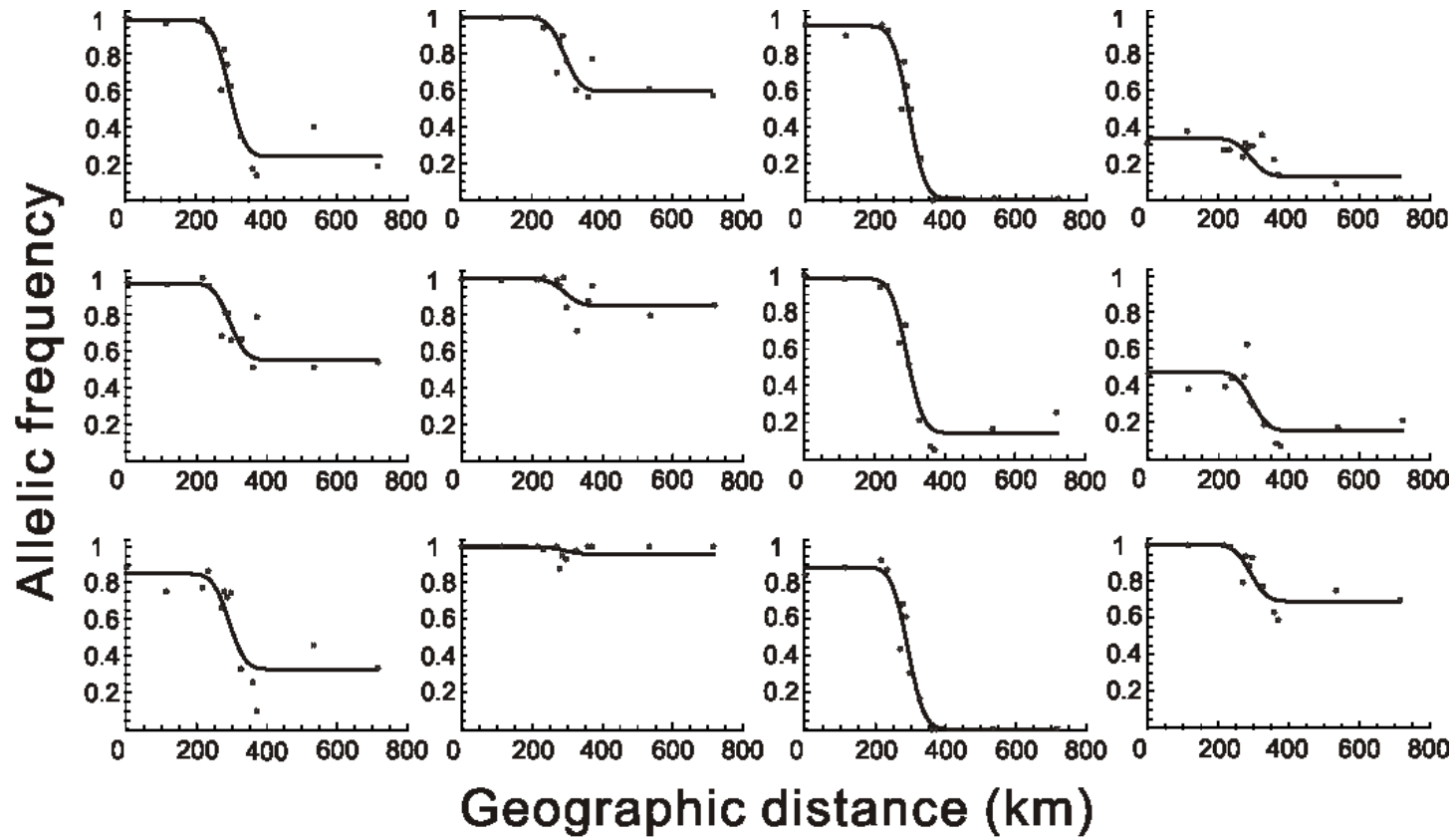


Figure 2: Observed (dots) and fitted (lines) frequencies of the allelic groups of the 12 polymorphic microsatellites (from top to bottom and left to right: Dvv-D2, Dvv-D4, Dvv-D8, Dvv-D9, Dvv-D11, Dvv-T2, Dvv-ET1, Dba-01, Dba-05, Dvv-T3, Dvv-D12, Dba-07) from *D. v. virgifera* in the region of contact between the two main invasive populations in Europe. Allelic group frequencies are represented as functions of geographic distance, along the cline axis (see text for details).

Assuming that the western and eastern populations came into contact four generations before sampling (i.e. in 2008, Bermond *et al.* 2012), Endler’s method provides a ML estimate of the standard deviation σ of the parent-offspring distances measured in one dimension of 18.26 km/generation^{1/2} (16.11-20.56 km/generation^{1/2}). If uncertainty about the year of contact was taken into account (T varying from 2 to 5 years) then the ML estimate of σ and its support limits ranged between 14.4 and 29.1 km/generation^{1/2} (Figure 3A).

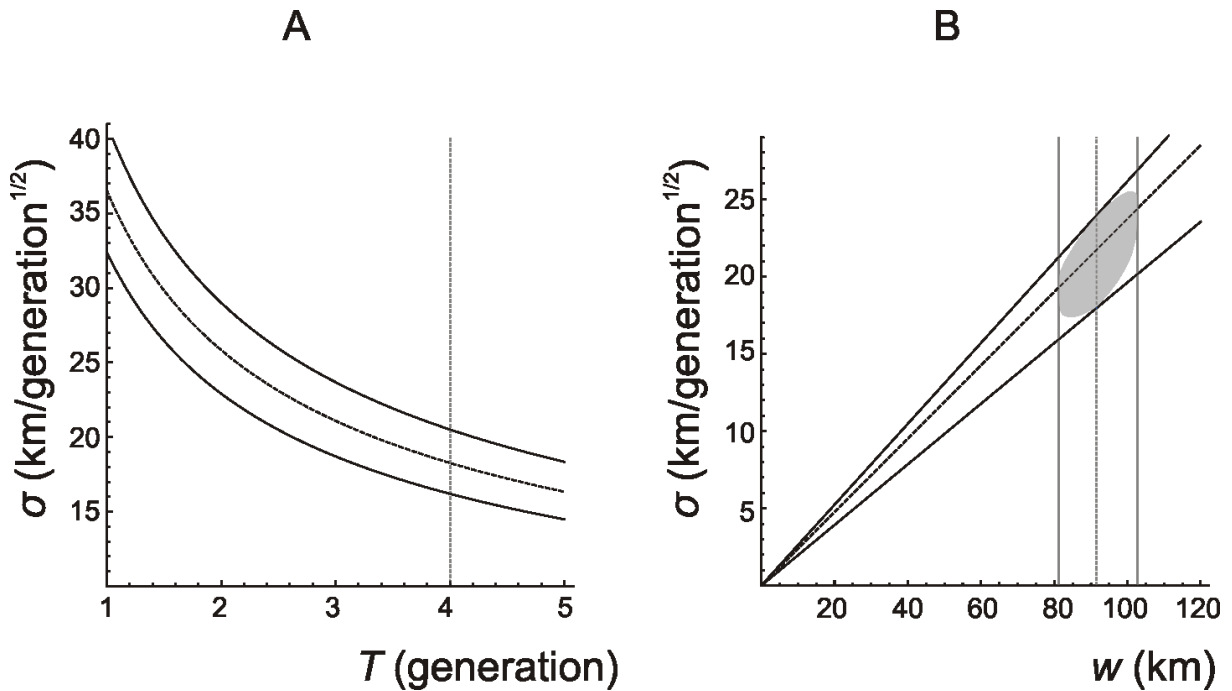


Figure 3: Maximum likelihood (ML) estimates of σ , the standard deviation of offspring-parent dispersal distances, obtained by the Endler’s and LD methods. (A) ML estimate of σ as a function of the time since contact T between the NW Italian and CSE European invasive populations of *D. v. virgifera*. Dashed and solid line indicate the ML estimate of σ obtained by the Endler’s method and its associated support limits, respectively. The vertical dashed line in gray corresponds to the value of T (4 years) suggested by historical information. (B) ML estimate of σ obtained by the LD method. The support area of σ and the width of the clines w is shown in gray and is obtained from the joint support limits of D_{max} (the maximum value of linkage disequilibrium at the center of the cline) and w . The dashed and solid black lines correspond to the ML estimates of D_{max} and its associated support limits, respectively. The dashed and solid gray lines correspond to the ML estimates of w and its associated support limits, respectively.

Linkage disequilibrium analysis

Most ML estimates of average linkage disequilibrium $\kappa_{0,2}$ had support limits including 0, except for five samples close to the center of the clines (Figure 4). The maximum estimate of $\kappa_{0,2}$ was 0.11 (0.09-0.12), in Summaga, Veneto (Figure 4). ML estimation of the

parameters of the spatial fit of linkage disequilibrium provided an estimate of 0.113 (0.078-0.136) for maximum linkage disequilibrium D_{max} (Figure 4).

As a model with concordant clines for all loci was selected (see above), {Eqn.5} (Barton & Gale 1993) can be reduced to $\sigma = w\sqrt{D_{max}r}$ where w is the cline width common to all loci. Figure 3B shows the support area of w and σ , assuming that all loci are physically independent ($r = 0.5$). The ML estimate of σ is 21.74 km/generation^{1/2} with a support area bounded by 17.30-25.52 km/generation^{1/2} (Figure 3b, Mathematica notebook).

By combining Endler's and LD's method, we obtained an estimate of the time since contact T of 2.82 years (2.33-4.08 years).

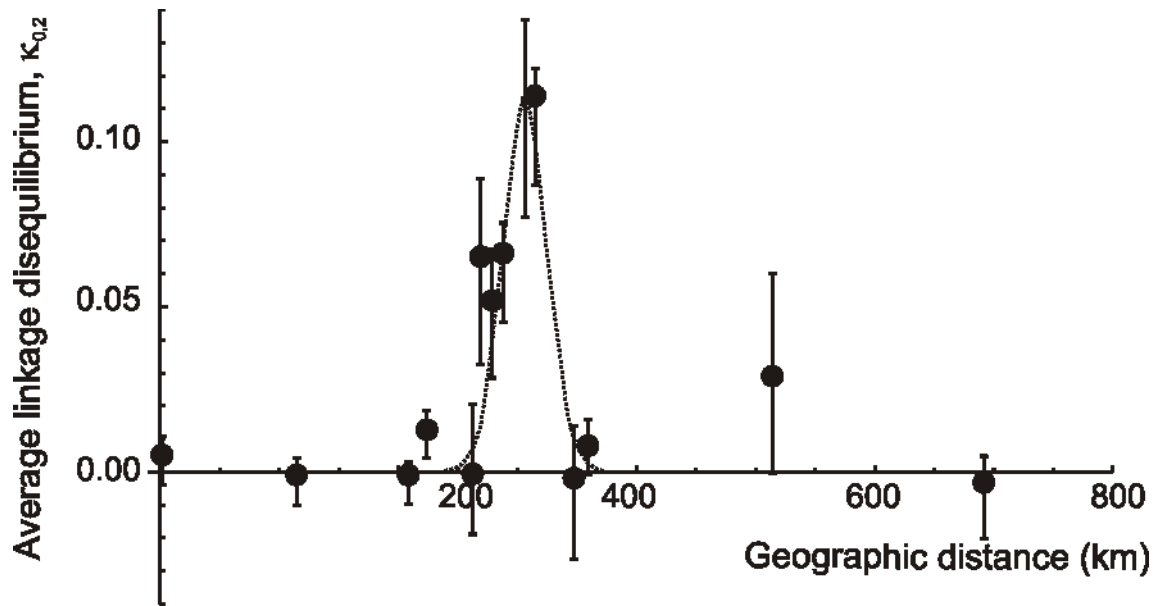


Figure 4: Maximum likelihood estimates (full circles) and spatial fit (dotted line) for the average linkage disequilibrium, $\kappa_{0,2}$ (Barton 2000), of *D. v. virgifera* in the contact zone between the two main European invasive populations in North West Italy and Central Europe, as a function of geographic distance along the cline axis. Vertical bars indicate the support limits for the estimated values of $\kappa_{0,2}$ and the maximum average linkage disequilibrium D_{max} .

Simulations

With Gaussian and exponential kernels and no difference in population size between the two sides of the contact zone, relative biases of the estimate of the standard deviation of the parent-offspring distance were consistently below 12% (Figure 5), and those of the time since contact were between 10 and 15%. There were systematic differences between estimation methods, with LD method producing the largest estimates of σ and the largest biases. Overall, the estimates of σ and T were only weakly biased, whatever the kurtosis value and population density differences, with errors always below 30%. The largest relative biases of σ estimation were found for extreme values of kurtosis (with biases of about 25% for $\gamma = 30$) for the mixture of Gaussians and for the highest density differences between the two sides of the contact zone (with biases of about 20% for $K_1/K_2=0.2$). The largest relative biases of T estimation (of about 30 %) were found for the highest density differences between the two sides of the contact zone ($K_1/K_2=0.2$).

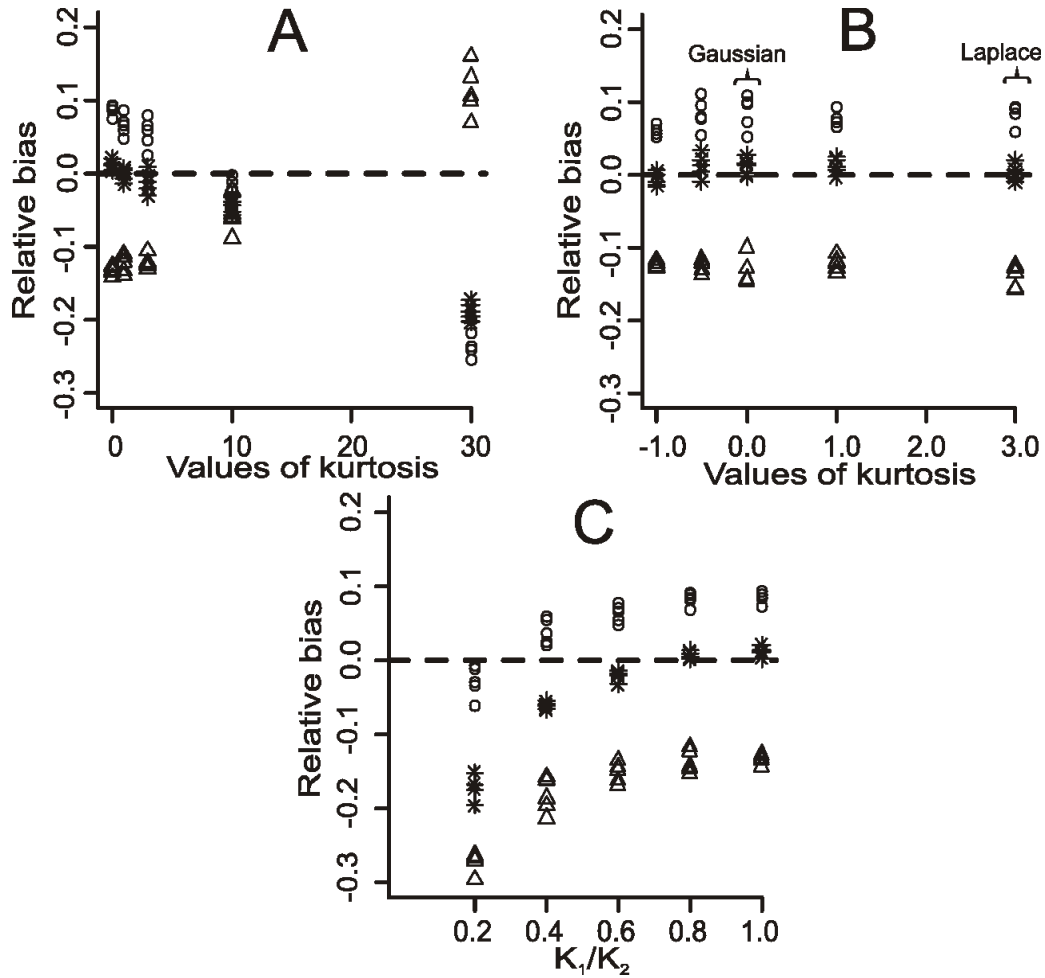


Figure 5: Influence of dispersal kernels (with mixtures of Gaussian kernels in A and exponential power kernels in B) and population density differences between the two sides of the contact zone (C) on the relative bias of ML estimates of the dispersal σ and the time since contact T as revealed by simulations. The relative bias of parameter X is defined as $\frac{\hat{X} - X}{X}$, with \hat{X} the estimated value and X , the value of the parameter used to simulate the data. Number of demes = 700, number of individuals per deme = from 200 to 1000, $\sigma = 20$ km/generation^{1/2}, number of simulated generations = 4 (see the R script used for the simulations in Appendix S2). The characteristics of dispersal kernels are presented in Supporting Information (Table S1). Dashed lines indicate a relative bias of zero. The ratio K_1/K_2 indicates the ratio of deme density between the two outbreaks. For each part of Figure 5 (A, B and C), asterisks, circles and triangle represent the relative bias of σ obtained with Endler's and LD method and the relative bias of T , respectively.

DISCUSSION

This study was carried out on a secondary contact zone between two invading European populations of *D. v. virgifera*, to obtain estimates of dispersal distances for this species. Studies of the population genetic structure of the contact zone and of the spatial variation of microsatellite allelic frequency and linkage disequilibrium generated two

estimates of the standard deviation of parent-offspring distances for *D. v. virgifera*. For both methods, we obtained ML estimates between 18 and 22 km/generation^{1/2}, with support limits roughly bounded by 16 and 26 km/generation^{1/2}. Assuming a Gaussian dispersal kernel, this standard deviation corresponds to a mean distance of about 13 to 21 km/generation (the mean of a folded Gaussian distribution is $\sigma\sqrt{2/\pi}$). We also obtained an estimate of the time since contact between the two invasive outbreaks of Northern Italy and Central and South Eastern Europe of about 3 years (bounded by 2.33 and 4.08 years). This is in fair agreement with historical information that dated the contact 4 years before sampling.

Validity of the estimate

The methods for estimating dispersal used here are based on the following assumptions: genotypic frequencies result from a pure migration process, with migration amounting to a diffusion process. As a consequence, genetic drift is assumed to be negligible, and loci are assumed to be neutral and independent. Finally, population size is assumed to remain constant over space.

Microsatellite loci are expected to be neutral ((Ellegren 2004), but see (Sutherland & Richards 1995; Watkins *et al.* 1995)), but it is possible that they are linked to genomic regions subject to selection (Charlesworth *et al.* 1993; Zhong *et al.* 1994). Several lines of evidence strongly suggest that selection was not involved in shaping the clines described here: selection on specific genomic regions may affect the various loci differently and create heterogeneity in cline slopes and centers. All the clines in this study were concordant, coincident, smooth sigmoid clines, whereas stepped clines would be expected in case of strong selection on multiple loci (Barton & Gale 1993). Moreover, selection against hybrids is unlikely because both outbreaks (NW Italy and CSE Europe) began too recently (founded in 2000 and 1992,

respectively) and their isolation time was short (about 20 generations) (Edmands 2002; Orr & Turelli 2001).

Long-distance dispersal is frequent in many species (Lockwood *et al.* 2007). For instance, Kim and Sappington (2004) found that recurrent large-scale migrations between populations of the boll weevil, *Anthonomus grandis* took place in North America. More recently, several authors have shown that *D. v. virgifera* can also disperse over long distances (Ciosi *et al.* 2011; Miller *et al.* 2005). Consequently, dispersal kernels are often not diffusive, instead having heavy tails and positive kurtoses (Liebhold & Tobin 2008; Lockwood *et al.* 2007). We simulated genotypic data resulting from various kernels (mixture of Gaussian or exponential power laws), with light or heavy tails, and with a kurtosis of between -1 and 30. If dispersal is not too leptokurtic, simulations showed that the true value of σ is closed to that estimated under diffusion hypothesis. Simulations thus indicated that dispersal estimates were not very sensitive to the type of dispersal kernel and kurtosis, with a relative bias of less than 25%. These results are consistent with previous findings showing that kurtosis had little effect on the estimation of σ with comparable methods (Rieux *et al.* 2013). If long distance dispersal is very rare, then the dispersal parameter estimated here is the same as the diffusion parameter. However, if there is enough long-range movement, the expansion dynamics can significantly change. The co-occurrence of both local diffusion and long-distance dispersal – a phenomenon called stratified dispersal (Shigesada *et al.* 1995) – can lead to a faster expansion than expected under local diffusion alone. With stratified dispersal, remote colonies are stochastically sent outside the expanding population. They eventually merge with the main population and contribute to the advance of the front (Shigesada *et al.* 1995).

In very large study areas, such as that studied here (more than 800 km wide), there is probably environmental heterogeneity rendering population density heterogeneous over space. Moreover, the two invasive populations of *D. v. virgifera* are of different ages, the NW

Italian outbreak occurring about eight years after the CSE European outbreak (Miller *et al.* 2005), and this may account for differences in density. A difference in density between the two colonizing populations may also result from differences in anthropogenic selection intensity (insecticide treatments and crop rotations). When densities are equal, the slope of the cline is a function of dispersal. Stepwise differences in densities across the contact zone would be expected to lead to a steeper slope of the cline, because of the stronger contribution to gene flow of the population with the highest density (Barton 1979). We thus simulated a worst-case scenario in which there were constant differences in population size between the two populations. Simulations showed that these differences had little influence on dispersal estimates, as biases were always below 20 %, even for a density ratio of 500 %.

We assumed a recombination rate of 0.5 between all microsatellites. This seems reasonable, because all linkage disequilibrium values were zero outside the contact zone. However, this assumption may be false for some microsatellites, resulting in a mean recombination rate slightly below 0.5. As $\sigma^2 \propto r$ with the LD method (Barton & Gale 1993), this may result in an overestimation of σ when considering $r = 0.5$. Note that for a wide range of recombination rates, between 0.3 and 0.5, the assumption of complete independence between loci would result in a limited bias of the estimation of σ , of less than 25%.

The methodologies based on Endler's method and LD (Barton & Gale 1993; Endler 1977) are both based on the assumption that drift has a limited effect. We were confident from the outset that this was the case for *D. v. virgifera* in Northern Italy, as very large population sizes are often reported for invasive outbreaks (Borioni *et al.* 2006; Szalai *et al.* 2011) and previous studies have shown negligible genetic differentiation between samples within outbreaks (Bermond *et al.* 2012; Ciosi *et al.* 2008). We are also confident, with hindsight, that drift had little effect on the genetic structure of *D. v. virgifera* in the contact zone because (i) we found concordant, coincident and parallel clines, whereas drift would be

expected to create some heterogeneity in cline shape (width, center and angle) between loci, and (ii) 86 % of the deviance was explained by the M_1 model, whereas drift would be expected to generate considerable background noise around expected allele frequencies, resulting in a poor fit.

Comparison with other dispersal estimates for *D. v. virgifera*

D. v. virgifera is one of the worst pests of maize worldwide. The control strategies used against it, including the use of GM maize producing *Bacillus thuringiensis* toxins and refuge zones, are based on the hypothesis that *D. v. virgifera* disperses effectively. The combination of different selective pressures (i.e. various pesticides) over space, to increase the degree of treatment heterogeneity (REX 2013) against *D. v. virgifera*, requires a rough knowledge of the distances over which *D. v. virgifera* can disperse. *D. v. virgifera* is invading Europe, and the invasive outbreaks are expanding geographically at a speed dependent on the dispersal capacities of *D. v. virgifera*. Risk analyses and simulations of various scenarios of human defense against *D. v. virgifera* spread (Baufeld &ENZIAN 2001; Carrasco *et al.* 2010) require the parameterization of *D. v. virgifera* dispersal capacities. This is why many previous studies have focused on the dispersal of *D. v. virgifera* in the USA (Coats *et al.* 1986; Gray *et al.* 2009) and in Europe (Baufeld &ENZIAN 2001; Ciosi *et al.* 2011).

Carrasco *et al.* (2010) performed capture-mark-recapture experiments and showed that half the individuals tested in the field moved between 117 and 425 m/day (with a median of 117 to 188 m/day). Given that the adult lifespan of *D. v. virgifera* is three months (Branson & Johnson 1973; Hill 1975), these daily distances correspond to between 11 and 38 km/generation (or a median of 11 to 17 km/generation), consistent with our estimates. Baufeld andENZIAN (2001) fitted a model to the expansion of the CSE Europe population since the start of this invasion near Belgrade in 1992 and estimated the expansion rate at 60 to

100 km/year. Assuming a diffusion approximation with a Gaussian dispersal kernel, Fisher's wave of advance model suggests that the expansion rate of a colonizing population with an intrinsic growth rate r is $C = 2\sigma r^{1/2}$ (Shigesada & Kawasaki 1997). In 2011, Szalai *et al.* (2011) showed that the mean annual growth rate of the CSE European outbreak was about four, giving a value of σ between 15 and 25 km.year^{-1/2}. Our estimates are consistent with these findings.

The temporary admixture zone as an opportunity for estimating dispersal

Temporary and vanishing admixture zones following secondary contact are rarely used to estimate dispersal. One recent study made use of the relationship found by Endler (1977) between dispersal and the slope of neutral clines to estimate dispersal distances in secondary contact zones between two colonizing populations of the banana pest *Mycosphaerella fijiensis* (Rieux *et al.* 2013). Rieux *et al.* (2013) fitted a scaled logistic model to their microsatellite frequencies samples at two different times. The slope difference between the two periods was used to estimate σ through Endler's equation. In our case, we did not use a temporal sampling because 1) we had a good idea of the time since contact (which was not the case of Rieux *et al.* (2013)) and 2) we could estimate σ from linkage disequilibrium without knowing T with a single year of sampling. Most studies using linkage disequilibrium in a secondary contact zone to estimate σ focus on hybrid zones maintained by selection (e.g. Barton 1982; Barton & Gale 1993; Gay *et al.* 2008; Singhal & Moritz 2012). To our knowledge, this study is also the first to use a spatial fit of the multilocus linkage disequilibrium in a maximum likelihood framework. This fit was computationally tractable, because all samples were at Hardy-Weinberg equilibrium, such that it was not necessary to estimate the between-genome within-gene association parameter $\kappa_{1,1}$ (Barton 2000) along with $\kappa_{0,2}$.

The methodology used here in the context of secondary contact between invasive populations has several advantages: i) it provides estimates of effective dispersal distance subject to only weak bias; ii) it provides estimates that are robust to departure from several hypotheses; iii) it provides narrow support limits, indicating a low level of uncertainty on the estimates; iv) it is cost- and time-effective as it requires the genotyping of only a few polymorphic genetic (diallelic) markers in a few hundred individuals sampled only once; v) there is no need to know precisely the time for which the populations have been in contact, and this factor can be estimated. The main constraint associated with this method is that the location of the secondary contact zone must be roughly known before sampling, to ensure accurate estimations of cline width.

This methodology seems to be suitable for studies of many cases of biological invasions *in natura*. Many invasions originate from multiple introductions in disconnected areas (Barney 2006; Dlugosch & Parker 2008; Hufbauer & Sforza 2008; Lombaert *et al.* 2010; Urbanelli *et al.* 2000) and many lead to secondary contact (e.g. Kolbe *et al.* 2004; Lavergne & Molofsky 2007; Lombaert *et al.* 2010; Lucek *et al.* 2010; Rieux *et al.* 2013). Temporary and vanishing allelic frequency clines are thus probably very frequent *in natura*, particularly in the context of biological invasions, and could be used to estimate effective dispersal distance accurately in many species provided time since contact is not too large.

ACKNOWLEDGMENTS

We thank Laura Maïche and Stefan Toepfer for sampling campaigns and technical assistance. We also thank Thomas Lenormand for initial ideas and Arnaud Estoup, Eric Lombaert, Etienne Klein, and Stefan Toepfer for their contributions and assistance with data analysis and discussion. This work was supported by grants from the *Agence Nationale de la Recherche* (ANR-06-BDIV-008-01 and ANR-09-BLAN-0145-01) and Agropolis Fondation (Montpellier, France) under the reference ID “BIOFIS” 1001-001.

REFERENCES

- Andow D, Kareiva P, Levin SA, Okubo A (1990) Spread of invading organisms. *Landscape Ecology* **4**, 177-188.
- Arim M, Abades SR, Neill PE, Lima M, Marquet PA (2006) Spread dynamics of invasive species. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 374-378.
- Barney J (2006) North American History of Two Invasive Plant Species: Phytogeographic Distribution, Dispersal Vectors, and Multiple Introductions. *Biological Invasions* **8**, 703-717.
- Barton NH (1979) The dynamics of hybrid zones. *Heredity* **43**, 341-359.
- Barton NH (1982) The Structure of the Hybrid Zone in *Uroderma-Bilobatum* (Chiroptera, Phyllostomatidae). *Evolution* **36**, 863-866.
- Barton NH (2000) Estimating multilocus linkage disequilibria. *Heredity* **84**, 373-389.
- Barton NH, Gale KS (1993) Genetic analysis of hybrid zone. In: *Hybrid zones and the evolutionary process* (ed. Harison RG), pp. 13-45. Oxford University Press, New York.
- Baufeld P, Enzian S (2001) Simulations model for spreading scenarios of western corn rootworm (*Diabrotica virgifera virgifera*) in case of Germany. *IWGO Newsletter* **22**, 14-15.
- Bermond G, Ciosi M, Lombaert E, *et al.* (2012) Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe. *PloS one* **7**, e50129.
- Bierne N, Borsa P, Daguin C, *et al.* (2003) Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M-galloprovincialis*. *Molecular Ecology* **12**, 447-461.
- Boriani M, Agosti M, Kiss J, Edwards CR (2006) Sustainable management of the western corn rootworm, *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae), in infested areas: experiences in Italy, Hungary and the USA. *EPPO Bulletin* **36**, 531-537.
- Bossart JL, Pashley Prowell D (1998) Genetic estimates of population structure and gene flow: Limitations, lessons and new directions. *Trends in Ecology & Evolution* **13**, 202-206.
- Bossdorf O, Auge H, Lafuma L, *et al.* (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**, 1-11.
- Branson TF, Johnson RD (1973) Adult western corn rootworms: oviposition, fecundity, and longevity in the laboratory. *Journal of Economic Entomology* **66**, 417-418.
- Branson TF, Krysan JL (1981) Feeding and oviposition behavior and life cycle strategies of *Diabrotica*: an evolutionary view with implications for pest management. *Environmental Entomology* **10**, 826-831.
- Broquet T, Petit EJ (2009) Molecular Estimation of Dispersal for Ecology and Population Genetics. *Annual Review of Ecology Evolution and Systematics*, **40**, 193-216.
- Burnham Kenneth P., R. AD (2010) *Model selection and multimodel inference: a practical information-theoretic approach*, Second edn. Springer, New York, NY.
- Carrasco LR, Harwood TD, Toepfer S, *et al.* (2010) Dispersal kernels of the invasive alien western corn rootworm and the effectiveness of buffer zones in eradication programmes in Europe. *Annals of Applied Biology* **156**, 63-77.
- Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289-1303.

- Ciosi M, Miller NJ, Kim KS, *et al.* (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology* **17**, 3614-3627.
- Ciosi M, Miller NJ, Toepfer S, Estoup A, Guillemaud T (2011) Stratified dispersal and increasing genetic variation during the invasion of Central Europe by the western corn rootworm, *Diabrotica virgifera virgifera*. *Evolutionary Applications* **4**, 54-70.
- Coats SA, Tollefson JJ, Mutchmor JA (1986) Study of migratory flight in the western corn rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology* **15**, 620-625.
- Daguin C, Bonhomme F, Borsa P (2001) The zone of sympatry and hybridization of *Mytilus edulis* and *M. galloprovincialis*, as described by intron length polymorphism at locus *mac-1*. *Heredity* **86**, 342-354.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* **17**, 431-449.
- Dobzhansky T, Wright S (1943) Genetics of Natural Populations. X. Dispersion Rates in *Drosophila Pseudoobscura*. *Genetics* **28**, 304-340.
- Edmunds S (2002) Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution* **17**, 520-527.
- Edwards AWF (1972) *Likelihood* Cambridge University Press, Cambridge.
- Ellegren H (2004) Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics* **5**, 435-445.
- Endler, JA (1977) Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, New Jersey.
- Facon B, Pointier J-P, Glaubrecht M, *et al.* (2003) A molecular phylogeography approach to biological invasions of the New World by parthenogenetic Thiarid snails. *Molecular Ecology* **12**, 3027-3039.
- Fisher RA (1937) The wave of advance of advantageous genes.. *Annals of Eugenics* **7**, 355-369.
- Gay L, Crochet PA, Bell DA, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution* **62**, 2789-2806.
- Gray ME, Sappington TW, Miller NJ, Moeser J, Bohn MO (2009) Adaptation and Invasiveness of Western Corn Rootworm: Intensifying Research on a Worsening Pest. *Annual Review of Entomology* **54**, 303-321.
- Gregory P (2013) "Markov Chain Monte Carlo Simulation Using the Metropolis Algorithm" from the Wolfram Demonstrations Project:
<http://demonstrations.wolfram.com/MarkovChainMonteCarloSimulationUsingTheMetropolisAlgorithm/>
- Hill RE (1975) Mating, oviposition patterns, fecundity and longevity of the western corn rootworm. *Journal of Economic Entomology* **68**, 311-315.
- Hufbauer RA, Sforza R (2008) Multiple introductions of two invasive *Centaurea* taxa inferred from cpDNA haplotypes. *Diversity and Distributions* **14**, 252-261.
- Johnson ML, Gaines MS (1990) Evolution of Dispersal: Theoretical Models and Empirical Tests Using Birds and Mammals. *Annual Review of Ecology and Systematics* **21**, 449-480.
- Kim KS, Sappington TW (2004) Boll weevil (*Anthonomus grandis* Boheman)(Coleoptera: Curculionidae) dispersal in the southern United States: evidence from mitochondrial DNA variation. *Environmental Entomology* **33**, 457-470.
- Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist* **150**, 1-23.

- Kolbe JJ, Glor RE, Schettino LRG, *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177-181.
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3883-3888.
- Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling Survival and Testing Biological Hypotheses Using Marked Animals - a Unified Approach with Case-Studies. *Ecological Monographs* **62**, 67-118.
- Lenormand T, Guillemaud T, Bourguet D, Raymond M (1998) Evaluating gene flow using selected markers: A case study. *Genetics* **149**, 1383-1392.
- Lenormand T, Raymond M (2000) Analysis of clines with variable selection and variable migration. *American Naturalist* **155**, 70-82.
- Liebholt AM, Tobin PC (2008) Population Ecology of Insect Invasions and Their Management*. *Annu. Rev. Entomol.* **53**, 387-408.
- Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution* **20**, 223-228.
- Lockwood JL, Hoopes MF, Marchetti MP (2007) *Invasion ecology* Blackwell, Malden, MA.
- Lodge DM, Shrader-Frechette K (2003) Nonindigenous species: Ecological explanation, environmental ethics, and public policy. *Conservation Biology* **17**, 31-37.
- Lombaert E, Guillemaud T, Cornuet JM, *et al.* (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PloS one* **5**, e9743.
- Lubina JA, Simon AL (1988) The Spread of a Reinvading Species: Range Expansion in the California Sea Otter. *The American Naturalist* **131**, 526-543.
- Lucek KAY, Roy D, Bezault E, Sivasundar A, Seehausen OLE (2010) Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Molecular Ecology* **19**, 3995-4011.
- Mack RN, Simberloff D, Lonsdale WM, *et al.* (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689-710.
- Mallet J, Barton N, Lamas G, *et al.* (1990) Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* **124**, 921-936.
- McKinney ML, Lockwood JL (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* **14**, 450-453.
- Miller N, Estoup A, Toepfer S, *et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**, 992.
- Miller NJ, Ciosi M, Sappington TW, *et al.* (2007) Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation tolerance. *Journal of Applied Entomology* **131**, 378-385.
- Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution* **19**, 18-24.
- Orr HA, Turelli M (2001) The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085-1094.
- Peacock MM (1997) Determining natal dispersal patterns in a population of North American pikas (*Ochotona princeps*) using direct mark-resight and indirect genetic methods. *Behavioral Ecology* **8**, 340-350.
- Phillips BL, Brown GP, Webb JK, Shine R (2006) Invasion and the evolution of speed in toads. *Nature* **439**, 803-803.

- Pimentel D, McNair S, Janecka J, *et al.* (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture Ecosystems & Environment* **84**, 1-20.
- R Development Core Team (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Raymond M, Rousset F (1995) Genepop (version. 1.2), a population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.
- Rehage JS, Sih A (2004) Dispersal behavior, boldness, and the link to invasiveness: a comparison of four *Gambusia* species. *Biological Invasions* **6**, 379-391.
- REX (2013) Heterogeneity of selection and the evolution of resistance. *Trends in Ecology & Evolution* **28**, 110-118.
- Rieux A, Lenormand T, Carlier J, Lapeyre de Bellaire L, Ravigné V (2013) Using neutral cline decay to estimate contemporary dispersal: a generic tool and its application to a major crop pathogen. *Ecology letters*, **16**, 721-730.
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**, 454-464.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.
- Rousset F (2001) Genetic approaches to the estimation of dispersal rates. . In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), pp. 18-28, Oxford University Press, New York, NY.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.
- Ruiz GM, Rawlings TK, Dobbs FC, *et al.* (2000) Global spread of microorganisms by ships - Ballast water discharged from vessels harbours a cocktail of potential pathogens. *Nature* **408**, 49-50.
- Sax DF, Stachowicz JJ, Gaines SD (2005) *Species Invasions: Insights into Ecology, Evolution, and Biogeography* Sinauer Associates, Sunderland, MA, U.S.A.
- Shigesada N, Kawasaki K (1997) *Biological invasions : theory and practice* Oxford University Press, Oxford [etc.].
- Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. *American Naturalist* **146**, 229-251.
- Singhal S, Moritz C (2012) Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution* **66**, 1474-1489.
- Sivcev I, Manojlovic B, Krnjajic S, *et al.* (1994) Distribution and harmfulness of *Diabrotica virgifera* LeConte (Coleoptera, Chrysomelidae), a new maize pest in Yugoslavia. *Zastita Bilja* **45**, 19-26.
- Skellam JG (1951) Random Dispersal in Theoretical Populations. *Biometrika* **38**, 196-218.
- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 787-792.
- Sokal RR, Rohlf FJ (1995) *Biometry. The Principles and Practice of Statistics in Biological Research.*, Third Edition edn. W.H. Freeman and Company, New York.
- Spencer JL, Hibbard BE, Moeser J, Onstad DW (2009) Behaviour and ecology of the western corn rootworm (*Diabrotica virgifera virgifera* LeConte). *Agricultural and Forest Entomology* **11**, 9-27.
- Sutherland GR, Richards RI (1995) Simple tandem DNA repeats and human genetic disease. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3636-3641.

- Szalai M, Papp Komáromi J, Bažok R, *et al.* (2011) Generational growth rate estimates of *Diabrotica virgifera virgifera* populations (Coleoptera: Chrysomelidae). *Journal of Pest Science* **84**, 133-142.
- Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in southern Poland. *Evolution* **40**, 1141-1159.
- Urbanelli S, Bellini R, Carrieri M, Sallicandro P, Celli G (2000) Population structure of *Aedes albopictus* (Skuse): the mosquito which is colonizing Mediterranean countries. *Heredity* **84**, 331-337.
- Watkins WS, Bamshad M, Jorde LB (1995) Population genetics of trinucleotide repeat polymorphisms. *Human Molecular Genetics* **4**, 1485-1491.
- Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution* **24**, 136-144.
- Wolfram Research Inc (2010) Mathematica Edition: Version 8.0. Wolfram Research, Inc., Champaign, Illinois.
- Wright S (1943) Isolation by Distance. *Genetics* **28**, 114.
- Zhong N, Ye LL, Dobkin C, Brown WT (1994) Fragile-X founder chromosome effects-linkage disequilibrium or microsatellite heterogeneity. *American Journal of Medical Genetics* **51**, 405-411.

DATA ACCESSIBILITY

Sample locations, microsatellite data, Mathematica notebook and R script: DRYAD entry doi:10.5061/dryad.10kd4.

AUTHOR CONTRIBUTIONS

This work forms part of GB's PhD thesis on the invasion of Europe by *Diabrotica virgifera virgifera*, supervised by TG. GB and TG conceived and designed the experiments. Sampling and molecular experiments were performed by GB and SM, and GB and AB, respectively. All coauthors were involved in data analysis and the writing of the paper.

SUPPORTING INFORMATION

Table S1: Construction of dispersal distributions from mixtures of Gaussian or exponential power family kernels. All kernels yield the same variance ($\sigma^2 = 400$) but different values of kurtosis γ . The probability density function of the Gaussian mixture is

$$f(x) = A \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}} + B \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_2^2}}, \text{ with } \begin{cases} A = 1 \\ B = 0 \end{cases} \text{ and } \begin{cases} A = 0 \\ B = 1 \end{cases} \text{ with probabilities } \omega \text{ and } 1-\omega, \text{ respectively.}$$

With such a mixture, $\sigma^2 = \omega\sigma_1^2 + (1-\omega)\sigma_2^2$. The probability density function

$$\text{of the exponential power kernel is } f(x) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{x}{\alpha}\right)^\beta}.$$

Gaussian mixture	Type	σ_1^2	σ_2^2	ω	γ
	Gaussian	400	0	1	0
	Light-tailed	300	936.9	0.843	1
	Heavy-tailed	200	1200	0.8	3
	Heavy-tailed	100	2183	0.856	10
	Heavy-tailed	50	4980	0.929	30
Exponential power	Type	β	A	γ	
	Light-tailed	6	35.4	-1	
	Light-tailed	2.8	32.2	-0.5	
	Gaussian	2	28.3	0	
	Heavy-tailed	1.4	21.9	1	
	Heavy-tailed (Laplace)	1	14.1	3	

VI. – CHAPITRE III : ABSENCE D'IMPACT DE L'HYBRIDATION *IN* *NATURA* AU COURS DE L'INVASION EUROPÉENNE DE *DIABROTICA* *VIRGIFERA*

Ce chapitre repose sur une étude dans laquelle nous cherchons à évaluer l'impact de l'hybridation *in natura* entre les deux foyers envahissants de *Dv* en Europe sur des caractères liés à la valeur sélective des individus. Cette étude se rapporte à l'**article n°3**.

Article n°3: **Bermond, G.**, Cavigliasso, F., Spencer J.L., Mallez, S., Guillemaud, T. No detectable effect of admixture between two European invading outbreaks of *Diabrotica virgifera virgifera in natura (in prep)*.

Contexte de l'étude, objectif visé et questions posées

Au cours des invasions biologiques, les pertes de diversité génétique associées aux étapes précoces de l'invasion (Sakai *et al.*, 2001; Wares *et al.*, 2005) sont parfois compensées dans l'aire envahie par des évènements d'hybridation entre populations génétiquement différenciées. De tels processus sont souvent considérés comme des facteurs stimulants des invasions (Drake, 2006; Ellstrand, Schierenbeck, 2000), favorisant l'adaptation à de nouveaux environnements par le biais de différents mécanismes tels que l'hétérosis (augmentation de la valeur sélective des individus hybrides par rapport aux parents, Dobzhansky, Gowen, 1952; Shull, 1914).

Bien que l'hybridation soit une idée séduisante pour contrebalancer le défi adaptatif lié aux introductions — la perte de diversité génétique générée par les goulots d'étranglement peut se traduire par une réduction du potentiel adaptatif de la population introduite, voir **section I.1.3** — et qu'elle soit considérée comme un stimulus des invasions, elle est de plus en plus critiquée et remise en question (Chapple *et al.*, 2013). Par ailleurs, l'hybridation entre populations génétiquement différenciées peut également aboutir à des phénomènes inverses telle que la dépression hybride (diminution de la valeur sélective des individus hybrides par rapport aux parents) pouvant s'expliquer par différents mécanismes (voir **section I.2.2.a**, Schierup, Christiansen, 1996; Tallmon *et al.*, 2004; Turelli, Orr, 2000). On peut donc s'attendre à avoir des conséquences diverses à l'issue d'un évènement d'hybridation.

Les deux études présentées précédemment dans les **Chapitres I** (La Vénétie, une zone hybride entre deux foyers envahissants européens de *Diabrotica virgifera virgifera*, localisée en Italie du Nord) et **II** (Estimation de la dispersion de *Diabrotica virgifera virgifera* en Europe) ont permis de mettre en évidence (grâce à l'utilisation de marqueurs génétiques neutres) une zone hybride localisée en Italie du Nord (en Vénétie) entre les deux foyers envahissants majeurs de *Diabrotica virgifera* en Europe. La forme des clines de fréquences alléliques observés indique une absence d'hétérogénéité entre loci et suggère ainsi qu'au sein de cette zone hybride les marqueurs utilisés ne sont pas liés avec des zones du génome sous sélection. Cependant à ce stade nous ne pouvons pas encore déterminer si cette zone hybride résulte en une zone hybride neutre ou une zone de tension (les attendus concernant la forme des clines et le nombre de paramètres estimés sont les mêmes pour ces deux types de zones).

L'objectif de cette nouvelle étude est donc de déterminer si des différences significatives de valeur sélective existent entre individus hybrides et parentaux. Pour cela nous avons testé des hypothèses selon lesquelles l'hybridation n'a pas d'impact sur des

caractères liés à la valeur sélective des individus et sur le système d'appariement des individus sexuellement matures. Les questions sous-jacentes sont les suivantes :

- 1) Quel est l'impact de l'hybridation sur des caractères phénotypiques lié à la valeur sélective des individus ?
- 2) Le statut hybride des individus leur confère-ils un avantage quelconque pour l'accès à la reproduction?
- 3) La reproduction se fait-elle au hasard entre individus hybrides et non hybrides

Matériel utilisé et méthodes employées

L'échantillonnage de cette étude a été réalisé au même moment que celui de l'étude correspondant à l'**article n°2** et le temps imparti pour les réaliser était assez limité. J'ai donc décidé, pour atteindre l'objectif visé de cette étude, de choisir des caractères facilement mesurables et ne demandant pas beaucoup d'investissement sur le terrain. La survie à jeun (sans apport de nourriture et d'eau) entre la date d'échantillonnage et la date de mort et le taux d'accouplement des individus (définissable sur des individus morts, au laboratoire) ont ainsi été sélectionnés. Mais comment comparer des populations parentales et hybrides qui évoluent dans des environnements potentiellement différents et soumis à des pressions de sélection très différentes (traitements insecticides, rotation des cultures) ? Pour contourner cette difficulté, 3 champs, Tv11, Tv19 et Vi9 situés au cœur de la zone hybride (Figure 13) et proches géographiquement, ont été sélectionnés et échantillonnés de manière à obtenir la plus grande variabilité possible du taux d'hybridation des individus qui s'y trouvent. En effet une représentation de tous les types génotypiques (parental Ouest, parental Est et hybride) est indispensable à la comparaison des populations parentales et hybrides. Une analyse d'assignation génétique des échantillons vénitiens (collectés en 2011) aux populations

parentales (Ouest ou Est) a donc été réalisée pour cela. Deux cents individus par échantillon ont finalement été prélevés en 2012.

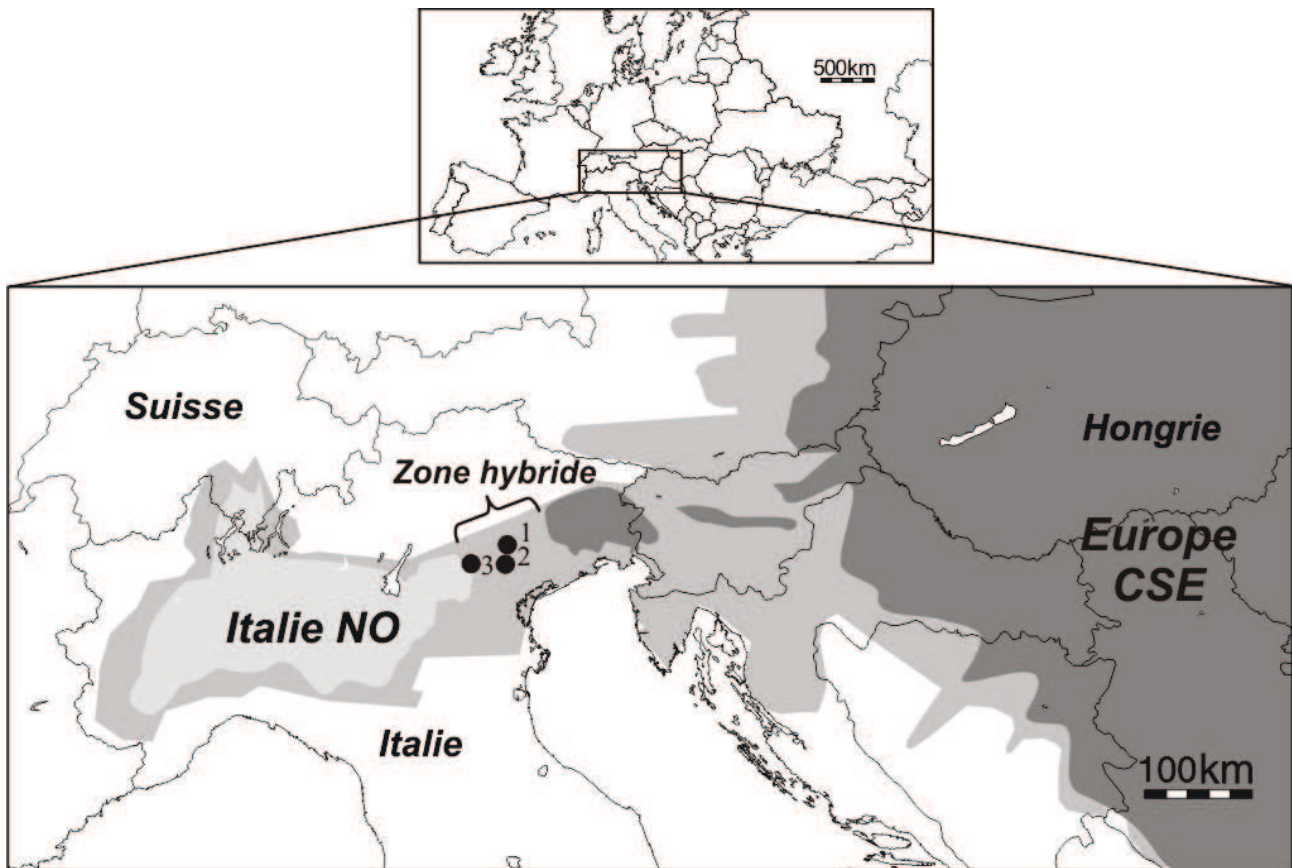


Figure 13 : Localisation des deux principaux foyers européens envahissants de *Diabrotica virgifera virgifera* (*Dvv*) en Italie du Nord-Ouest et en Europe Centrale et du Sud-Est (Europe CSE) et localisation des échantillons de *Dvv* collectés en 2012 dans la zone hybride. Les nombres se réfèrent au nom des échantillons : 1 = Tv11, 2 = Tv19, 3 = Vi9. L'aire de répartition de *Dvv* avant le contact entre les populations des foyers Ouest et Est (en 2007) est représentée en gris clair (Italie NO) et gris foncé (Europe CSE). L'aire de répartition après le contact (2012) est représentée en gris intermédiaire.

Dans un premier temps, nous avons disséqué les 600 individus afin de déterminer leur sexe. Puis, nous avons extrait et amplifié l'ADN des individus et du sperme contenu dans les spermathèques des femelles afin de les génotyper à 13 loci microsatellites. Cela a nécessité la mise en place d'un protocole d'isolement et de génotypage du sperme, ce qui n'avait jamais été fait auparavant chez cette espèce.

Une fois les génotypes connus, nous avons cherché à déterminer le type génotypique des individus (hybrides ou parents) et leur taux d'hybridation. Dans un premier temps, nous avons inféré la structure génétique de chacun des 3 échantillons vénitiens en utilisant la

méthode Bayésienne implémentée dans le logiciel STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000) avec un modèle d'admixture. Pour cela, nous avons utilisé plusieurs échantillons d'Italie NO et d'Europe CSE représentatifs des populations parentales (Bermond *et al.*, 2012). Cette première analyse a permis de calculer un coefficient Q (le coefficient de « coancestry » de STRUCTURE) variant de 0 à 1 et correspondant à la proportion du génome d'un individu qui ici appartient à la population parentale Ouest. Nous avons ensuite défini 5 classes non-chevauchantes du coefficient Q (classe 1 = [0-0.2] ; classe 2 = [0.2-0.4] ; classe 3 = [0.4-0.6] ; classe 4 = [0.6-0.8] ; classe 5 = [0.8-1]) afin de caractériser le statut hybride de chaque individu. Les classes 1 et 5 correspondent respectivement aux génotypes parentaux Est et Ouest ; les classes 2 et 4 correspondent aux génotypes « backcross » et enfin la classe 3 correspond aux génotypes F_1 et hybrides de 2^e, 3^e, 4^e et 5^e génération (le contact secondaire a eu lieu 5 ans auparavant). Notons qu'il existe probablement d'autres classes dans la nature et que la classification proposée ici est forcément fautive. Par exemple, le croisement entre un individu backcross-Est et un individu backcross-Ouest devrait résulter en un hybride appartenant à la classe 3 (précédemment présentée) or cet hybride ne correspond pas à la définition d'un véritable F_1 . Cette classification est donc une approximation, cependant compte tenu de la mise en évidence de fortes pentes de clines, les classes hybrides réelles doivent probablement correspondre aux 5 présentées ici. Ces classes sont désignées par la suite comme les classes de STRUCTURE. Enfin, nous avons calculé un taux d'hybridation (TH) à partir du coefficient Q , variant de 0 à 1 et défini tel que : $TH=2Q$ (pour $Q \leq 0.5$) et $TH=2(1-Q)$ (pour $Q > 0.5$). Dans un second temps, nous avons utilisé une approche Bayésienne de regroupement génétique implémentée dans le logiciel NewHybrid 1.1 (Anderson, Thompson, 2002). Ce programme utilise des simulations basées sur la méthode de Monte Carlo par Chaines de Markov pour calculer la probabilité *a posteriori* que chaque individu d'un échantillon tombe dans une des six classes génétiques prédéfinies : deux classes parentales

(Pure-Ouest et Pure-Est), deux classes hybrides F_1 et F_2 et enfin deux classes de croisement « backcross » (backcross-Ouest et backcross-Est, correspondant au descendant d'un F_1 et d'un individu Pure-Ouest et Pure-Est, respectivement. Ces classes sont désignées par la suite comme les classes de NewHybrid. Le statut hybride de chaque individu a été ainsi caractérisé par deux variables qualitatives obtenues par deux méthodes de regroupements génétiques (les classes de STRUCTURE et de NewHybrid) et une variable quantitative (*TH*), intimement liée aux classes de STRUCTURE.

Enfin nous avons testé l'hypothèse d'un lien entre le statut hybride des individus et leur probabilité d'accouplement ou leur temps de survie par l'intermédiaire de plusieurs tests statistiques tels que des tests exacts de Fisher, des modèles linéaires généralisés mixtes et des modèles de Cox. L'hypothèse d'un accouplement aléatoire entre individus appartenant à des types génotypiques différents a également été testée grâce à des tests exacts de Fisher.

Résultats

Les résultats obtenus au cours de cette étude sont remarquables par leur simplicité : aucun effet du statut hybride n'a été constaté sur les caractères liés à la valeur sélective des individus, quels que soient la mesure du statut et le caractère concerné (survie à jeun et taux d'accouplement). Ceci suggère donc que le statut hybride ne semble pas conférer d'avantages pour l'accès à la reproduction. Par ailleurs, l'hypothèse d'un accouplement aléatoire entre individus appartenant à des types génotypiques différents n'a pas été rejetée.

1 **No detectable effect of admixture between two European invading outbreaks of**
2 *Diabrotica virgifera virgifera in natura*

3
4
5
6 Gérald Bermond^{1, 2, 3}, Fanny Cavigliasso^{1, 2, 3}, Sophie Mallez^{1, 2, 3}, Joseph L. Spencer⁴ and
7 Thomas Guillemaud^{1, 2, 3}

8
9 ¹INRA, UMR 1355 Institut Sophia Agrobiotech, F-06903 Sophia Antipolis, France

10 ²Université de Nice Sophia Antipolis, UMR Institut Sophia Agrobiotech, F-06903 Sophia
11 Antipolis, France

12 ³CNRS, UMR 7254 Institut Sophia Agrobiotech, F-06903 Sophia Antipolis, France

13 ⁴Section of Ecological Entomology, Division of Biodiversity and Ecological Entomology,
14 Illinois Natural History Survey, 1816 South Oak Street, Champaign, IL 61820-6960, U.S.A

15

16 **ABSTRACT**

17 The reduction of genetic diversity that often occurs during introduction of future
18 invasive populations can decrease their adaptive potential. Multiple introductions followed by
19 admixture events may increase the genetic diversity and therefore facilitate invasion. In this
20 study, we challenged the hypothesis that admixture may have had a positive impact in the
21 context of the European invasion of the western corn rootworm (WCR), *Diabrotica virgifera*
22 *virgifera*. This beetle was introduced several times in Europe since the 1980's from the USA.
23 The multiple introductions of this major pest of cultivated corn led to the formation of two
24 major outbreaks that eventually merged into a secondary contact zone where insects from both
25 outbreaks interbreed. We collected about 600 insects from this contact zone and genotyped
26 them using 13 microsatellite markers. Three types of information were obtained from the
27 collected individuals: (i) their starving survival; (ii) their admixed status, determined through a
28 Bayesian method of genetic clustering; (iii) and their mating success, studied *via* the detection,
29 isolation and genotyping of sperm in female spermathecae. In this study, we statistically tested
30 for a link between the individual admixed status and mating success on the one hand, and the
31 individual starving survival on the other hand. Our results do not reveal any impact of the
32 admixed status on the mating success or the survival of individuals. We conclude that this
33 particular zone of admixture is probably neutral and will thus vanish through time under the
34 effect of WCR dispersal.

35

36 **Key words:** western corn rootworm, fitness, sperm, GLMMs, Cox's models, rate of admixture

37

38

39 INTRODUCTION

40

41 Despite the increase of the number of invasive species, only a small fraction of
42 emigration events leads to successful introductions, settlements and proliferations into new
43 territories (Williamson, Fitter, 1996). To better control or simply avoid the colonization of new
44 territories by exotic species, biologists have tried for sixty years to investigate and understand
45 the complex ecological processes involved in invasions (Elton, 1958; Gurevitch *et al.*, 2011).
46 According to Fauvergue *et al.* (2012), three types of causes can explain the success or failure of
47 introductions. First, abiotic and biotic environmental conditions of the new area may be
48 unfavorable for the introduced species (*e.g.* Moyle, Light, 1996). Explanatory causes can be
49 also demographic: bottlenecks involved during introductions generate other processes such as
50 demographic stochasticity (Lande, 1993) and Allee effects (1949) that can lead to the extinction
51 of the introduced populations. Finally genetic causes can also be involved: population
52 introductions are often associated with founder events that are generally associated to a loss of
53 genetic diversity that can lead to a reduction of the adaptive potential of populations (Willi *et*
54 *al.*, 2006). The consequences of this reduction may be important for the introduced populations
55 since the invaded environment represents new selection pressures that will act on a limited
56 range of genetic variability. Bottlenecks can also induce an increase of consanguinity in
57 introduced populations, generate inbreeding depression and even the fixation of deleterious
58 alleles. Such processes will increase the probability of extinction of introduced populations
59 (Fauvergue *et al.*, 2012).

60 During biological invasions, the loss of genetic diversity associated with bottlenecks
61 (Sakai *et al.*, 2001; Wares *et al.*, 2005) is sometimes offset in the invaded area by intraspecific
62 hybridization (*i.e.* admixture) between populations that have been introduced from several

63 source populations of the native area (Verhoeven *et al.*, 2011). These multiple introductions
64 from different population sources are common (Dlugosch, Parker, 2008) and sometimes lead to
65 the formation of a single outbreak in which admixture occurs (Genton *et al.*, 2005). Such
66 introductions can also lead to the formation of several outbreaks, geographically disconnected
67 (Facon *et al.*, 2008; Kolbe *et al.*, 2008; Lavergne, Molofsky, 2007; Rosenthal *et al.*, 2008;
68 Taylor, Keller, 2007) and genetically differentiated (Ciosi *et al.*, 2008). Thereafter, the
69 expansion of these outbreaks can then lead to a secondary contact (*e.g.* Rieux *et al.*, 2013)
70 where populations meet, cross and form an admixed zone (*e.g.* Bermond *et al.*, 2012).

71 Such admixture can have beneficial consequences for the invasive populations
72 (Ellstrand, Schierenbeck, 2000; Verhoeven *et al.*, 2011) that promote adaptation to new
73 environments (*e.g.* Facon *et al.*, 2008; Roman, 2006). First, the admixture can increase the
74 extent of genetic variability displayed by populations (Kolbe *et al.*, 2008) and allow selection to
75 operate on a broader range of genetic variability. Second, recombination between genotypes
76 from genetically differentiated populations can create new genotypes and increase the
77 phenotypic range on which the natural selection will act. Finally, admixture can also lead to
78 heterosis – *i.e.* a larger fitness of the admixed individuals than that of the parental individuals
79 (Dobzhansky, Gowen, 1952; Shull, 1914). This may occur when the parental introduced
80 populations have fixed a set of different recessive deleterious mutations that become
81 heterozygous after the cross of the populations. Although other mechanisms such as
82 overdominance exist to explain heterosis, their effect is generally considered less important
83 than the masking of genetic load (Verhoeven *et al.*, 2011).

84 Admixture is thus often considered as a stimulating factor of successful invasions
85 (Drake, 2006; Ellstrand, Schierenbeck, 2000). However, some authors have suggested that the
86 stimulatory effect of admixture during invasions was often over-estimated (*e.g.* Chapple *et al.*,
87 2013). Admixture can lead to outbreeding depression, which is a decrease in the fitness of

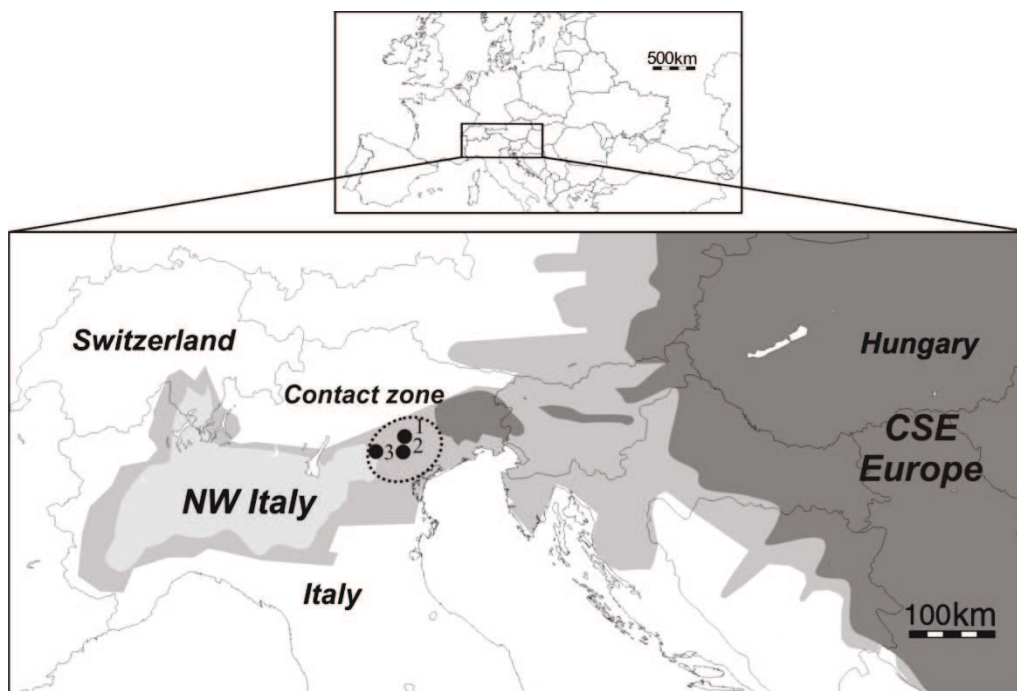
88 admixed individuals relative to their parents may be observed (Edmands, 1999; Lynch, 1991).
89 This can happen through (i) underdominance (Schierup, Christiansen, 1996), (ii) the breaking
90 down of co-adapted gene complexes (Tallmon *et al.*, 2004) and (iii) genetic incompatibilities of
91 Dobzhansky-Muller type (Turelli, Orr, 2000), *i.e.* the whole of independent genetic changes
92 that occur in isolated populations and that must be compatible with their own genetic
93 backgrounds but need not be compatible with other genetic backgrounds (Fitzpatrick, 2008).

94 Here, we used the western corn rootworm (WCR) as a biological model to study the
95 impact of admixture in an invasion context. This univoltine beetle, native to Central America,
96 is considered as one of the most damaging pests of corn in the USA and its associated damages
97 exceed US \$1billion per year (Branson, Krysan, 1981; Smith, 1966; Spencer *et al.*, 2009).
98 Several introductions of WCR from the USA have succeeded in Europe (Ciosi *et al.*, 2008;
99 Miller *et al.*, 2005) since the late 1980's. These introductions led to the formation of two main
100 outbreaks genetically differentiated: the outbreak of Northwestern Italy (NW Italy), first
101 observed in 2000, and the Central and South Eastern Europe (CSE Europe) outbreak, first
102 observed in 1992 (Ciosi *et al.*, 2008). From 2004, these two outbreaks converged to each other
103 and came into contact in 2008 in the Veneto region of Northern Italy, thus forming an admixed
104 zone (Bermond *et al.*, 2012). In this study, we collected 600 adult individuals at the centre of
105 the admixed zone and we genotyped them using 13 microsatellite markers. Three types of
106 information were obtained from the collected individuals: (i) their starving survival; (ii) their
107 admixed status, determined through a Bayesian method of genetic clustering, and (iii) their
108 mating success, studied *via* the observation, isolation and genotyping of sperm in female
109 spermathecae. We statistically tested for a link between the individual admixed status and
110 mating success on the one hand, and the individual starving survival on the other hand. Finally
111 we could also test the hypothesis of an assortative mating in the contact area, *i.e.* preferential
112 mating between individuals of certain admixed status.

113 **MATERIAL AND METHOD**

114 **Samples collection and survival measures**

115 WCR adults were sampled in three locations, Vi9, Tv11 and Tv19, in summer 2012,
116 between the 19th and the 20th of July in the region of Veneto, in northern Italy (Figure 1) (see
117 details in Table 1). For each sampling site, 200 adults WCR were collected from a single corn
118 field. Sampling was conducted in a way to obtain a roughly balanced sex-ratio (see details in
119 Table 1). Adult beetles were caught by hand or by using a mouth vacuum and were then placed
120 individually without food in ventilated (perforated) flat-top 1.5-mL polypropylene microtubes
121 to let a sufficient amount of oxygen to survive until further analyses. Then, microtubes were
122 stored in open plastic storage boxes and placed in the same controlled conditions in climatic
123 rooms at the temperature of 23°C with a photoperiod of 16 hours until the death of all
124 individuals.



125

126 **Figure 1:** Location of the main expanding outbreaks of *Diabrotica virgifera virgifera* in Europe (North
127 West (NW) Italy and Central South Eastern (CSE) Europe outbreaks) and location of the samples
128 collected in 2012 in the contact zone (Veneto, Northern Italy), represented by the dotted circle. Numbers
129 refer to sample names (see Table 1). The distribution areas of WCR before contact between the western
130 and eastern outbreak populations (in 2007) are shown in light gray (NW Italy) and dark gray (CSE
131 Europe). The distribution area of 2012 (after contact) is shown in medium gray.

132 **Table 1:** Description of the *Diabrotica virgifera virgifera* samples from Veneto, in Northern Italy. *N* is the sample size. The number of
 133 genotyped spermatheca represents the number of males that successfully mated and for which multilocus genotypes were determined from
 134 spermatheca.

135

Sample name/ number	Province/Region/Country	<i>N</i>	Number of sampled females	Number of genotyped females	Number of mated females among genotyped	Number of sampled males	Number of genotyped males	Number of genotyped spermathecae	GPS coordinates (Latitude; Longitude)
Tv11/1	Treviso/Veneto/Italy	252	97	96	83	103	103	53	(45°47'52.80"N,11°53'33.50"E)
Tv19/2	Treviso/Veneto/Italy	274	122	118	113	78	78	72	(45°43'26.69"N,11°52'57.61"E)
Vi9/3	Vicenza/Veneto/Italy	261	108	105	100	92	90	65	(45°33'58.70"N, 11°24'59.40"E)

136

137 For each collected adult we measured the survival time in starvation conditions (without food and water). We counted twice a day, late in the
 138 morning (11 am) and in the evening (8 pm), the number of dead individuals per sample site and we randomly moved the remaining living adults
 139 within each storage box, to avoid any effect of the position in the box on insect survival. After their death, individuals were stored in 96% ethanol
 140 at 4°C.

141

142 **Determination of sex, dissection of spermathecae and isolation of sperm**

143 Determining the sex of WCR adults is difficult (but see Hammack, French, 2007). To
144 reliably determine it, we dissected all individuals to look for female or male reproductive
145 organs. If present, spermatheca is easily isolable under stereo-microscope by pulling on the
146 pygidium of the insect with a pair of fine forceps. After isolation, we placed the spermatheca in
147 10 μ L of 0.065% NaCl aqueous solution on a microscope slide. Observations were done under
148 microscope (Zeiss, Axiomager Z1) at 200X and 400X magnification to reveal the presence of
149 spermatozoa (Figure S1A and S1B, respectively). We cut into two pieces the spermatheca and
150 strongly pushed its extremity to release the sperm outside the spermatheca. To completely
151 isolate the sperm, we covered the drop with a coverslip and used the back of the forceps to
152 gently tap the coverslip several times. Spermatozoa were then visible as a mass of hair-like
153 threads extruding from fissures in the spermatheca. We then removed the coverslip and added
154 10 μ L of NaCl to facilitate the manipulation of the sperm mass using forceps and/or needle.
155 The tough sheath of the spermatheca was then completely removed and the sperm was directly
156 transferred with a P10 micropipette in a well of a 96 well-PCR plate for DNA extraction.

157

158 **Genotyping**

159 DNA was extracted from spermatozoa using the commercial kit prepGeM (ZyGeM Ltd,
160 Hamilton, New Zealand) according the manufacturer's instructions with an elution volume of
161 20 μ L. Adults were washed three times in 0.065% NaCl before extraction to remove the excess
162 of ethanol from the tissues. For all individuals, we extracted DNA from the thorax or half body
163 cut lengthwise. The adults' DNA extraction was performed with the DNeasy® tissue kit
164 (Qiagen, Hilden, Germany) following the manufacturer's instructions and with an elution
165 volume of 100 μ L. Thirteen microsatellite loci of WCR (Bermond *et al.*, 2012) were amplified

166 using three separate multiplex PCRs performed in a S1000™ Thermal Cycler and were
167 analyzed as described by Miller et al. (2007). DNA amplification consisted of 25 and 33 PCR
168 cycles for individuals and sperm, respectively.

169 The absence of maternal contamination in the sperm genotypes was verified by
170 checking the following conditions: firstly, the sperm genotypes should contain at most 2 alleles
171 because WCR is diploid. Secondly we compared the genotypes of sperm and those of the
172 associated females. The genotypes of the sperm should not be strictly identical to that of the
173 mother. Note that this is a stringent condition because, for particularly frequent alleles, those
174 genotypes may be identical even without contamination.

175

176 **Data analysis**

177 *Genetic assignment of WCR sample to parental populations and estimation of the rate of*
178 *admixture.*

179 In order to determine the degree of admixture of the individuals and their genotypic
180 classes (Pure-East or Pure-West parental, F₁, admixed genotypes of 2nd, 3rd, 4th and 5th
181 generations (because the contact took place 5 generations ago) and backcrosses (B_x)) we
182 sought to assign each individual of each sample to parental populations/outbreaks using the
183 Bayesian method implemented in STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000). For this,
184 we used 6 reference samples from NW Italy and 7 reference samples from CSE Europe
185 outbreaks (Table S1 and Figure S2) as representative samples of parental populations
186 (Bermond *et al.*, 2012). Because the Veneto region is an admixed zone between two parental
187 populations (Bermond *et al.*, 2013; Bermond *et al.*, 2012), we performed 20 runs with the
188 number of clusters (*K*) fixed to 2. Each run consisted of a burn-in of 2x10⁵ iterations, followed
189 by 10⁶ iterations. We used the admixture model together with the correlated allele frequencies

190 model (Falush *et al.*, 2003), without the use of sampling location as prior information (Hubisz
191 *et al.*, 2009). Default values were maintained for all other parameters. Five non-overlapping
192 classes of values of the co-ancestry coefficient Q were defined to characterize the individual
193 admixed status: (class 1 = [0-0.2]; class 2 = [0.2-0.4]; class 3 = [0.4-0.6]; class 4 = [0.6-0.8];
194 class 5 = [0.8-1]). As a rule of thumb, classes 1 and 5 should correspond to Pure-East and Pure-
195 West parental genotypes, respectively; classes 2 and 4 should correspond to backcross
196 genotypes and class 3 should correspond to F_1 and admixed genotypes of 2nd, 3rd, 4th and 5th
197 generations. These classes are referred to as Structure classes hereafter.

198 The coefficient of co-ancestry Q was also used to compute the individual rate of
199 admixture (ROA) ranging from 0 (100% of the genome is of the east or west parental type) to 1
200 (half the genome is of the east parental type, and the other half is of the west parental type).
201 ROA is calculated as follows: $ROA = 2Q$ (for $Q \leq 0.5$) and $ROA = 2(1 - Q)$ (for $Q > 0.5$). ROA
202 is thus maximal for a Q value of 0.5 (pure admixed type) and minimal for a Q value of 0 or 1
203 (parental type).

204 Therefore the admixed status of each individual was characterized by one qualitative
205 variable obtained using the clustering method (the genotypic classes or Structure classes), and
206 one quantitative variable (ROA) intimately linked to the Structure classes.

207 A second method of genetic clustering using NewHybrids (Anderson, Thompson, 2002)
208 was also used to confirm our results (data shown in SUPPORTING INFORMATION:
209 *GENETIC CLUSTERING USING NEWHYBRIDS*).

210

211 **Statistical analysis**

212 *Impact of the female admixed status on the mating probability*

213 To test the hypothesis of a relationship between admixed status of females and their
214 mating status (mated or not), we compared the distributions of mated and unmated females
215 among admixed and non-admixed genotypic classes in each sample (Tv11, Tv19 and Vi9) with
216 Fisher’s exact tests on 2x2 contingency tables. Admixed classes matched to class 2, 3 or 4
217 while parental classes matched to class 1 or 5. Thus, we performed 6 Fisher’s exact tests when
218 considering the Structure classes. Because these tests involve multiple comparisons, levels of
219 significance were corrected by the procedure of Benjamini and Hochberg (1995) implemented
220 in the function “p.adjust” of R (R Development Core Team, 2012).

221 We also tested the hypothesis of a relationship between admixed status of females and
222 their mating status (mated or not) with generalized linear mixed models (GLMM) with the
223 mating status as the response variable (two modalities: “mated females” and “non-mated
224 females”) with a binomial error distribution, the genotypic classes (Structure classes) as a
225 qualitative explanatory variable or the ROA as a quantitative explanatory variable and the
226 sample site as a random explanatory variable:

227 Mating success of females \sim Structure classes + (1|sample) {1}

228 Mating success of females \sim ROA + (1|sample) {2}

229

230 *Impact of the male admixed status on their mating probability*

231 To test the hypothesis of a relationship between the admixed status of males and their
232 reproductive success, we performed the same kind of analyses as above (Fisher’s exact tests
233 and GLMM). However, as mating status of males cannot be determined by direct observation,
234 inferences had to be made using sperm – the genotypes of sperms correspond to the genotypes
235 of mated males. We therefore compared the distributions of sperms (corresponding to mated

236 males only) and adult males (corresponding to both mated and unmated males) among admixed
237 and non-admixed genotypic classes using Fisher’s exact tests in each sample. Levels of
238 significance of p -values were corrected as described previously. In the GLMM, the dependent
239 variable “Reproductive success of males” included two modalities, “sperm” and “sampled
240 males”.

241 Mating success of males \sim Structure classes + (1|sample) {3}

242 Mating success of males \sim ROA + (1|sample) {4}

243

244 *Impact of the admixed status of individuals on their survival time*

245 The time between sampling and death in starvation conditions of each of the 600 collected
246 adults was measured. The link between the admixed status of individuals and their survival
247 time was tested with 2 mixed Cox’s models (Cox, 1972) with R (R Development Core Team,
248 2012), with the sample as random effect:

249 Survival probability \sim Structure classes + sex + (1|sample) {5}

250 Survival probability \sim ROA + sex + (1|sample) {6}

251

252 *Impact of the genotypic type of WCR individuals on their mating choices*

253 We tested the hypothesis of random mating between individuals according to their
254 genotypic class within each sample. For this, we tested the independence between the
255 genotypic class (admixed or parental) of sperm and that of the female mated with that sperm
256 with Fisher exact tests on 2x2 contingency tables. As before, Structure classes were clustered
257 into 2 modalities: admixed (classes 2, 3 and 4) and parental (classes 1 and 5).

258

259

260

261

262 **RESULTS**

263

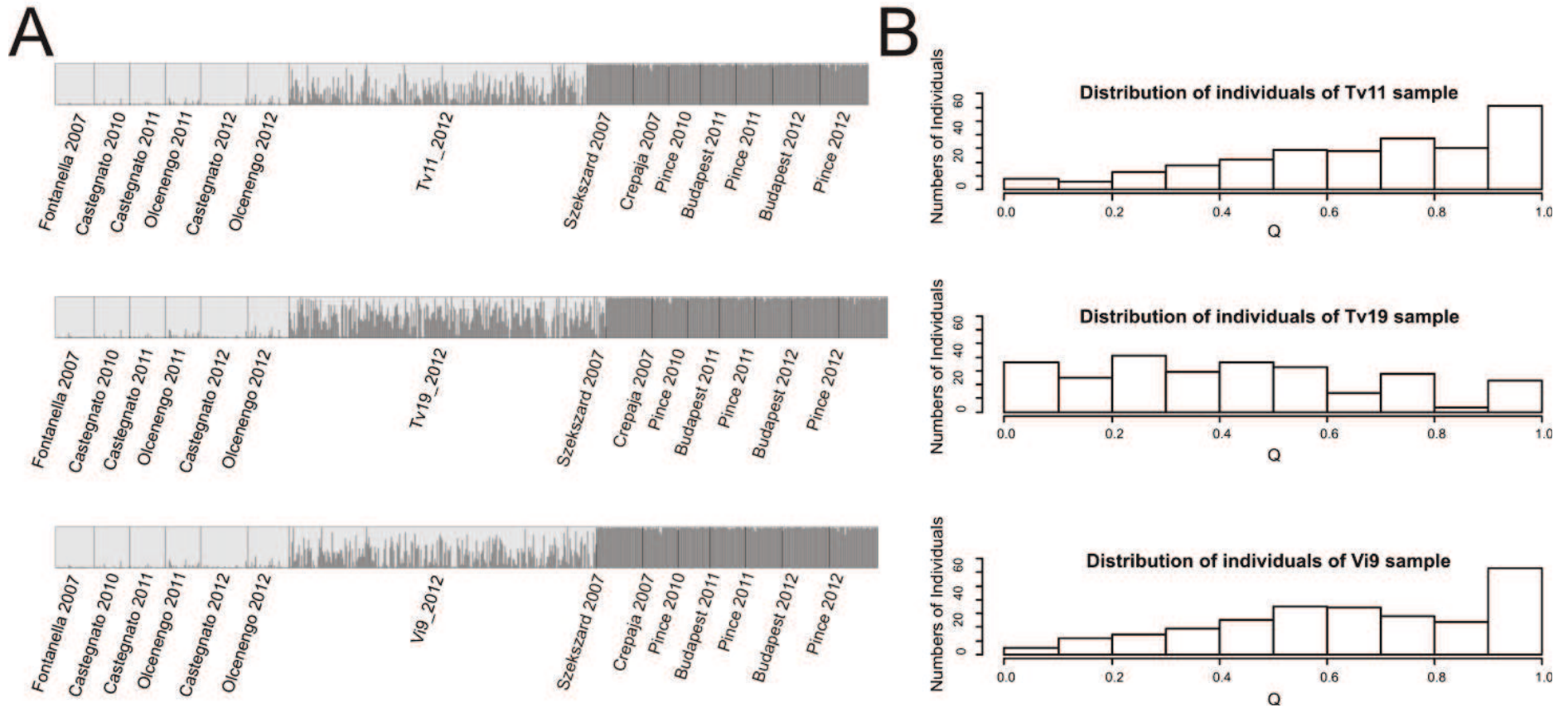
264 *Determination of sex, dissections of spermathecae and sperm isolation*

265 A total of 327 females and 273 males were collected in Tv11, Tv19 and Vi9 (see Table
266 1 for more details). Among the 319 genotyped females, 93% were mated (296/319). The
267 number of genotyped males and spermathecae was 217 and 190, respectively.

268

269 *Genetic assignment of WCR samples to parental populations and estimation of the rate of*
270 *admixture.*

271 The Structure analysis sorted, as expected, the 6 samples representative of the NW Italy
272 outbreak within one cluster (Pure-West, all $Q > 0.8$) and the 7 samples representative of the
273 CSE Europe outbreak within the other cluster (Pure-East, all $Q < 0.2$) (see Figure 2A). Q
274 coefficients of individuals sampled in Tv11, Tv19 and Vi9 vary from 0.014 to 0.989. Overall,
275 26 % of the individuals from these samples were assigned to the Pure-West cluster (with $Q >$
276 0.8), 12% to the Pure-East cluster (with $Q < 0.2$) and 62% were admixed between both clusters
277 ($0.2 < Q < 0.8$) (Figure 2 and Figure 3). For Tv11 and Vi9, a bimodal distribution of Q was
278 found with a mode at 1 (a group of individuals is assigned to the Pure-West) and the other
279 mode at 0.8 and 0.6 for Tv11 and Vi9, respectively (Figure 2B). In Tv19, the distribution of Q
280 looks like a mixture between a unimodal and a uniform distribution with a flat mode in the
281 interval of 0 to 0.6 (Figure 2B).



282

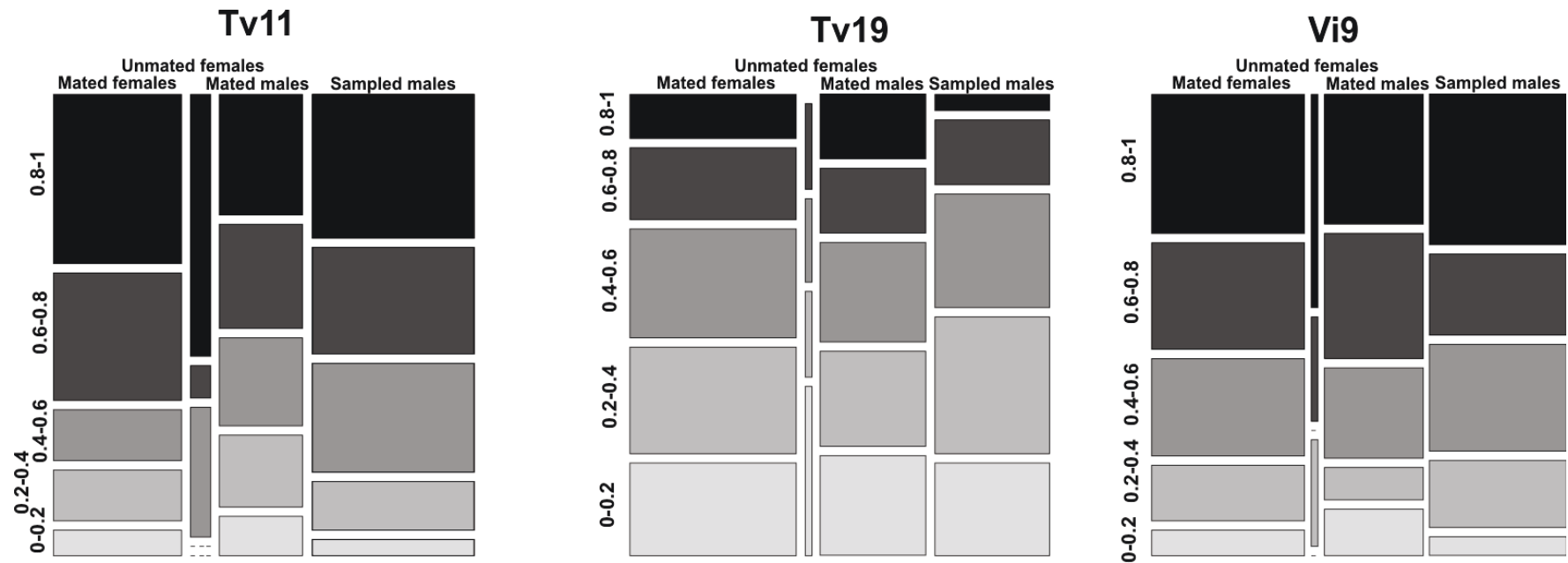
283 **Figure 2:** Admixture analysis of Tv11, Tv19 and Vi9 samples resulting from a Structure analysis based on 13 microsatellite loci. The most likely run is here represented with
 284 DISTRUCT version 1.0 (Rosenberg, 2004). A- Assignment of WCR individuals and sperms to $K = 2$ genetically distinct clusters. For each adult or sperm sample, a vertical
 285 bar displaying two colours represents the proportion of the genome of the individuals that belongs to each of the two clusters. Adults or sperm samples are grouped by
 286 sampling location (the name of which is above the plot). Q. B- Distribution of the individual coefficient of coancestry Q for each sample, Tv11, Tv19 and Vi9.

287

288 **Influence of admixed status on the fitness of individuals**

289 *Impact of the female admixed status on their mating probability*

290 The female mating probability was not significantly influenced by the Structure classes (Figure 3, $p = 0.49$ for GLMM {1}) or by the
 291 ROA (GLMM {2}, $p = 0.35$). Fisher’s exact tests on Structure classes showed that “parental” (class 1 or class 5) and “admixed” classes (class 2,
 292 class 3 or class 4) displayed very similar frequencies among mated and unmated females (Figure 3, $p > 0.05$ for all 6 tests). Note that unmated
 293 females are probably too few ($n = 23$) to get a large statistical power.



294

295 **Figure 3:** Admixed class distribution (Structure classes) of samples Tv11, Tv19 and Vi9 among mated and unmated females and among sperm and males (see text
 296 for details). Bx= backcross.

297 *Impact of male admixed status on their mating probability*

298 The Structure classes and the ROA did not statistically influence the reproductive
299 success of males whatever the considered variable (Figure 3, $p = 0.33$ and 0.22 for GLMM {3}
300 and {4}, respectively). Fisher's exact test on Structure classes showed that "parental" and
301 "admixed" classes displayed very similar frequency among mated and unmated males
302 (Figure 3, $p > 0.05$ for most tests).

303

304 *Impact of the admixed status of individuals on their survival time*

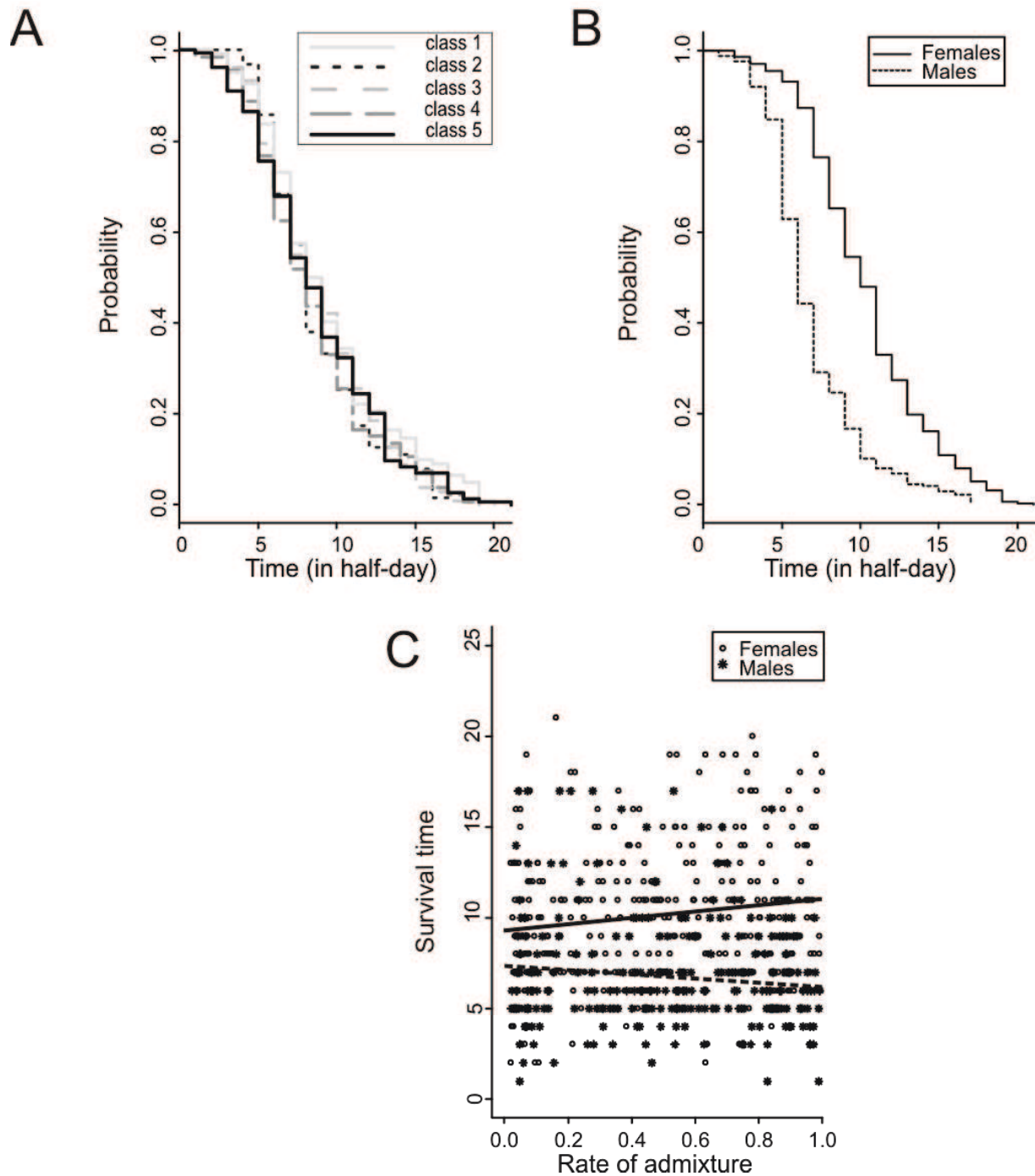
305 The different Structure classes and the ROA did not display a significant influence on
306 the survival time of starving WCR adults (Figure 4, $p = 0.10$ and 0.61 , for Cox's models {5},
307 and {6}, respectively). The mortality hazard is 1.3 higher in males than in females (Fig 4B, 4C,
308 Cox's models, $p < 10^{-5}$).

309

310 *Impact of the genotypic type of individuals on the mating choices within WCR samples*

311 Fisher's exact tests indicated that, in all samples, the genotypic classes (admixed or
312 parental determined from the Structure analyses) of sperm and that of the female mated with
313 that sperm were not statistically dependent of each other ($p > 0.05$ for all comparisons).
314 Therefore, no choice of the sexual partner according to its hybrid status has been detected.

315



316

317 **Figure 4:** Survival of starving individuals as a function of their hybrid status. A- Survival probability of
 318 individuals belonging to the different classes of Structure (class1, class2, class3, class4, class5) as a function of
 319 time (mixed Cox's models $p > 0.05$). B- Survival probability of males and females as a function of time (mixed
 320 Cox's model, $p < 10^{-3}$). C- Survival time of individuals sampled in the population Tv11, Tv19 and Vi9 as a
 321 function of their ROA (mixed Cox's model, $p > 0.05$). Solid line: trend of survival time of females. Dashed line:
 322 trend of survival time of males.

323

324 **DISCUSSION**

325

326 In this study, we tested whether the admixed status influenced the individual fitness in
327 the context of the European invasion of WCR. Our results show that mating success and
328 survival of adults in starvation conditions are very probably not influenced by their admixed
329 status. Moreover, the choice of sexual partners does not seem to depend on the admixed class
330 of the adults.

331 **Admixed individuals are frequent in the centre of the WCR contact zone**

332 Under the hypothesis of a selection against admixed individuals, frequencies of parental
333 genotypes are classically higher than frequencies of admixed genotypes in an admixed zone
334 (Jiggins, Mallet, 2000). As a consequence, the distribution of the Q coefficients is expected to
335 be U-shaped, with numerous low (0-0.2) and high (0.8-1) Q values and relatively few
336 intermediate (0.2-0.8) Q values. This is not what we observed here: (i) results of Structure
337 analyses show that the 3 samples contain a majority of admixed individuals; (ii) Tv11 and Vi9
338 contain very few individuals belonging to the class 1 (0-0.2), probably because the samples
339 were collected in the western part of the contact zone rather than in its centre.

340 **Admixed status probably does not influence individual fitness**

341 In this study, we sought to determine whether the genotypic classes (Structure classes)
342 and the ROA of individuals influenced their probability of mating and their ability to survive to
343 stressful conditions. To do this, we used statistical tests (Fisher's exact tests, GLMM and mixed
344 Cox's models) (i) to compare the genotypes of mated females with those of non-mated females,
345 (ii) to compare the genotypes of all males with those of successful mated males, *via* the
346 genotype of sperm contained in mated females, (iii) to test whether the mating is random

347 according to the genotypic classes, and (iv) to test if the genotypic classes and the ROA had an
348 impact on the survival time of starving adults.

349 Concerning the two first points, no significant effect was detected and we therefore
350 conclude that, overall, the genotype of the males did not influence the probability of mating.
351 The same conclusion holds for females. Therefore there is no indication that admixed status
352 influences mating probability. However statistical tests associated with female mating success
353 may not display sufficient power to detect small effects because few unmated females were
354 observed. It is noteworthy that we have measured only two traits related to fitness – the mating
355 rate and the survival time of starving adults – that can hardly be generalized to all components
356 of fitness. Additional analyses on other phenotypic traits such as fertility, speed of development
357 or the rate of eggs survival are thus necessary. However unpublished results on other
358 phenotypic traits (Stefan Toepfer, HongMei Li, Gerald Bermond, unpublished data) confirm
359 our results.

360 Concerning the hypothesis of random mating, no association was detected between the
361 genotypic classes of males and that of females within mated couples: the genotypic classes of
362 the mated females and that of the sperm in the females were not statistically associated.
363 Statistical power issues are not likely the cause of the absence of an effect here as the genotypic
364 class frequencies were most often equilibrated among mated individuals (sperm and mated
365 females).

366 Similarly, analyses of Cox's models show that the genotypic classes (Structure classes)
367 and the ROA do not influence survival time in stressful conditions. However, sex has a strong
368 and significant influence: males have a mortality risk 1.3 times higher than females. This result
369 was unexpected because the literature suggests that WCR males live longer than females *in*
370 *natura* (102 days for males against 78 for females) (Hill, 1975). Two hypotheses may explain

371 this result: (i) our samples were collected after mating of the majority of females, at a time
372 when they are feeding while males are looking for sexual partners and have probably fewer
373 resources than females (Hill, 1975; Spencer *et al.*, 2009). Given that the survival test took in
374 starving conditions, males have probably suffered a disadvantage compared to females; (ii) at
375 the time of mating, the males transfer the sperm to the females *via* the spermatophore. In some
376 insects, the spermatophore is associated to the spermatophylax, a nutritive vesicle given by
377 males to females as nuptial gift that extends the mating time and ensures the sperm transfer
378 (Arnaud, 1999; Vahed, 1998). This nutritive vesicle is present in WCR (Murphy, Krupke,
379 2011). If females also use the spermatophylax as nutritive resource, they may benefit of an
380 energy bonus relative to males, explaining their larger survival.

381 It is important to note that the analysis of the survival time is semi-controlled. We do
382 not have any information on the age of individuals or on the date of their last food intake.
383 However, the hypothesis that age and satiety of individuals are randomly distributed in the
384 different hybrid classes seems reasonable and thus the detection of an effect of the hybrid status
385 should not be related to age or satiety.

386 The absence of detection of admixture effect on fitness components of WCR might be
387 explained by the timing of the study within the invasion process. Indeed, most studies that
388 revealed a positive impact of admixture on invasion were conducted in situations where
389 admixture occurred in the first steps of invasions, soon after introductions (Drake, 2006;
390 Ellstrand, Schierenbeck, 2000; Keller, Taylor, 2010). In the present study, our sampling and
391 tests occur about 12 and 20 generations after the introductions of WCR into the two invaded
392 European areas. Impacts of admixture might greatly differ between early contacts soon after
393 primary introductions and secondary contacts. Assuming that the invasion success partly
394 depends on the capacity of a population to respond to new selective pressures, in cases of
395 secondary contacts such as observed in Veneto, both populations may have already proved

396 successful and may have evolved towards larger fitness. Hence, it may be more difficult to
397 detect positive effects of admixture in this context.

398

399 **Neutral contact zones to estimate dispersal**

400 Recently, Bermond et al (2013) has used the admixed zone of Veneto to estimate
401 dispersal capabilities of WCR by studying the geographic clines of allelic frequencies of
402 microsatellite markers. The dispersal of WCR was estimated at about 20 km/generation. The
403 idea beyond the analyses of Bermond et al. (2013) is that the size of the admixed zone (the
404 width of the clines) increases at a speed that depends on the dispersal. This estimate relies on
405 the idea that the admixed zone is neutral, *i.e.* (i) individuals of parental and admixed types do
406 not differ in fitness and (ii) mating is random.

407 Previous laboratory tests on phenotypic traits of WCR did not show any significant
408 difference in measured traits between parental (NW Italy and CSE Europe) strains and F1
409 crosses (Stefan Toepfer, HongMei Li, Gerald Bermond, unpublished data). The observed
410 phenotypic traits were the individual weight, the length and width of the elytra, the fecundity,
411 the egg survival during the winter and the larval and pupal survival. The average value of
412 phenotypic traits of admixed populations was either an intermediate value between the values
413 of the two parental populations or equal to that of one of the parents. These results are
414 consistent with those of the present study, *i.e.* neither heterosis nor outbreeding depression was
415 found. Our results suggest that the contact area is actually neutral or close to neutrality and that
416 mating is random. Therefore they agree with the hypothesis assumed by Bermond et al. (2013).

417

418

419 **Conclusion**

420 The present study is original because most results on fitness were obtained without
421 manipulation or experimentations and were directly obtained on individuals collected *in natura*.
422 In addition, the results we obtained are remarkable by their simplicity: no effect of the admixed
423 status on traits related to fitness whatever the measure of the admixed status and whatever the
424 trait. Whether admixture generally has a positive impact on invasion remains to be carefully
425 studied and more case studies on ongoing invasions are necessary to reach definitive
426 conclusions (Kolbe *et al.*, 2007; Rieux *et al.*, 2013; Verhoeven *et al.*, 2011). They constitute
427 many field models on which invasion biologists can test the role of admixture, keeping in mind
428 the critical importance of negative results in the literature.

429

430 **ACKNOWLEDGMENTS**

431 We thank Laura Maïche and Stefan Toepfer for technical assistance during the sampling
432 campaign. We are also grateful to Arnaud Estoup, Eric Lombaert, Etienne Klein, Elodie
433 Vercken, Virginie Ravigné, Benoit Facon, Nicolas Ris and Yannis Michalakis for their
434 contribution and their help in elaborating experimental design and discussion. This work was
435 supported by grants from the *Agence Nationale de la Recherche* (ANR-06-BDIV-008-01 and
436 ANR-09-BLAN-0145-01) and from the Agropolis Foundation (RTRA-Montpellier, BIOFIS
437 project number 1001-001).

438

439 REFERENCES

- 440 Allee WC, Park O, Emerson AE, Park T, Schmidt KP (1949) *Principles of animal ecology* WB
441 Saundere Co. Ltd.
- 442 Anderson E, Thompson E (2002) A model-based method for identifying species hybrids using
443 multilocus genetic data. *Genetics* 160, 1217-1229.
- 444 Anderson EC (2003) User's guide to the program NewHybrids version 1.1 beta. Technical
445 report, 7 April.
- 446 Arnaud L (1999) La compétition spermatique chez les insectes: les stratégies d'assurance de la
447 paternité et la préséance du sperme. *Biotechnol. Agron. Soc. Environ* 3, 86-103.
- 448 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful
449 approach to multiple testing. *Journal of the Royal Statistical Society Series B-
450 Methodological* 57, 289-300.
- 451 Bermond G, Blin A, Vercken E, *et al.* (2013) Estimation of the dispersal of a major pest of
452 maize by cline analysis of a temporary contact zone between two invasive outbreaks.
453 *Molecular Ecology*.
- 454 Bermond G, Ciosi M, Lombaert E, *et al.* (2012) Secondary Contact and Admixture between
455 Independently Invading Populations of the Western Corn Rootworm, *Diabrotica
456 virgifera virgifera* in Europe. *PloS one* 7, e50129.
- 457 Branson TF, Krysan JL (1981) Feeding and oviposition behavior and life cycle strategies of
458 *Diabrotica*: an evolutionary view with implications for pest management.
459 *Environmental Entomology* 10, 826-831.
- 460 Chapple DG, Miller KA, Kraus F, Thompson MB (2013) Divergent introduction histories
461 among invasive populations of the delicate skink (*Lampropholis delicata*): has the
462 importance of genetic admixture in the success of biological invasions been
463 overemphasized? *Diversity and Distributions* 19, 134-146.
- 464 Ciosi M, Miller NJ, Kim KS, *et al.* (2008) Invasion of Europe by the western corn rootworm,
465 *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various
466 reductions of genetic diversity. *Molecular Ecology* 17, 3614-3627.
- 467 Cox DR (1972) Regression models and life-tables. *Journal of the Royal Statistical Society.
468 Series B (Methodological)*, 187-220.
- 469 Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation,
470 adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17, 431-
471 449.
- 472 Dobzhansky T, Gowen J (1952) Nature and origin of heterosis. *Heterosis.*, 218-223.
- 473 Drake JM (2006) Heterosis, the catapult effect and establishment success of a colonizing bird.
474 *Biology Letters* 2, 304-307.
- 475 Edmands S (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a
476 wide range of divergence. *Evolution* 53, 1757-1768.
- 477 Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of
478 invasiveness in plants? *Proceedings of the National Academy of Sciences of the United
479 States of America* 97, 7043-7050.
- 480 Elton CS (1958) *The Ecology of Invasions by Animals and Plants* Methuen & Co LTD,
481 London.
- 482 Facon B, Pointier JP, Jarne P, Sarda V, David P (2008) High genetic variance in life-history
483 strategies within invasive populations by way of multiple introductions. *Current
484 Biology* 18, 363-367.
- 485 Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus
486 genotype data: Linked loci and correlated allele frequencies. *Genetics* 164, 1567-1587.

- 487 Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced
488 populations, with special reference to biological control. *Evolutionary applications* 5,
489 424-443.
- 490 Fitzpatrick BM (2008) Hybrid dysfunction: Population genetic and quantitative genetic
491 perspectives. *The American Naturalist* 171, 491-498.
- 492 Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations
493 of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of
494 introduction. *Molecular Ecology* 14, 4275-4285.
- 495 Gurevitch J, Fox GA, Wardle GM, Inderjit, Taub D (2011) Emergent insights from the
496 synthesis of conceptual frameworks for biological invasions. *Ecology letters* 14, 407-
497 418.
- 498 Hammack L, French BW (2007) Sexual dimorphism of basitarsi in pest species of *Diabrotica*
499 and *Cerotoma* (Coleoptera : Chrysomelidae). *Annals of the Entomological Society of*
500 *America* 100, 59-63.
- 501 Hill RE (1975) Mating, oviposition patterns, fecundity and longevity of the western corn
502 rootworm. *Journal of Economic Entomology* 68, 311-315.
- 503 Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure
504 with the assistance of sample group information. *Molecular Ecology Resources* 9, 1322
505 -1332.
- 506 Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology &*
507 *Evolution* 15, 250-255.
- 508 Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion.
509 *Journal of Evolutionary Biology* 23, 1720-1731.
- 510 Kolbe JJ, Glor RE, Schettino LR, *et al.* (2007) Multiple sources, admixture, and genetic
511 variation in introduced *Anolis* lizard populations. *Conservation Biology* 21, 1612-1625.
- 512 Kolbe JJ, Larson A, Losos JB, de Queiroz K (2008) Admixture determines genetic diversity
513 and population differentiation in the biological invasion of a lizard species. *Biology*
514 *Letters* 4, 434-437.
- 515 Lande R (1993) Risks of population extinction from demographic and environmental
516 stochasticity and random catastrophes. *American Naturalist* 142, 911-927.
- 517 Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the
518 success of an invasive grass. *Proceedings of the National Academy of Sciences of the*
519 *United States of America* 104, 3883-3888.
- 520 Lynch M (1991) The genetic interpretation of inbreeding depression and outbreeding
521 depression. *Evolution* 45, 622-629.
- 522 Miller N, Estoup A, Toepfer S, *et al.* (2005) Multiple transatlantic introductions of the western
523 corn rootworm. *Science* 310, 992.
- 524 Miller NJ, Ciosi M, Sappington TW, *et al.* (2007) Genome scan of *Diabrotica virgifera*
525 *virgifera* for genetic variation associated with crop rotation tolerance. *Journal of*
526 *Applied Entomology* 131, 378-385.
- 527 Moyle PB, Light T (1996) Fish invasions in California: do abiotic factors determine success?
528 *Ecology* 77, 1666-1670.
- 529 Murphy AF, Krupke CH (2011) Mating Success and Spermatophore Composition in Western
530 Corn Rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology* 40, 1585-
531 1594.
- 532 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
533 multilocus genotype data. *Genetics* 155, 945-959.
- 534 R Development Core Team (2012) R: A Language and Environment for Statistical Computing.
535 R Foundation for Statistical Computing, Vienna, Austria.

- 536 Rieux A, Lenormand T, Carlier J, Lapeyre de Bellaire L, Ravigné V (2013) Using neutral cline
537 decay to estimate contemporary dispersal: a generic tool and its application to a major
538 crop pathogen. *Ecology letters*.
- 539 Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range.
540 *Proceedings of the Royal Society B-Biological Sciences* 273, 2453-2459.
- 541 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure.
542 *Molecular Ecology Notes* 4, 137-138.
- 543 Rosenthal DM, Ramakrishnan AP, Cruzan MB (2008) Evidence for multiple sources of
544 invasion and intraspecific hybridization in *Brachypodium sylvaticum* (Hudson) Beauv.
545 in North America. *Molecular Ecology* 17, 4657-4669.
- 546 Sakai AK, Allendorf FW, Holt JS, *et al.* (2001) The population biology of invasive species.
547 *Annual Review of Ecology and Systematics* 32, 305-332.
- 548 Schierup MH, Christiansen FB (1996) Inbreeding depression and outbreeding depression in
549 plants. *Heredity* 77, 461-468.
- 550 Shull GH (1914) Duplicate genes for capsule-form in *Bursa bursa-pastoris*. *Molecular and*
551 *General Genetics MGG* 12, 97-149.
- 552 Smith RF (1966) Distributional patterns of selected western north american insects: the
553 distribution of diabroticites in western north America. *Bulletin of Entomology of society*
554 *of America* 12, 108-110.
- 555 Spencer JL, Hibbard BE, Moeser J, Onstad DW (2009) Behaviour and ecology of the western
556 corn rootworm (*Diabrotica virgifera virgifera* LeConte). *Agricultural and Forest*
557 *Entomology* 11, 9-27.
- 558 Tallmon DA, Luikart G, Beaumont MA (2004) Comparative evaluation of a new effective
559 population size estimator based on approximate Bayesian computation. *Genetics* 167,
560 977-988.
- 561 Taylor DR, Keller SR (2007) Historical range expansion determines the phylogenetic diversity
562 introduced during contemporary species invasion. *Evolution* 61, 334-345.
- 563 Turelli M, Orr HA (2000) Dominance, epistasis and the genetics of postzygotic isolation.
564 *Genetics* 154, 1663-1679.
- 565 Vahed K (1998) The function of nuptial feeding in insects: a review of empirical studies.
566 *Biological Reviews* 73, 43-78.
- 567 Verhoeven KJF, Macel M, Wolfe LM, Biere A (2011) Population admixture, biological
568 invasions and the balance between local adaptation and inbreeding depression.
569 *Proceedings of the Royal Society B-Biological Sciences* 278, 2-8.
- 570 Wares JP, Hughes AR, R.K. G (2005) Mechanisms that drive evolutionary change: insights
571 from species introductions and invasions. In: *Species Invasions: Insights into Ecology,*
572 *Evolution and Biogeography* (eds. Sax DF, Stachowicz JJ, Gaines SD), pp. 229-257.
573 Sinauer Associates Inc, Sunderland, MA USA.
- 574 Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small
575 populations. *Annual Review of Ecology, Evolution, and Systematics*, 433-458.
- 576 Williamson M, Fitter A (1996) The varying success of invaders. *Ecology* 77, 1661-1666.
- 577

SUPPORTING INFORMATION

Table S1

Sample number	Sample	Outbreak	Origin(Region(Country))	Sampling date	<i>N</i>	GPS coordinates
1	Olcenengo	NW Italy	Lombardia (Italy)	2011	30	(45°21'59.50"N; 08°18'18.40"E)
2	Olcenengo	NW Italy	Lombardia (Italy)	2012	35	(45°21'59.50"N; 08°18'18.40"E)
3	Fontanella	NW Italy	Lombardia (Italy)	2007	33	(45°26'00.40"N ; 9° 47'48.54"E)
4	Castegnato	NW Italy	Lombardia (Italy)	2010	30	(45°34'06.52"N ; 10° 6'35.63"E)
5	Castegnato	NW Italy	Lombardia (Italy)	2011	30	(45°34'06.52"N ; 10° 6'35.63"E)
6	Castegnato	NW Italy	Lombardia (Italy)	2012	40	(45°34'06.52"N ; 10° 6'35.63"E)
7	Pince	CSE Europe	Prekmurje (Slovenia)	2010	27	(46°31'15.38"N; 16°31'23.70"E)
8	Pince	CSE Europe	Prekmurje (Slovenia)	2011	31	(46°31'15.38"N; 16°31'23.70"E)
9	Pince	CSE Europe	Prekmurje (Slovenia)	2012	41	(46°31'15.38"N; 16°31'23.70"E)
10	Szekszard	CSE Europe	Tolna (Hungary)	2007	39	(46°30'59.52"N; 18°37'57.72"E)
11	Budapest	CSE Europe	Pest (Hungary)	2011	30	(47°31'12.90"N; 18°46'33.10"E)
12	Budapest	CSE Europe	Pest (Hungary)	2012	40	(47°31'12.90"N; 18°46'33.10"E)
13	Crepaja	CSE Europe	Voïvodine (Serbia)	2007	30	(45°2'55.38"N; 20°38'23.52"E)

Note: NW: North West. CSE: Central and South-Eastern. For all samples, DNA was extracted with the DNeasy tissue kit (Qiagen, Hilden, Germany), as explained in Materials and Methods, *Genotyping*.

GENETIC CLUSTERING USING NEWHYBRIDS

Materials and Methods

Genetic assignment of WCR sample to parental populations and determination of the admixed status.

As alternative method of genetic clustering, we used the Bayesian genetic clustering approach implemented in Newhybrids 1.1 (Anderson, Thompson, 2002). This program uses Markov chain Monte Carlo simulations to compute the posterior probability that each individual in a sample falls into six genotypic classes: two parental classes (Pure-West and Pure-East), F_1 and F_2 admixed, two classes of backcrosses (Bx-West and Bx-East corresponding to the offspring of F_1 and Pure-West matings or F_1 and Pure-East matings, respectively). Since our loci were semi-diagnostic (Bermond *et al.*, 2013) they made it possible to distinguish easily the 4 hybrid classes (F_1 vs. F_2 and backcrosses) from each other (Anderson, Thompson, 2002). The genotypic class of each individual was defined by the class with the largest posterior probability. Because low-frequency alleles were abundant in our dataset, uniform priors for allele frequencies and mixing proportion were used rather than Jeffreys priors as suggested by Anderson (2003). Posterior probabilities were obtained from the mean of 4 runs of 2×10^5 burn-in iterations followed by 10^6 million iterations. These classes are referred to as NewHybrids classes hereafter.

Statistical analyses

Impact of the female admixed status on the mating probability

To test the hypothesis of a relationship between admixed status of females and their mating status (mated or not), we compared the distributions of mated and unmated females among admixed and non-admixed genotypic classes in each sample (Tv11, Tv19 and Vi9) with Fisher's exact tests on 2×2 contingency tables. Admixed classes matched to F_1 , F_2 , Bx-

West or Bx-East while parental classes matched to Pure-West or Pure-East. Thus, we performed 8 Fisher’s exact tests when considering NewHybrids classes. Because these tests involve multiple comparisons, levels of significance were corrected by the procedure of Benjamini and Hochberg (1995) implemented in the function “p.adjust” of R (R Development Core Team, 2012).

We also tested the hypothesis of a relationship between admixed status of females and their mating status (mated or not) with generalized linear mixed models (GLMM) with the mating status as the response variable (two modalities: “mated females” and “non-mated females”) with a binomial error distribution, the genetic classes (NewHybrids classes) as a qualitative explanatory variable and the sample site as a random explanatory variable:

$$\text{Mating success of females} \sim \text{NewHybrids classes} + (1|\text{sample}) \quad \{1\}$$

Impact of the male admixed status on their mating probability

To test the hypothesis of a relationship between the admixed status of males and their reproductive success, we performed the same kind of analyses as above (Fisher’s exact tests and GLMM). However, as mating status of males cannot be determined by direct observation, inferences had to be made using sperm. We therefore compared the distributions of sperms (corresponding to only mated males) and adult males (corresponding to both mated and unmated males) among admixed and non-admixed genotypic classes using Fisher’s exact tests in each sample. Levels of significance of *p*-values were corrected as described previously. In the GLMM, the dependent variable “Reproductive success of males” included two modalities, “sperm” and “sampled males”.

$$\text{Mating success of males} \sim \text{NewHybrids classes} + (1|\text{sample}) \quad \{2\}$$

Impact of the admixed status of individuals on their survival time

The time between sampling and death in starvation conditions of the 600 collected adults was measured. The link between the admixed status of individuals and their survival time was tested with a mixed Cox's model (Cox, 1972) with R (R Development Core Team, 2012), with the sample as random effect:

$$\text{Survival probability} \sim \text{NewHybrid classes} + \text{sex} + (1|\text{sample}) \quad \{3\}$$

Impact of the genotypic type of WCR individuals on their mating choices

We tested the hypothesis of random mating between individuals according to their genotypic class within each sample. For this, we tested the independence between the genotypic class (admixed or parental) of sperm and that of the female mated with that sperm with Fisher exact tests on 2x2 contingency tables. As before, NewHybrids classes were clustered into 2 modalities: admixed (F₁, F₂, Bx-West, Bx-East) and parental (Pure-West, Pure-East).

RESULTS

The analysis of NewHybrid reveals that the majority of expected classes (Pure-West, Pure-East, F₁, F₂, Bx-West, Bx-East) are present in each sample (Figure S3). However, their frequency varies from one sample to another (Fisher's exact test, $p = 5.10^{-4}$). Tv11 and Vi9 have a majority of F₂ and Pure-west types whereas Tv19 mainly contains F₂ and Pure-east types (Figure S3).

Impact of the female admixed status on their mating probability

The female mating probability was not significantly influenced by the NewHybrids classes (Figure S3, $p = 0.57$ for GLMM {1}). Fisher's exact tests on NewHybrids classes showed that "parental" and "admixed" classes displayed very similar frequencies among

mated and unmated females (Figure S3, $p > 0.05$ for all 8 tests). Note that unmated females are probably too few for a satisfactory statistical power.

Impact of male admixed status on their mating probability

The genotypic classes (NewHybrids classes) did not statistically influence the reproductive success of males whatever the considered variable (Figure S3, $p = 0.06$ for GLMM {2}). Fisher's exact test on NewHybrids classes showed that "parental" and "admixed" classes displayed very similar frequency among mated and unmated males (Figure S3, $p > 0.05$ for most tests). Two tests in Tv11 revealed that the parental classes (Pure-West or Pure-East) are over-represented in mated individuals, when compared to the Bx-West class (Figure 3B, + 31 % and $p < 0.05$, and + 66 % and $p < 10^{-2}$, for Pure-West and Pure-East, respectively).

These results are difficult to interpret because they involve (i) a single sample, (ii) a single Bx type, and (iii) a single measure of the genotypic classes (NewHybrids class). In addition, the opposite effect (Bx-West males significantly mate more than Pure-West males) was found to be marginally significant ($p = 0.036$ before correction for multiple testing) in Vi9 sample.

Impact of the admixed status of individuals on their survival time

The different classes of NewHybrids did not display a significant influence on the survival time of starving WCR adults (Figure S4, $p = 0.87$ for Cox's model {3}).

Impact of the genotypic type of individuals on the mating choices within WCR samples

Fisher's exact tests indicated that, in all samples, the genotypic classes (admixed or parental determined from the NewHybrid analyses) of sperm and that of the female mated

with that sperm were not statistically dependent of each other ($p > 0.05$ for all comparisons). Therefore, no choice of the sexual partner according to its hybrid status has been detected.

Conclusion

Globally, genotypic classes provided by this second method of genetic clustering lead to the same results obtained with the previous ones (the Structure classes) concerning the influence of the admixed status on traits positively correlated to the fitness. The admixed status has no effect and we can conclude that both methods are corroborant. Otherwise, even if admixed genotypic classes are different between both methods we can compare the parental ones. Are the assigned individuals to parental classes the same between both methods? In 85 % of cases they are (data not shown here).

FIGURES

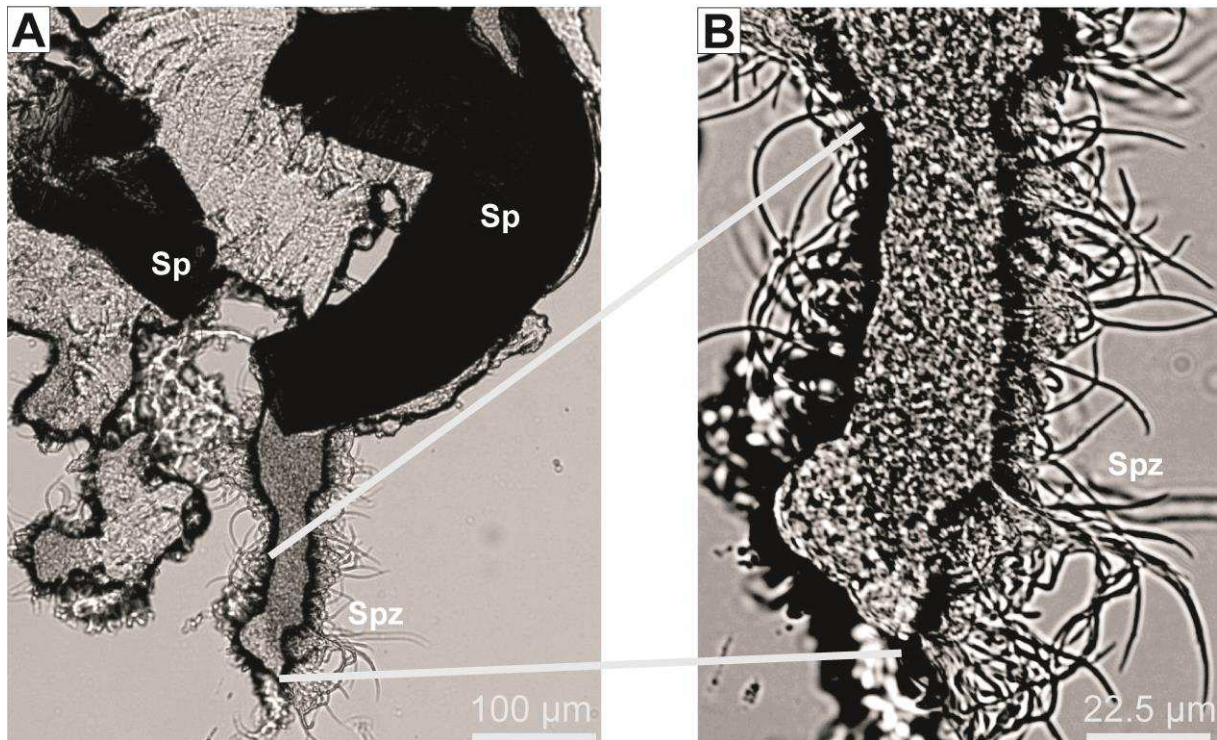


Figure S1: (A) Dissected spermatheca of *Diabrotica virgifera virgifera*. (B) Zoom on spermatozoa present in spermatheca. SP: spermatheca; SPZ: spermatozoa.

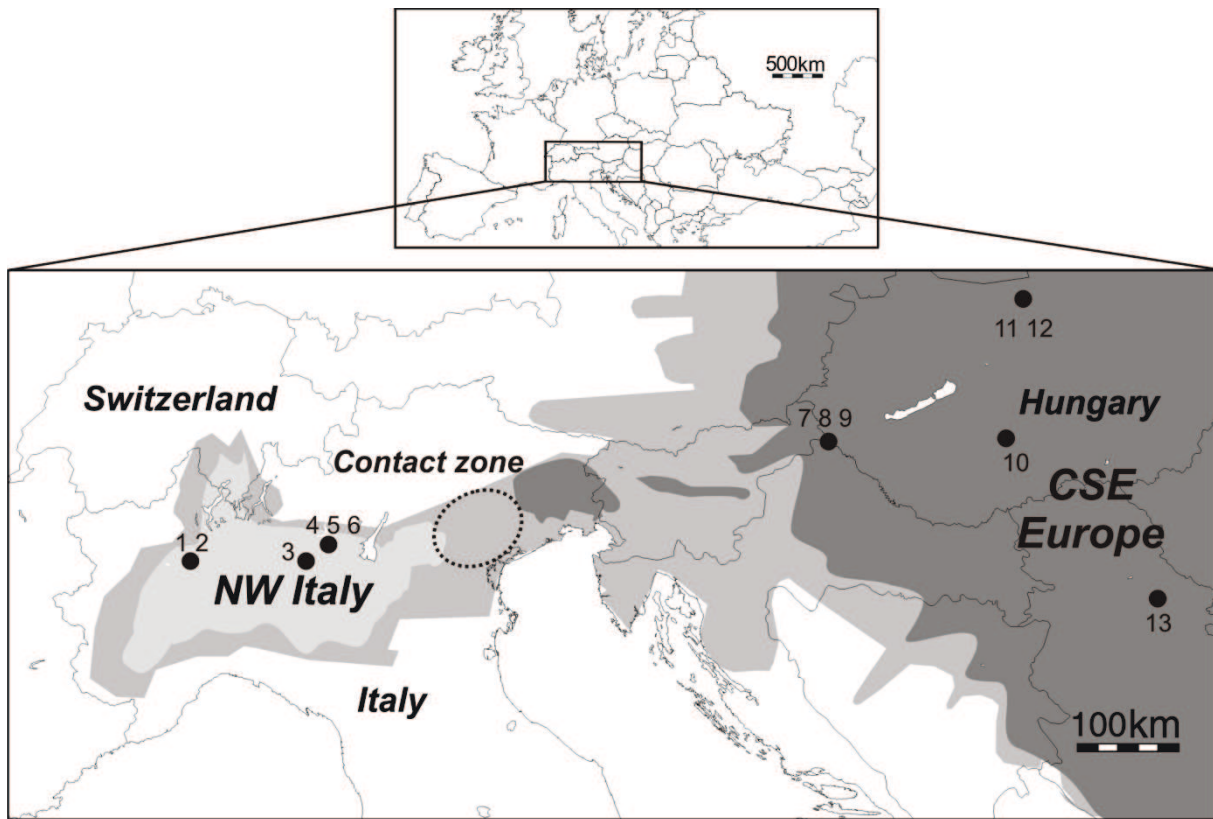


Figure S2: Location of the additional western corn rootworm sampling sites in North-West (NW) Italy and in Central South-Eastern (CSE) Europe. Numbers refer to sample names (see Table S1). The distribution areas of *Diabrotica virgifera virgifera* before contact between the western and eastern outbreak populations (in 2007) are shown in light gray (NW Italy) and dark gray (CSE Europe). The distribution area in 2012 (after contact) is shown in medium gray.

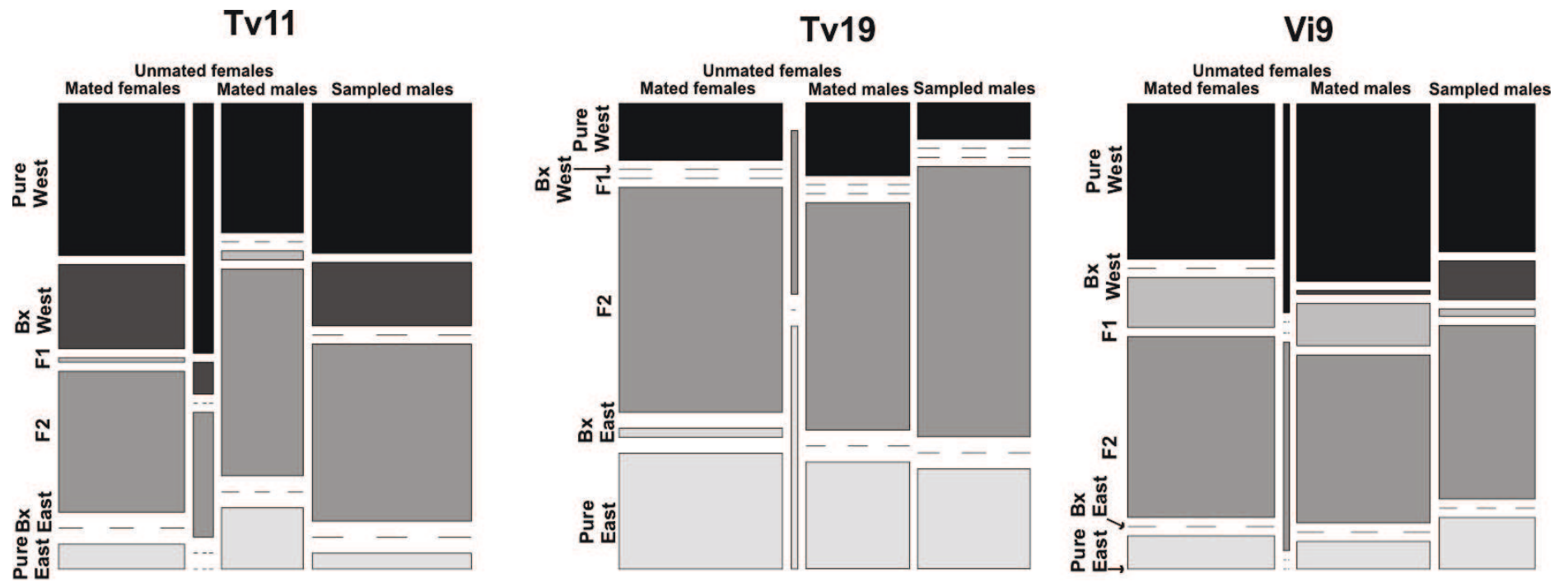


Figure S3: Admixed class distribution (Structure classes) of samples Tv11, Tv19 and Vi9 among mated and unmated females and among sperm and males (see text for details). Bx= backcross.

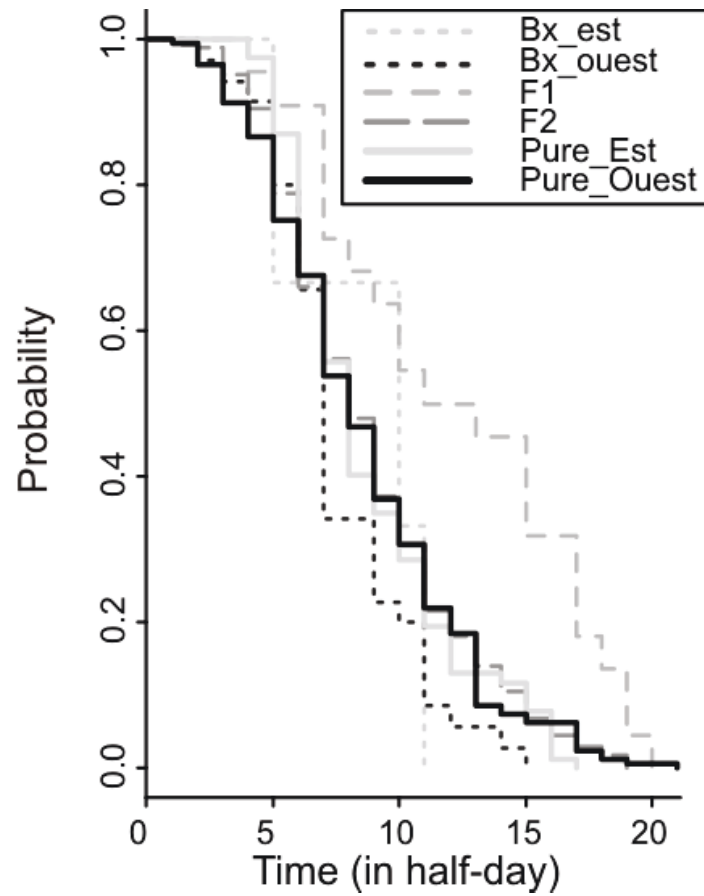


Figure S4: Survival probability of individuals belonging to the different classes of NewHybrids (Pure West, Pure-East, F1, F2, Bx-West, Bx-East) as a function of time (mixed Cox's models $p > 0.05$).

VII. – CHAPITRE IV : IMPACT DE L’HYBRIDATION ENTRE POPULATIONS ENVAHISSANTES DE *DIABROTICA VIRGIFERA VIRGIFERA* EN EUROPE SUR DES CARACTÈRES GÉNÉTIQUES ET PHÉNOTYPIQUES MESURÉS AU LABORATOIRE.

Comme indiqué précédemment (dans l’**Avant-propos**), ce chapitre repose sur des mesures réalisées en 2008, avant mon arrivée en thèse (2010), mais pas encore analysées. Le contexte global de cette étude (les conséquences évolutives de l’hybridation au cours des invasions biologiques) étant le même que celui du chapitre précédent et également largement développé en introduction de cette thèse, je ne m’y attarderai pas. En introduction de ce chapitre, je présenterai néanmoins le contexte particulier de l’étude, les objectifs visés et le protocole expérimental choisi. Dans un second temps, j’exposerai dans la section « Matériels et Méthodes » le design expérimental mis en place et je montrerai en quoi ce dernier ne permet pas un traitement statistique optimal des données. Enfin après avoir expliqué les différentes analyses que j’ai choisies d’effectuer pour y remédier, je présenterai les premiers résultats que j’ai obtenus.

INTRODUCTION

Dans cette étude, la chrysomèle des racines du maïs, *Diabrotica virgifera virgifera* (*Dvv*), a été utilisé comme modèle biologique pour étudier l'impact potentiel de l'hybridation sur la valeur sélective des individus au cours d'un cas particulier d'invasion biologique : son invasion européenne. Comme précédemment évoqué, les introductions répétées de cette peste en provenance des Etats-Unis ont conduit à la formation de deux foyers principaux, génétiquement différenciés : les foyers d'Italie du Nord-Ouest (Italie NO) et d'Europe Centrale et du Sud-Est (Europe CSE) (Ciosi *et al.*, 2008). La perte de variabilité génétique par rapport à la population source et associée aux goulots d'étranglement, atteint environ 30 à 40% au sein de chacun des deux foyers (Ciosi *et al.*, 2008). Récemment, ces derniers ont convergé et formé une zone hybride que l'on pense être neutre dans le Nord de l'Italie, en Vénétie (Bermond *et al.*, 2013; Bermond *et al.*, 2012). Malgré une restauration avérée des niveaux de variabilité génétique neutre dans certaines populations vénitiennes (Bermond *et al.*, 2012), le statut hybride ne semble pas avoir d'incidence particulière sur des caractères corrélés à la valeur sélective des individus (**Chapitre III**). Dans cette précédente étude, nous n'avons analysé que certains caractères (la survie à jeun, le succès d'accouplement, le régime de reproduction (homo ou hétérogamie)) positivement liés à la valeur sélective mais d'autres sont aussi intéressants et n'ont pas encore été analysés (la fécondité, la survie des individus du stade œuf jusqu'à l'âge adulte...). Par ailleurs, l'échantillonnage sur le terrain ne garantit pas la présence d'individus F_1 parmi ceux récoltés et ne permet donc pas d'accéder à l'hétérosis (dont l'effet est maximal pour lors la première génération d'hybride (F_1), voir **section I.2.2.b**).

L'objectif de cette étude était donc d'analyser et comparer la diversité génétique neutre (mesurée sur marqueurs microsatellites), les valeurs moyennes et les variances de plusieurs caractères phénotypiques entre populations parentales et hybrides. Pour cela, un évènement d'hybridation a été réalisé au laboratoire pour mimer l'hybridation réelle en cours

in natura et entre les foyers d'Italie NO et d'Europe CSE. Compte tenu de la difficulté liée à l'élevage de *Dv*v au laboratoire, une seule génération d'hybridation a été réalisée. Les comparaisons ont donc porté sur deux types de populations : des populations parentales (Italie NO et Europe CSE) et des populations hybrides (qui diffèrent selon le sens de croisement et l'origine du parent Europe CSE, voir Figure 13). Les résultats obtenus permettront de déterminer si le croisement de ces populations génétiquement différenciées a une influence sur le potentiel adaptatif de *Dv*v et rend ainsi son invasion encore plus problématique. Les questions sous-jacentes étaient :

1) L'évènement d'hybridation simulé au laboratoire permet-il d'augmenter la variabilité génétique neutre des populations hybrides par rapport à celle des populations parentales ?

2) Y a-t-il un impact de l'hybridation sur les différents caractères phénotypiques mesurés ?

3) Cela se traduit-il par des changements significatifs de moyenne ou de variance des caractères phénotypiques mesurés au sein des populations hybrides ?

MATÉRIELS ET MÉTHODES

Origine, croisements et élevage de masse de *Dv*v

Des adultes issus des foyers d'Europe CSE et Italie NO ont été collectés directement sur le maïs (les épis ont été secoués au-dessus d'un entonnoir sur lequel était fixé un sac en mousseline) dans des champs hautement infestés. Nous désignerons cette génération comme la génération F_0 . Deux origines différentes ont été sélectionnées en Europe CSE (Crvenka au Nord de la Serbie et Kondoros au Sud de la Hongrie) contre une seule en Italie NO (Como dans la région de la Lombardie, Figure 14).

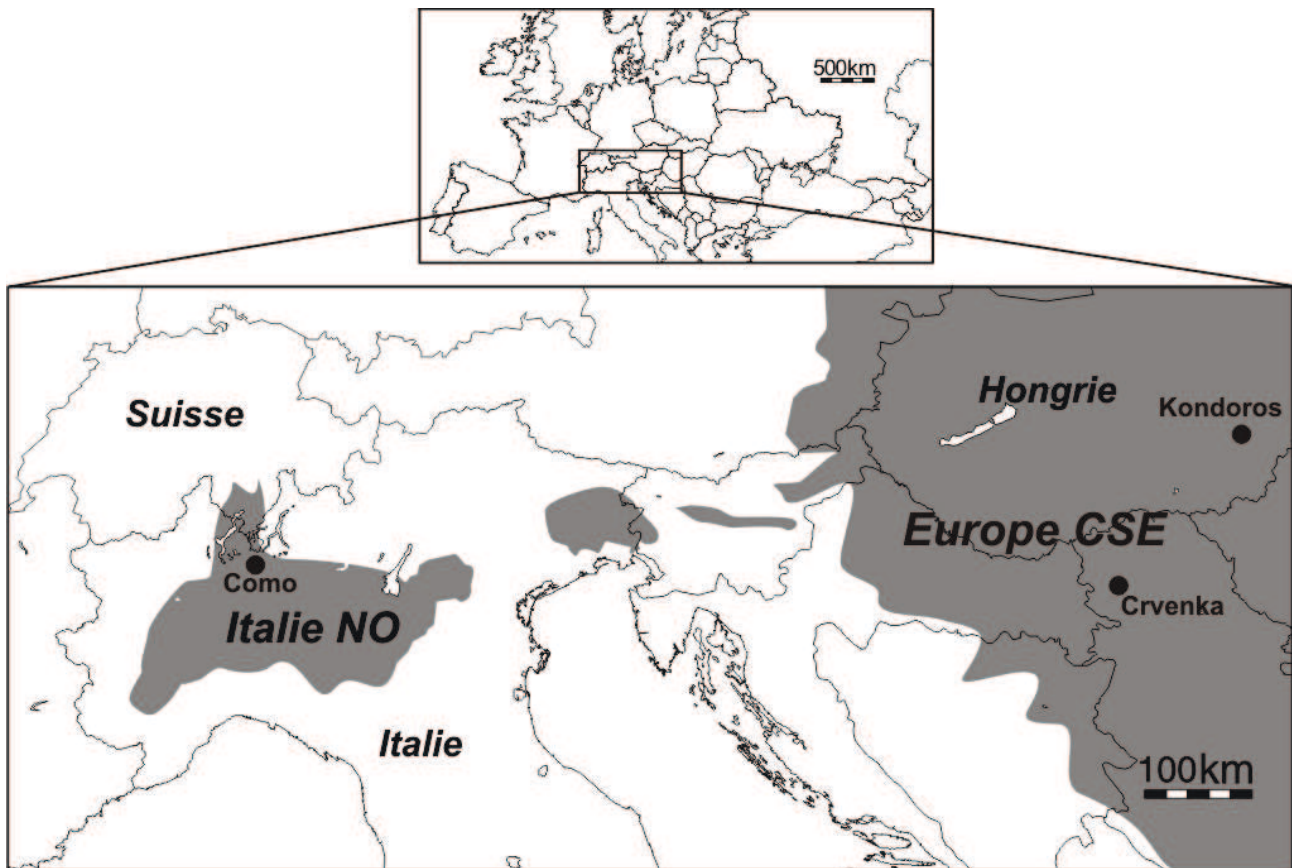


Figure 14 : Localisation des deux principaux foyers européens envahissants de *Diabrotica virgifera virgifera* (*Dvv*) en Italie du Nord-Ouest et en Europe Centrale et du Sud-Est (Europe CSE) et localisation des échantillons de *Dvv* collectés en 2006 en Italie NO et Europe CSE.

Afin de limiter les effets maternels des différentes populations échantillonnées, une génération d'élevage au laboratoire a été effectuée. Dans un premier temps, les adultes collectés sur le terrain ont été ramenés au laboratoire pour procéder à un élevage de masse. Chaque origine (Italie NO, Europe CSE Serbie, Europe CSE Hongrie) fut à cette étape isolée des autres. Les adultes furent ensuite autorisés à pondre et les œufs récoltés ont été stockés jusqu'à leur transformation en larves puis en adultes dans des conditions adéquates (Branson, 1976; Krysan, 1976; Krysan, 1972; Krysan, 1982). À partir de la nouvelle génération d'adultes obtenus, différents croisements entre Italie NO et Europe CSE ont été effectués pour obtenir tous les types d'hybrides possibles (origine du foyer Europe CSE et sens de croisement, cf Figure 15). Parallèlement à ces croisements, notons que, parmi les deux origines géographiques en Europe CSE, seule la souche d'origine Serbe a été maintenue en

élevage. Ainsi pour cette nouvelle génération que nous désignerons comme la génération F_1 , six types de croisement ont été obtenus : 2 croisements parentaux (Italie NO et Europe CSE (Serbie)) et 4 croisements hybrides (Italie NO X Hongrie ; Hongrie X Italie NO, Italie NO X Serbie et Serbie X Italie NO). [Par convention, l'origine à gauche de « X » au sein des différents croisements correspond toujours au genre féminin. Pour plus de commodité nous utiliserons par la suite, uniquement « Serbie » ou « Hongrie » pour caractériser les origines du foyer Europe CSE au sein des différents croisements hybrides]. Pour chaque type de croisement, plusieurs couples (entre 100 et 200) ont été isolés durant 5 jours pour permettre la fécondation puis ont été regroupés dans de grosses cages à populations (45 x 45 x 60 cm) et autorisés à pondre. Selon le même principe que précédemment, les œufs ont été récoltés et amenés jusqu'à l'âge adulte. C'est sur cette nouvelle génération d'individus, désignée comme la génération F_2 , que la majorité des différents caractères phénotypiques a été mesurée. La fécondité réalisée des femelles F_2 a été mesurée en comptant le nombre d'œufs observés en génération F_3 .

[Plus de détails sur les conditions d'échantillonnage et d'élevage de masse sont disponibles dans l'ANNEXE située à la fin de ce chapitre.]

Mesure de caractères phénotypiques

Les différents caractères phénotypiques mesurés se sont déclinés en caractères morphométriques (masse corporelle fraîche, longueur des élytres, largeur des élytres), en caractères liés à la valeur sélective des individus (fécondité réalisée, survie des œufs pendant l'hiver, survie des œufs jusqu'à l'âge adulte, valeur sélective globale, durée de vie adulte) et en caractères d'activité (proportion d'adultes volants, temps moyen d'attente jusqu'au décollage de l'insecte). Ces caractères ont été choisis parce qu'ils sont liés à la valeur sélective de manière directe (fécondité réalisée, survie des œufs pendant l'hiver, survie des œufs jusqu'à l'âge adulte, valeur sélective globale, durée de vie adulte) ou indirecte (masse

corporelle fraîche, longueur des élytres, largeur des élytres) ou parce qu'ils reflètent la capacité de *Dvv* à disperser et envahir (proportion d'adultes volants, temps moyen d'attente jusqu'au décollage de l'insecte) (Li *et al.*, 2010). [J'invite le lecteur à se rapporter à l'ANNEXE située à la fin de ce chapitre pour avoir plus de détails sur la mesure de ces caractères].

Toutes les manipulations effectuées au sein de cette étude ont été répétées une fois. Deux blocs expérimentaux ont donc été mis en place.

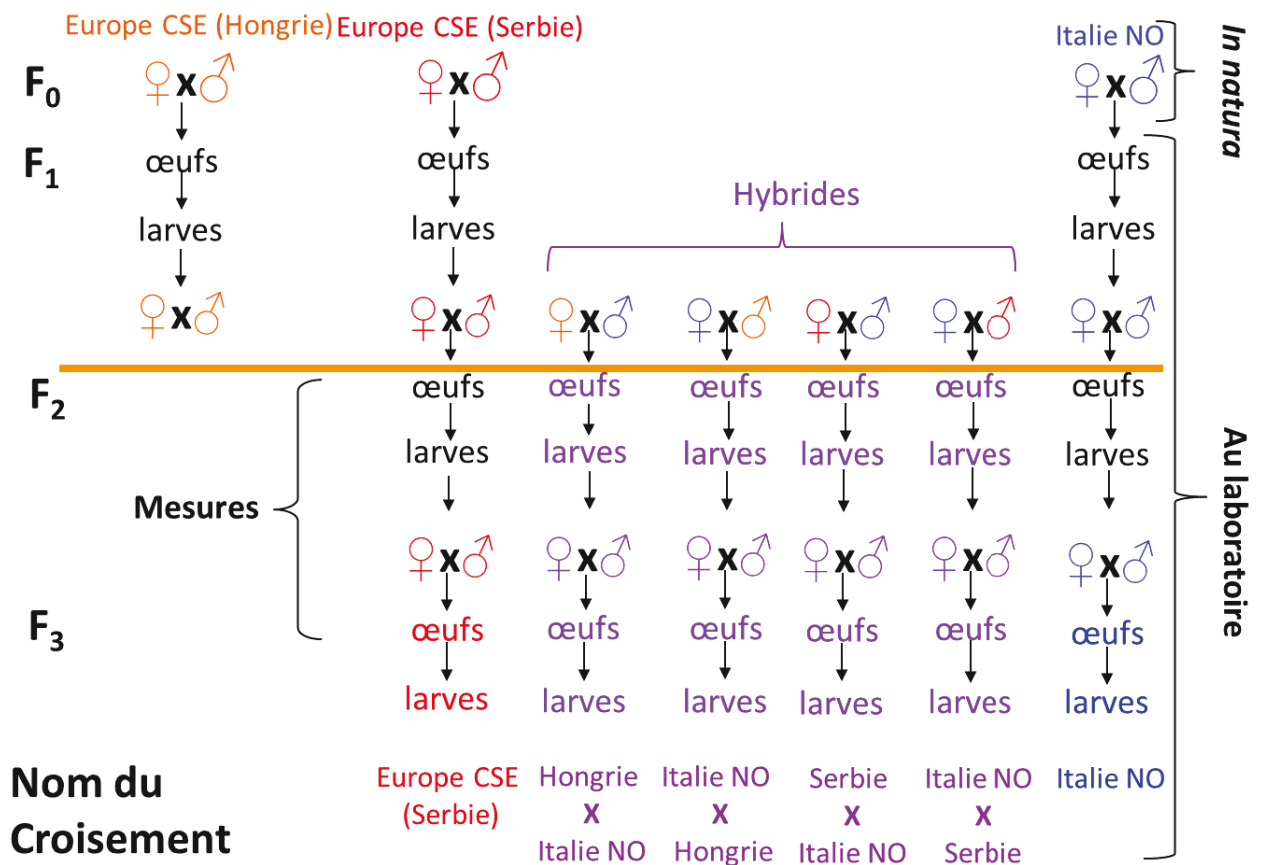


Figure 15 : Design expérimental de l'étude. Les populations échantillonnées en Europe CSE (Hongrie), Europe CSE (Serbie) et Italie NO sont respectivement représentées en orange, rouge et bleu. Les hybrides sont représentés en violet quel que soit le croisement réalisé. F₀, F₁, F₂, F₃ représentent les différentes générations utilisées au sein de cette étude.

La contribution du laboratoire à ce travail

Les données collectées au sein de chaque type de croisement et mesurées sur les différents caractères phénotypiques ne permettent pas à ce stade un traitement statistique optimal. La variance intra-croisement des caractères mesurés comporte une composante intra-famille qui ne nous intéresse pas et une composante inter-famille qui nous intéresse. Cette variance inter-familles intra-croisement est la variance génétique intra-croisement. Les croisements ont été réalisés de telle sorte que les descendance des couples parentaux (de la génération F_1) ont été regroupées par type de croisement. Les descendance (*i.e.* les familles) sont donc indiscernables en l'état et nous ne pouvons donc pas séparer variance intra et inter-familles. Si l'on parvient à tenir compte de cet effet famille, nous pourrons alors correctement estimer la variabilité intra-croisement inter-familles et ainsi prendre en compte un proxy de la variabilité génétique de chacun des croisements.

Notre laboratoire a été chargé du génotypage des individus et de l'analyse statistique des données permettant de faire le lien entre données microsatellites et données phénotypiques. En plus d'une analyse simple consistant à vérifier que les souches hybrides présentent une plus forte variabilité génétique neutre que les souches pures, nous avons choisi de traiter les données sous un aspect de « génétique quantitative » : les marqueurs génétiques serviront ici à tenir compte de l'effet famille dans les croisements. Nous avons ainsi entrepris une analyse de reconstruction des familles sur la base des génotypes multilocus des individus.

Estimation de la diversité génétique neutre

Extraction ADN et analyses microsatellites

Tous les échantillons de *Dvv* ont été stockés dans de l'éthanol absolu jusqu'à l'extraction de l'ADN. Les individus ont ensuite été rincés dans trois bains de NaCl à 0.065 %

puis coupés de manière longitudinale. L'extraction de l'ADN de chaque individu ne s'est faite que sur une seule moitié du corps de l'individu à l'aide du Kit Dneasy[®]Blood& Tissue (Qiagen, Hilden, Allemagne), suivant les recommandations du fabricant. Le génotypage des individus a enfin été mené à l'aide de 13 marqueurs microsatellites amplifiés *via* trois PCR multiplexées décrites par Bermond et al. (2012).

Variabilité génétique neutre au sein des échantillons de *Dvv*

La variabilité génétique neutre au sein des échantillons a été résumée en déterminant le nombre moyen d'allèles par locus (A) et l'hétérozygotie moyenne attendue (He) (Nei, 1987). Les variables A et He ont été calculées avec GENECLASS 2 ver 2.0.h (Piry *et al.*, 2004). Pour comparer les valeurs de A entre différents échantillons de populations, nous avons estimé la richesse allélique (RA) basée sur l'échantillon de plus petite taille en utilisant la méthode de raréfaction (Petit *et al.*, 1998) implémentée dans Fstat 2.9.3 (Goudet, 2001). Les F_{IS} ont été calculés avec « Genepop on the web » (Raymond, Rousset, 1995b). Les écarts à l'équilibre d'Hardy-Weinberg ont été testés par l'approche du test de probabilité utilisée dans « Genepop on the Web » (aRaymond, Rousset, 1995a)

La variabilité génétique neutre a ensuite été analysée en tenant compte de la structure en famille des individus étudiés. L'idée générale était de se concentrer sur la variabilité génétique neutre inter-familles. Cent échantillons de chaque croisement ont été générés en échantillonnant au hasard un individu par famille reconstruite avec COLONY. Cette procédure de ré-échantillonnage a été réalisée avec R (R Development Core Team, 2012). Cent jeux de A , He , F_{IS} et probabilités associées au test d'équilibre d'Hardy Weinberg ont été calculés en utilisant Genepop 4.0 (Rousset, 2008) à partir des cent échantillons générés. Pour calculer RA , 7 (le nombre minimum de familles dans chaque croisement) familles par type de croisement ont été aléatoirement tirées et un seul individu a été tiré par famille.

L'homogénéité de *RA* et *He* entre échantillons a été testée avec des tests non-paramétriques de Friedman (avec le locus comme unité de répétition). Après avoir détecté une hétérogénéité globale, nous avons testé l'hypothèse d'une variabilité génétique égale entre chaque croisement hybride et parental (avec une plus grande variabilité génétique dans les populations hybrides comme hypothèse alternative) avec des tests de rangs signés de Wilcoxon, unilatéraux. Le niveau de significativité de ces tests a ensuite été corrigé avec la procédure de Benjamini and Hochberg (1995) pour tenir compte du nombre de comparaisons.

Reconstruction de familles

Pour chaque type de croisement, les femelles accouplées ont été regroupées de telle sorte que les descendants de différentes femelles se sont retrouvés mélangés dans la même cage. Dans le but d'analyser correctement la variabilité phénotypique, nous avons pris en compte la structure familiale au sein de chaque type de croisement. Les marqueurs microsatellites ont été utilisés pour reconstruire les relations de parenté entre individus et nous ont ainsi permis de reconstruire des fratries et ainsi de considérer l'effet famille au sein des différents croisements. La reconstruction des familles a été réalisée grâce au logiciel COLONY ver 2.0.1.1 (Wang, 2004). Ce dernier implémente une méthode de maximum de vraisemblance permettant de regrouper les individus en fratries sur la seule base de leur génotype multilocus à des marqueurs microsatellites. Les configurations de fratries les plus vraisemblables sont obtenues grâce à l'utilisation d'un algorithme de recuit simulé, une méthode de Monte Carlo similaire aux méthodes de Monte-Carlo par chaînes de Markov (MCMC). La méthode développée par Wang prend également en compte les différents types d'erreurs de génotypage que l'on peut rencontrer lorsque l'on travaille avec des marqueurs microsatellites (dropout, mutations, contaminations, faux allèles...) et peut s'employer avec ou sans connaissance des génotypes parentaux. Dans notre cas, nous n'avons pas accès aux

génotypes parentaux et quatre runs indépendants ont été lancés pour chaque type de croisement. Les fratries ont été définitivement constituées après avoir observé des regroupements d'individus consensus entre les différents runs.

Analyses statistiques des caractères phénotypiques

La structure familiale fut prise en compte dans nos analyses statistiques grâce à l'utilisation de modèles linéaires mixtes (pour une erreur distribuée normalement), de modèles linéaires généralisés mixtes (GLMMs, pour toute autre distribution de l'erreur) ou de modèles de Cox mixtes (adéquats pour des données censurées comme la « durée de vie adulte » et « le temps moyen avant le décollage de l'insecte »), avec l'effet « famille » comme effet aléatoire. De cette manière, l'éventuelle variance génétique associée à chacun des caractères est prise en compte.

Compte tenu de la complexité des analyses à mener, du nombre de variables explicatives disponibles et du nombre de variables à expliquer (les caractères mesurés), j'ai cherché à expliquer la variance de mes données par un modèle simple, qui a du sens d'un point de vue biologique et qui me permet de déterminer si l'hybridation a un impact sur la moyenne et la variance des caractères liés à la valeur sélective des individus. J'ai pour cela choisi un modèle générique, utilisable pour chacun des caractères, avec des facteurs fixes déterminés en amont de l'analyse statistique. La stratégie adoptée ici ne correspond pas tout à fait à une démarche classique de sélection de modèles. Les variables explicatives « Traitement » (avec les modalités Italie NO, Europe CSE (Serbie) et Hybrides), « Sexe » (avec les modalités Femelles et Mâles) et « Bloc » (avec les modalités 1 et 2) ont donc été traitées comme effets fixes. Notons que le « Bloc » aurait dû être considéré comme une variable aléatoire mais face aux difficultés rencontrées nous ne l'avons pas traité comme telle. Enfin, nous avons aussi cherché à déterminer pour chacun des caractères, si la variance inter-familles (le proxy de la variance génétique) variait ou pas entre les modalités de la variable

« Traitement ». Pour ce faire j'ai comparé pour chacun des caractères un modèle possédant une structure aléatoire emboîtée « Traitement|famille » avec un modèle possédant une structure aléatoire simple « 1|famille » (Tableau 5). Pour certains caractères, un seul sexe est concerné par les mesures phénotypiques (fécondité, survie des œufs...). Je présente donc ci-dessous les deux types de modèles génériques permettant de couvrir l'ensemble des caractères étudiés :

Caractère phénotypique \sim Traitement*Sexe+Bloc+(1|famille) {1}

Caractère phénotypique \sim Traitement+Bloc +(1|famille) {2}

À partir de ces modèles préétablis nous avons cherché à (i) déterminer la meilleure structure aléatoire et sa significativité ; (ii) estimer et déterminer la significativité des effets fixes ; (iii) tester l'égalité des caractères pour toutes les comparaisons « parents versus hybrides ».

Sur les 11 caractères phénotypiques initialement mesurés, 10 (masse corporelle fraîche, longueur des élytres, largeur des élytres, fécondité réalisée, survie des œufs pendant l'hiver, survie des œufs jusqu'à l'âge adulte, valeur sélective globale, durée de vie adulte, proportion d'adultes volants, temps moyen avant le décollage de l'insecte) ont pu être analysés statistiquement. La variable composite « valeur sélective globale » a été calculée de la sorte (voir ANNEXE) :

Valeur sélective globale = survie des larves jusqu'au stade adulte*sex-ratio des adultes*fécondité réalisée*survie des œufs pendant l'hiver

Notons que la procédure expérimentale n'a permis de mesurer qu'une valeur moyenne de la variable « survie des larves jusqu'au stade adulte » par « Traitement » et par « Bloc ».

Par ailleurs, seule la structure aléatoire « 1|famille » a été considérée pour le caractère « temps moyen avant le décollage de l'insecte » car nous n'avons pas trouvé d'outil statistique

permettant de prendre en considération des structures aléatoires emboîtées telles que « Traitement|famille » pour des données censurées.

Excepté les caractères « proportion d'adultes volants» (caractérisé par une distribution binomiale de l'erreur et donc analysé grâce un GLMM) et « durée de vie adulte » (analysé grâce à un modèle de Cox) tous les autres ont été analysés grâce à des modèles linéaires mixtes (avec distribution normale de l'erreur). Pour cela, les distributions des caractères « fécondité réalisée » et « survie des œufs jusqu'à l'âge adulte » ont été rapprochées d'une distribution normale avec une transformation racine carrée et le caractère « valeur sélective globale » a été rapproché d'une distribution normale en utilisant une transformation logarithmique.

La surdispersion des données a été testée pour le caractère « proportion d'adultes volants» en calculant Φ , le paramètre de surdispersion des données :

$$\Phi = \frac{DR}{ddlR} ; \text{ avec } DR = \text{ la déviance résiduelle du modèle et } ddlR = \text{ le nombre de degrés de}$$

liberté résiduels.

RÉSULTATS

Variabilité génétique neutre au sein des échantillons de *Dvv*

Les analyses de reconstruction de familles obtenues à partir des génotypes multilocus à 13 microsatellites ont conduit à la formation de 7 à 28 familles par croisement dans les deux blocs expérimentaux (moyenne =18.5 ; écart-type =7.9 ; voir Tableau 1).

Tableau 1 : Nombre de familles reconstruites par croisement et par bloc.

Populations	Bloc	Nombre de familles
Europe CSE (Serbie)	1	23
	2	20
Italie NO	1	7
Hongrie X Italie NO	1	23
	2	26
Italie NO X Hongrie	1	20
	2	25
Serbie X Italie NO	1	28
	2	17
Italie NO X Serbie	1	20
	2	13

Parce qu'il n'y avait pas de différence génétique neutre entre les blocs expérimentaux, la variation génétique neutre a été calculée après avoir regroupés les individus des 2 blocs pour chaque croisement. Aucun autre regroupement n'a été effectué. Les quatre croisements hybrides n'ont pas été regroupés pour les analyses ultérieures en raison d'une différenciation génétique mesurée sur marqueurs microsatellites (données non présentées ici). Les échantillons hybrides et parentaux (Italie NO, Europe CSE (Serbie)) ont montré un polymorphisme modéré avec une moyenne de 6.08 allèles par locus (écart-type = 4.01). Le nombre moyen d'allèles variait de 3.23 pour Italie NO à 5.20 pour Italie NO X Hongrie (Tableau 2). L'hétérozygotie attendue était modérée et variait de 0.42 pour Italie NO à 0.51 pour Italie NO X Hongrie et Hongrie X Italie NO (Tableau 2).

Tableau 2 : Statistiques résumant la variation génétique au sein de chaque croisement de *Diabrotica virgifera virgifera*

Populations échantillonnées	<i>N</i>	<i>A</i>		<i>He</i>	<i>p-HW</i>	<i>F_{IS}</i>
		Dénombrement directs (<i>DD</i>)	Richesse Allélique (<i>RA</i>)			
Italie NO	37	3.23 (1.83)	3.23 (1.83)	0.42	0.00	-0.06
Europe CSE (Serbie)	132	3.46 (1.89)	3.00 (1.42)	0.44	<10-3	-0.04
Hongrie X Italie NO	140	4.53 (3.17)	4.15 (2.66)	0.51	<10-3	-0.15
Italie NO X Hongrie	142	5.23 (3.67)	4.57 (2.68)	0.51	<10-3	-0.08
Italie NO X Serbie	142	4.69 (2.86)	4.08 (2.36)	0.50	<10-3	-0.09
Serbie X Italie NO	116	4.30 (3.09)	4.03 (2.61)	0.50	<10-3	-0.13

Note : Populations Parentales = Italie NO et Europe CSE (Serbie); Populations hybrides = Hongrie femelles X Italie NO males; Italie NO femelles X Hongrie males; Italie NO femelles X Serbie males and Serbie femelles X Italie NO males. *A* : nombre moyen d'allèles par locus. *A* est donné par des dénombrements directs (*DD*) et la richesse allélique (*RA*). *RA* est basée sur la plus petite taille d'échantillon (*N* = 37 pour plusieurs loci dans l'échantillon italien). *He* : hétérozygotie moyenne attendue (Nei, 1987). *F_{IS}*: corrélation de fréquence de gènes entre individus par population (Weir, Cockerham, 1984). *p-HW* : probabilités associées aux tests exacts d'écart à l'équilibre d'Hardy-Weinberg. Les écart-types entre loci sont montrés entre parenthèses.

Aucune différence significative d'hétérozygotie attendue (*He*) n'a été observée (test de Friedman, $\chi^2 = 6.10$; ddl = 5; $p = 0.30$), sauf pour la comparaison entre Italie NO X Serbie et Europe CSE (Serbie) ou une plus grande hétérozygotie a été trouvée pour le croisement hybride (Tableau 3). Des différences significatives de richesse allélique (*RA*) ont été détectées entre les échantillons (test de Friedman, $\chi^2 = 23.96$; ddl = 5; $p = 0.00022$). Nous avons trouvé un nombre beaucoup plus grand d'allèles dans les croisements hybrides que dans les croisements parentaux (Tableau 3). Un excès d'hétérozygotie (*F_{IS}* négatifs dans le Tableau 2) et des tests exacts d'écart à l'équilibre d'Hardy-Weinberg significatifs (Tableau 2) ont été trouvés pour tous les croisements.

Tableau 3 : Tests de comparaison de la variabilité génétique entre échantillons quand l'effet famille est pris ou non en considération

Avec ou sans effet famille	Croisement	<i>He</i>		<i>RA</i>	
		Croisements parentaux		Croisements parentaux	
		Italie NO	Europe CSE (Serbie)	Italie NO	Europe CSE (Serbie)
Sans	Hongrie X Italie NO	0.22	0.07	0.02	0.00
	Italie NO X Hongrie	0.10	0.05	0.00	0.00
	Italie NO X Serbie	0.28	0.04	0.01	0.01
	Serbie X Italie NO	0.36	0.10	0.03	0.01
Avec	Hongrie X Italie NO	0.19 (0.13) [0]	0.00 (0.00) [0.56]	0.26 (0.21) [0]	0.12 (0.14) [0.01]
	Italie NO X Hongrie	0.29 (0.16) [0]	0.04 (0.04) [0.14]	0.18 (0.22) [0]	0.13 (0.19) [0.01]
	Italie NO X Serbie	0.44 (0.21) [0]	0.06 (0.07) [0.10]	0.39 (0.28) [0]	0.17 (0.20) [0.01]
	Serbie X Italie NO	0.39 (0.20) [0]	0.07 (0.05) [0.08]	0.34 (0.25) [0]	0.15 (0.16) [0.01]

Note : Sans effet famille : probabilités associées aux tests des rangs signés de Wilcoxon permettant de tester une égalité d'hétérozygotie moyenne attendue (*He*) et de richesse allélique (*RA*) par paire d'échantillons. Seules les comparaisons entre croisements hybrides et parentaux sont représentées dans cette table. Les différences significatives après correction de Benjamini and Hochberg (1995) sont montrées en gras. Avec l'effet famille : moyenne des probabilités (sur 100 échantillons aléatoires, voir le texte pour plus de détails) des tests des rangs signés de Wilcoxon permettant de tester une égalité d'hétérozygotie moyenne attendue (*He*) et de richesse allélique (*RA*) par paire d'échantillon. Les écart-types associés sont montrés entre parenthèses. La proportion de probabilités significatives après correction de Benjamini and Hochberg (1995) est montrée entre crochets.

L'ensemble du jeu de données ré-échantillonnées, qui prend en compte la structure familiale, montre un faible polymorphisme avec un nombre moyen d'allèles variant de 2.89 pour Italie NO à 4.94 pour Italie NO X Hongrie (Tableau 4). L'hétérozygotie moyenne attendue est modérée et varie de 0.44 pour Italie NO à 0.59 pour Hongrie X Italie NO (Tableau 4). Les plus grandes hétérozygoties et richesses alléliques moyennes ont été observées dans des échantillons hybrides. En considérant la structure familiale dans le jeu de données, l'hétérozygotie et la richesse allélique étaient significativement plus élevées dans les échantillons hybrides que dans les échantillons parentaux dans respectivement 11 et 0.5% des cas (valeurs entre crochets dans le Tableau 3). Notez que RA a été estimée pour un très petit nombre d'individus (7), car un croisement contenait uniquement 7 familles. En général, les tests exacts d'écart à l'équilibre d'Hardy-Weinberg ne sont pas significatifs.

Les résultats obtenus sur les jeux de données ré-échantillonnés semblent montrer que les croisements réalisés entre les deux populations parentales (Italie NO et Europe CSE Serbie) ont permis d'augmenter la variabilité génétique neutre chez les hybrides (He et RA) (Tableau 4) du moins en partie (dans respectivement 11 et 0.5 % des cas). Pour quantifier plus précisément cette augmentation, nous aurions pu comparer les distributions des proportions des tests de Wilcoxon significatifs montrées entre crochets dans le Tableau 3 avec celles obtenues en intra-classe hybride (comparaisons entre hybrides uniquement) sous l'hypothèse nulle que la diversité génétique (He et RA) est égale entre les différents croisements hybrides.

Tableau 4 : Statistiques résumant la variation génétique au sein du jeu de données ré-échantillonné de *Diabrotica virgifera virgifera*

Populations ré-échantillonnées	<i>N</i>	<i>A</i>		<i>He</i>	<i>p-HW</i>	<i>F_{IS}</i>
		<i>DD</i>	<i>RA</i>			
Italie NO	7	2.89 (0.07)	2.90 (0.06)	0.44(0.03)	0.85 (0.15)	0.01 (0.06)
Europe CSE (Serbie)	43	3.29 (0.06)	2.70 (0.12)	0.46 (0.01)	0.41 (0.22)	-0.02 (0.02)
Hongrie X Italie NO	49	4.31 (0.09)	3.33 (0.18)	0.59 (0.01)	<10 ⁻³ (<10 ⁻³)	-0.14 (0.01)
Italie NO X Hongrie	46	4.94 (0.12)	3.46 (0.25)	0.55 (0.01)	0.01 (0.02)	-0.07 (0.01)
Italie NO X Serbie	33	4.39 (0.08)	3.22 (0.21)	0.54 (0.01)	0.26 (0.22)	-0.08 (0.02)
Serbie X Italie NO	45	4.16 (0.05)	3.26 (0.15)	0.54 (0.01)	0.14 (0.15)	-0.07 (0.01)

Note : Populations Parentales = Italie NO et Europe CSE (Serbie); Populations hybrides = Hongrie femelles X Italie NO males; Italie NO femelles X Hongrie males; Italie NO femelles X Serbie males and Serbie femelles X Italie NO males. *A* : nombre moyen d'allèles par locus. *A* est donné par des dénombrements directs (*DD*) et la richesse allélique (*RA*). *RA* est basée sur la plus petite taille d'échantillon (*N* = 37 pour plusieurs loci dans l'échantillon italien). *He* : hétérozygotie moyenne attendue (Nei, 1987). *F_{IS}* : corrélation de fréquence de gènes entre individus par population (Weir, Cockerham, 1984). *p-HW* : probabilités associées aux tests exacts d'écart à l'équilibre d'Hardy-Weinberg. Les écart-types entre loci sont montrés entre parenthèses.

Analyses statistiques des caractères phénotypiques

Concernant les valeurs moyennes des caractères phénotypiques mesurés

L'effet « bloc » est significatif pour tous les caractères excepté la fécondité. Globalement, les valeurs moyennes obtenues dans le bloc 2 sont supérieures à celles du bloc 1. Le contraire a été constaté pour les caractères « survie des œufs pendant l'hiver et durée de vie adulte ». Ces différences sont présentées ici car le « bloc » a été traité comme effet fixe mais elles ne nous intéressent à proprement parler pour répondre aux questions posées. Elles sont vraisemblablement dues à une accumulation de différences (dans les gestes ou les manipulations réalisés) entre les deux séries expérimentales et générées par l'expérimentateur.

L'effet sexe est toujours significatif lorsqu'il est considéré et indique que les femelles sont toujours supérieures aux mâles excepté pour les caractères d'activité « proportion d'adultes volants et temps moyen avant le décollage de l'insecte ». Ces résultats sont cohérents avec ce que l'on peut observer dans la nature, les femelles sont habituellement plus grandes et plus volumineuses que les mâles surtout si elles sont gravides. Par ailleurs, il n'est pas étonnant d'observer que les mâles ont une activité supérieure à celle des femelles car à l'état naturel, ces derniers passent la majeure partie de leur temps à la recherche de partenaires sexuelles.

Les résultats (Tableau 5) montrent que l'effet traitement est significatif pour la majorité des caractères mesurés (7/10) « masse corporelle fraîche, longueur et largeur des élytres, survie des œufs pendant l'hiver, survie des œufs jusqu'au stade adulte, valeur sélective globale et temps moyen avant décollage de l'insecte ».

Tableau 5 : Résultats obtenus à partir du modèle linéaire (généralisé) mixte générique pour chacun des caractères phénotypiques mesurés

Facteurs explicatifs fixes	Masse corporelle fraîche (mg)	Longueur des élytres (mm)	Largeur des élytres (mm)	Fécondité réalisée (comptage)	Survie des œufs en hiver (%)	Survie œufs jusqu'à l'âge adulte (%)	Valeur sélective globale	Durée de vie adulte (jours)	Proportion d'adultes volants	Temps moyen avant le décollage de l'insecte (sec)
Traitement	<10 ⁻³	<10 ⁻³	<0.05	0.31	<10 ⁻³	<10 ⁻³	<0.05	0.18	0.27	<0.05
Sexe	B>A~C	B>A~C	B~A ; A~C.B>C	A~B~C	A>B>C	A>B ; A>C ; B~C	A>B ; A>C ; B~C	A~B~C	A~B~C	A>C>B
Bloc	<10 ⁻³	<10 ⁻³	<10 ⁻³	-	-	-	-	<10 ⁻³	<10 ⁻⁵	<10 ⁻⁵
Traitement:sexe	F>M	F>M	F>M	0.65	<10 ⁻³	<10 ⁻³	<10 ⁻³	F>M	M>F	M>F
	<10 ⁻³	<0.05	<0.05	1~2	1>2	2>1	2>1	<10 ⁻⁵	<10 ⁻³	<10 ⁻⁵
	2>1	2>1	2>1	-	-	-	-	1>2	2>1	2>1
	0.26	0.76	0.94	-	-	-	-	0.19	0.71	0.25
Facteurs explicatifs aléatoires										
1 famille	3034 (s)	34 (s)	-653 (s)	2617 (ns)	2670 (ns)	633 (ns)	527.4 (ns)	36.7 (ns)	381.5 (s)	175.8 (s)
Traitement famille	3044	43	-644	2625	2678	638	536.5	-	391.3	-
Sur-dispersion (>1)	-	-	-	-	-	-	-	-	0.88	-

Note : Les chiffres de la partie haute du tableau (« Facteurs explicatifs fixes ») représentent les probabilités associées au modèle statistique générique utilisé pour chacun des caractères phénotypiques mesurés. Les chiffres de la partie basse du tableau (« Facteurs explicatifs aléatoires ») représentent d'une part les valeurs d'AIC du modèle générique, obtenues pour les différentes structure aléatoires testées (1|famille ou Traitement|famille) et pour chacun des caractères mesurés et d'autre part la valeur du coefficient de sur-dispersion Φ , mesuré pour le caractère « Proportion d'adultes volant ». Traitement : A= Italie NO, B= Europe CSE (Serbie) et C= Hybrides ; Sexe : M(mâles) et F(femelles) ; Bloc : 1 et 2 ; Traitement :sexe=:interaction entre le Traitement et le sexe ; Traitement|famille : permet d'estimer une variance « inter-familles » pour chaque modalité de « Traitement » ; 1|famille : permet d'estimer une seule variance « inter-familles » pour l'ensemble des données ; (s) indique si l'effet aléatoire est significatif ; (ns) indique qu'un effet aléatoire est non-significatif ; (-) indique que des données ne sont pas disponibles, par exemple pas d'effet « sexe » car les mesures ne sont effectuées que sur les femelles.

Plusieurs différences significatives ont également été trouvées entre les deux croisements parentaux. Globalement les valeurs moyennes des caractères morphométriques (« masse corporelle fraîche et longueur des élytres ») du croisement parental Europe CSE (Serbie) sont supérieures à celles du croisement parental Italie NO. Le contraire a été constaté sur les caractères liés à la valeur sélective des individus (« survie des œufs jusqu'à l'âge adulte et valeur sélective globale ») et sur un caractère d'activité (« temps moyen avant le décollage de l'insecte »).

Nous avons constaté que les valeurs moyennes des hybrides sont (i) équivalentes à celles du croisement parental le moins performant pour 5 d'entre eux ; (ii) inférieures à celles des deux croisements parentaux pour seulement un caractère : « survie des œufs pendant l'hiver » ; (iii) comprises entre celles de deux croisements parentaux pour seulement un caractère : « temps moyen avant le décollage de l'insecte ».

L'interaction entre les effets fixes « Traitement et Sexe » n'a jamais été significative.

Concernant la variance des caractères phénotypiques mesurés

L'effet aléatoire « Traitement|famille » n'a jamais été retenu sur l'ensemble des caractères mesurés et indique donc que les variances inter-familles intra-croisement ne sont pas statistiquement différentes entre croisements.

Par ailleurs l'effet aléatoire « 1|famille » est significatif pour 5/10 caractères : « masse corporelle fraîche, longueur des élytres, largeur des élytres, proportion d'adultes volants et temps moyen avant le décollage de l'insecte » et indique qu'une variance génétique (inter-familles inter-croisements) existe bien pour ces mêmes caractères.

Aucun signe de sur-dispersion ne fut détecté pour le caractère « proportion d'adultes volants ».

Les résultats obtenus dans cette étude seront discutés dans leur globalité avec ceux obtenus dans le **Chapitre III** (Absence d'impact de l'hybridation *in natura* au cours de l'invasion européenne de *Diabrotica virgifera*) dans la **section *Discussion Générale***.

ANNEXE

Li H., Guillemaud T., Kuhlmann U., Toepfer S

Origin, crossing, and mass rearing of *Diabrotica v. virgifera*

Experiments were conducted with four laboratory-reared *D. v. virgifera* populations (= here synonym for genotypic classes, (Johnston *et al.*, 2004)), *i.e.* the central and southeastern European population (CSE, as defined by Miller *et al.*, (2005)), the northwest Italian population (NWIT, as defined by Miller *et al.*, (2005)), and two crossings of both. Adults of the central and southeastern European population were collected from heavily infested maize fields in Crvenka in Severnobački County in northern Serbia and in Kondoros in Bekes County in southern Hungary. Adults from the northwest Italian population were collected from an infested maize field in Como in Lombardy County in northern Italy. In order to mass collect adults from the field, a funnel with an attached gauze bag was used. Adults were shaken from the maize plants into the funnel and bag. In the laboratory, adults were reared in gauzed cages (45 x 45 x 60 cm) according to (Krysan, 1986; Singh, Moore, 1999) (temperature L:D, 24-26:18-20 °C, 40 to 60 % relative humidity, light regime L:D, 14:10).

Adults were allowed to lay eggs into petridishes with moist, sterilised and sieved field soil (<200 micrometer mesh size) up to two months after collection. The petridishes were changed every two weeks. The soil with eggs was washed with hand warm tap water through a 0.25 mm sieve. The number of recovered eggs on the sieve was estimated. Eggs were then stored in sterile moist soil and kept at 24°C during two weeks for pre-diapause (Branson, 1976; Krysan, 1976; Krysan, 1972), and then at 6 to 8°C during another 5 to 7 months for diapause (Krysan, 1982). To reduce parental effects of canalization or exposure to their

previous environments in the fields, all pure populations were mass-reared under laboratory conditions during one generation before crossing.

Adults of the CSE and NWIT population were crossed in the laboratory to simulate the natural overlap of these two independent *D. v. virgifera* populations. Spontaneously crossed wild *D. v. virgifera* were not used for three reasons: First, it is uncertain whether both populations actually started to overlap and cross (M. Augusto, 2008, pers. comm.). Second, *D. v. virgifera* are difficult to mass collect because population levels in the suggested overlapping zone are extremely low due to eradication efforts (Furlan, Italy, 2008). Third, in case that crossing already naturally occurred, several generations of crossed generations might be present in maize fields making it difficult to classify or replicate genotypic classes in a meaningful way (Johnston *et al.*, 2004). In laboratory, newly emerged adults from the CSE population and the NWIT population were sexed according to antenna length (Kuhar, Youngman, 1995; Staetz *et al.*, 1976) and tarsus characteristics (Hammack, French, 2007). For each cross, several pairs were set up in which one CSE female from each origin (Serbia and Hungary) was crossed with one NWIT male or one NWIT female with one CSE male from each origin (Serbia and Hungary) leading to four distinct hybrid crosses. Single pairs were kept in small gauze-covered plastic cylinders (dia.: 120 mm, h: 140 mm) for five days. Two soft, unripe kernels of organically-produced maize, one 13 x 13 x 13 mm piece of zucchini flesh, one 13 x 13 x 13 mm piece of pumpkin flesh and a 5 x 5 x 5 mm piece of artificial pollen diet (Branson *et al.*, 1988; Singh, Moore, 1999) were provided as food, as well as a 10 x 5 x 5 mm cube of 15% water agar. After 5 days, female-male pairs were double-checked for correct sex identification to avoid mistakes in crossings. Then, about 100-200 crossed pairs were pooled and transferred into a larger gauze cage (45 x 45 x 60 cm) for egg laying under conditions as described above. Each type of considered population (both parental populations and the four hybrid crosses) was reared separately.

Each *D. v. virgifera* was reared under similar laboratory conditions, as per George & Ortman (1965) and Branson et al. (1975). Eggs were incubated at a temperature of L:D, 25:21°C for 14 to 24 days to initiate egg hatching. About 2,000 to 3,000 seeds of the maize hybrid Gavott (UFA Semences, Bussigny, Switzerland) were soaked in water for 24h and germinated in a plastic tray (l: 330 mm, w: 190 mm, h: 110 mm) with a gauze lid. Three days after germination, ready-to-hatch eggs were transferred onto the seeds in the plastic trays (approx. 5,000 eggs per tray) and maintained at a temperature of L:D, 25:21 °C, and light regime of L:D, 14:10. Emerging larvae found unlimited food (maize roots). After 14 to 20 days, third instar larvae were transferred along with the maize seedlings into gauze-covered cylinders (dia.: 120 mm, h: 140 mm) containing sterilised field soil for pupation (soil sieved at < 5 mm mesh size, 30 % moisture). The transferred maize still provided food for the larvae until entering the soil for pupation.

Assessing phenotypic traits

All phenotypic traits were measured on the second generation (F₂) of parental populations (NWIT and CSE-Serbia) and on the first generation of hybrid populations after crossing, excepted for the realized fecundity that was measured on eggs of the third generation (F₃). Totally, 11 phenotypic traits were assessed, i.e. six fitness traits (adult lifespan, fecundity, overwintering survival of eggs, survival of larvae and pupae until adulthood, survival from eggs to adulthood, overall fitness), two activity traits (proportion of adults flying, flight take-off response) and three morphometric traits (fresh body weight, elytra length, elytra width).

Fitness traits

In order to assess adult lifespan and fecundity, newly emerged male-female pairs were transferred into small bioassay containers (temperature L:D, 24:18 °C, light regime L:D,

14:10). The containers consisted of two plastic urinalysis cups (dia.: 48 mm, h: 80 mm), stacked one inside the other providing approximately 175 cm³ of space (for details see Li *et al.*, 2009). The upper cup had a 10 - mm hole in the bottom to give the female access to the lower, soil-filled cup for egg-laying. Abundant food was provided in each container (two soft, unripe kernels of organically-produced maize , one 13 x 13 x 13 mm piece of zucchini flesh, one 13 x 13 x 13 mm piece of pumpkin flesh (Li *et al.*, 2009), and a 5 x 5 x 5 mm piece of artificial pollen diet (Branson *et al.*, 1988; Singh, Moore, 1999). A 10 x 5 x 5 mm cube of 15% water agar served as a water source for the adults. Food and agar were changed every 5 to 7 days. All pairs of *D. v. virgifera* were provided with the same amount of food, as *D. v. virgifera* fitness and activity is known to be influenced by recent diet experience (Levine *et al.*, 2002; Mabry *et al.*, 2004).

In order to assess adult lifespan, bioassay containers were daily checked for live and dead adults. The date of death was recorded, lifespan calculated, and dead adults removed. The experiment was stopped after 70 days because this period is long enough to reliably reflect the total adult lifespan (Li *et al.*, 2009). The proportion of females and males surviving until day 70 was calculated for each tested population.

In order to assess fecundity, two teaspoons of moist, sterile black field soil (sieved at 0.15 mm mesh size; 25-35% moisture) were placed into the lower cup of the bioassay containers (see above) after 7 days maturation period (Branson, Johnson, 1973; Hill, 1975). Then, every 14 days after adult emergence, the lower cup of the bioassay container (containing soil and eggs) was removed and replaced with a new one. The soil with eggs was washed with hand warm tap water through a 0.25 mm sieve, and recovered eggs were counted. The experiment was stopped after 70 days because this period is long enough to reliably reflect the lifetime fecundity (Li *et al.*, 2009). The realised fecundity was calculated for each individual female and eggs. Eggs were then stored in sterile moist soil and kept at

24°C during two weeks for pre-diapause (Branson, 1976; Krysan, 1976; Krysan, 1972), and then at 6 to 8°C during another 5 months for diapause (Krysan, 1982)

In order to determine the overwintering survival of eggs, the soil with eggs was washed with hand warm tap water through a 0.25 mm sieve after 5 months diapause. The eggs survival was checked according to (Modic *et al.*, 2008) under a stereomicroscope and recorded per individual parental female of each tested populations.

In order to determine survival of larvae and pupae until adulthood, 200 successfully overwintered viable eggs incubated at a temperature of L:D, 25:21 °C for 14 to 24 days to initiate egg hatching. As not each female had laid enough eggs to obtain 200 eggs after overwintering, this experimental part was not conducted on an individual female base, but with pooled egg batches. About 2,000 to 3,000 seeds of the maize hybrid Gavott (UFA Semences, Bussigny, Switzerland) were soaked in water for 24h and germinated in a plastic tray (l: 330 mm, w: 190 mm, h:110 mm) with a gauze lid. Three days after germination, ready-to-hatch eggs were transferred onto the seeds in the plastic trays (approx. 5,000 eggs per tray) and maintained at a temperature of L:D, 25:21 °C, and light regime of L:D, 14:10. Emerging larvae found unlimited food (maize roots). After 14 to 20 days, third instar larvae were transferred along with the maize seedlings into gauze-covered cylinders (dia.: 120 mm, h: 140 mm) containing sterilised field soil for pupation (soil sieved at < 5 mm mesh size, 30 % moisture). The transferred maize still provided food for the larvae until entering the soil for pupation. Adults started to emerge about one week later. Adult emergence was recorded daily until no emergence was anymore recorded during four consecutive days. The total number of emerged adults was divided by the 200 initially incubated viable eggs to calculate the survival from hatching larvae via pupal stage until adulthood.

Survival from eggs via all developmental stages until adulthood was calculated by combining the data from the overwintering survival of eggs with the survival from larvae and pupae until adulthood.

The overall fitness of each individual *D. v. virgifera* female was calculated according to (Lombaert *et al.*, 2008) as follows:

Fitness = Larvae to adult survival * Adult sex ratio * Realized fecundity * Egg overwintering survival

A sex ratio of 50 to 50 was assumed according to (Spencer *et al.*, 2009; Toepfer, Kuhlmann, 2006). The fitness index, as calculated here, equals the net reproductive rate of a population (Begon *et al.*, 1990).

Activity traits

Flight activity of adults was assessed as a measure of dispersal capability, using methodology of Li *et al.*, (2010). Flight activity was assessed of young adults (7 day old) (Naranjo, 1991) because activity measures on young adults better reflect activity differences between individuals and between *D. v. virgifera* populations than measures on mature adults which are too much influenced by nutritional statuses and egg loads (Li *et al.*, 2010). In detail, flight activity was assessed by measuring the proportion of adults flying and the flight take-off response (as described by Duan *et al.*, 1998; Li *et al.*, 2010; Toepfer *et al.*, 2005). The flight stands used consisted of a wooden pin (h: 40 mm, dia.: 10 mm) fixed onto the end of an inverted white plastic funnel (h: 160 mm, dia.: 135 mm at base). The base of the funnel was surrounded with water to prevent the adults from walking off of the stand. An individual adult was gently transferred onto the base of the funnel using an aspirator device. Following adult release, the incidence of take-off and the time from release to take off were recorded. The trial ended when the adults flew off the stand or when 300 seconds had elapsed (Li *et al.*, 2010; Toepfer *et al.*, 2005). Approximately 148 CSE adults, 38 NWIT adults, 292 adults from the

CSE x NWIT cross and 304 adults from the NWIT x CSE cross were observed under laboratory conditions of 24 °C and 60% relative humidity between 14:00 and 16:00.

The proportion of adults flying and mean time until take-off was calculated for young and mature, female and male adults (non fliers were recorded as taking more than 300 seconds). Tested adults were returned to the bioassay containers.

Morphometric traits

Fresh body weight, elytra length and elytra width were measured on each individual *D. v. virgifera* within 24 h following adult emergence (= initial morphometric traits without feeding), because measures on young adults more easily allow the detection of differences between individuals and *D. v. virgifera* populations than measures on mature adults which are too variable due to influences by nutritional status and egg loads (Li *et al.*, 2010). The tested young adults were assumed to no longer be teneral, as they were fully coloured and did not have the light grey soft body typical of adults just after emergence.

Fresh body weight was measured by transferring adults into a small plastic container (h: 50 mm, dia.: 27 mm) and weighting them on a 0.1 mg to 160 g precision scale (Fox, Czesak, 2006). Individuals were then returned to the bioassay containers.

Elytra length and width (i.e., single measurement of both elytrons together across the dorsum) were measured according to Mabry *et al.* (2004) and Li *et al.* (2009) Elytra length and width were chosen among other morphometric traits, such as hind tibia length or head capsule width because measures of elytra characters well reflect fecundity, life span and activity of *D. v. virgifera* (Li *et al.*, 2009; Li *et al.*, 2010). Adults were placed on a cool pad (cool but not frozen IcePack Migros, Delemont, Switzerland) to limit their activity during the measurements with a micrometer scale to the nearest 0.06 mm under a stereomicroscope (16 × magnification).

VIII. – DISCUSSION GÉNÉRALE

L'objectif global de cette thèse était d'identifier, d'étudier et de documenter un cas d'hybridation au cours de l'invasion européenne de *Dv* et d'en évaluer les conséquences évolutives. Pour atteindre cet objectif, quatre études ont été réalisées et ont permis (i) de détecter une zone hybride en Italie du Nord entre les deux foyers envahissants d'Italie NO et d'Europe CSE ; (ii) d'estimer la dispersion de *Dv* entre ces deux foyers et (iii) de montrer que l'hybridation n'a pas dans ce cas de conséquences sur des caractères positivement liés à la valeur sélective des individus.

Cette dernière partie me permettra de revenir sur chacun des résultats obtenus au cours de ces travaux de thèse et de les discuter.

VIII.1 – L'Italie du Nord, une zone de contact riche en informations

La Vénétie est une zone hybride...

La zone de contact actuelle entre les deux principaux foyers envahissants de *Dv* en Europe (Italie NO et Europe CSE) est située dans la région de la Vénétie. Cette zone, intensément contrôlée depuis la première capture de *Dv* dans la région en 1998 au moyen de pièges à phéromones, a même été le lieu de la mise en place d'un programme d'éradication. Depuis 2008, la distribution géographique de *Dv* est devenue continue en Italie du Nord (Edwards, Kiss, 2009; Edwards, Kiss, 2010) suite au contact entre les populations de ces foyers (Edwards, Kiss, 2008; Edwards, Kiss, 2009; Edwards, Kiss, 2010).

Les différentes analyses que nous avons réalisées dans l'**article n°1** (l'analyse bayésienne de la structure des populations des échantillons de *Dvv* récoltés en Italie du Nord et en Europe CSE, l'analyse ABC du choix de scénarios historiques expliquant l'origine des différents échantillons de Vénétie et la détermination du taux d'hybridation des individus vénitiens) ont toutes montré que la Vénétie est bien une zone où des événements d'hybridation se produisent entre les différents foyers de *Dvv* et qui contient des individus hybrides. En effet, le génome des individus échantillonnés en Vénétie est assigné aux deux différents clusters⁸ génétiques identifiés (un cluster Ouest correspondant au foyer d'Italie NO et un cluster Est correspondant au foyer d'Europe CSE) avec des taux d'hybridation variables. Selon l'analyse ABC, la plupart des échantillons provenant de la partie occidentale de la Vénétie ont une simple origine et sont issus d'Italie NO alors que tous les autres échantillons résultent d'événements d'hybridation entre les foyers parentaux.

L'assignation génétique d'un individu de Summaga (l'échantillon le plus à l'Est de la Vénétie) au foyer d'Italie NO implique une dispersion à longue distance de *Dvv*. Des études récentes ont suggéré qu'une telle dispersion est commune chez *Dvv* et que plusieurs de ses populations envahissantes s'étendent par le biais d'une dispersion stratifiée —définie par la présence conjointe de dispersion locale et dispersion à longue distance (Shigesada *et al.*, 1995) — en Europe et aux Etats-Unis (Ciosi *et al.*, 2011; Gray *et al.*, 2009).

Enfin, la zone hybride a une largeur très importante (environ 100 km) malgré le fait qu'elle soit récente (courte période de temps entre le contact secondaire (2008) et l'échantillonnage (2009)). Cela peut s'expliquer par les capacités de dispersion de *Dvv*, supposées considérables : le taux d'expansion du foyer d'Europe Centrale a été estimé à 60-100 km/an (Baufeld, Enzian, 2001).

⁸ Ce terme emprunté au vocabulaire anglais se réfère à un ensemble génétiquement homogène.

...qui peut s'avérer utile pour étudier la dispersion de *Dvv* et les conséquences évolutives de l'hybridation !

Ces premiers résultats ont appelé une analyse des clines plus approfondie (voir section suivante) sur la zone géographique étudiée nous permettant d'estimer des paramètres de dispersion (*e.g.* Barton, Gale, 1993). Sous le seul effet de la dispersion des individus, les fréquences alléliques ont tendance à devenir homogènes dans l'espace et entraînent une diminution de la pente des clines observés, directement dépendante de l'intensité de la dispersion et du temps depuis le contact secondaire (*e.g.* Endler, 1977).

Par ailleurs, les analyses que nous avons menées au cours de cette première étude ont montré que la variabilité génétique neutre de la zone hybride était plus importante que celle des populations parentales et approchait celle de la population source d'Amérique du Nord; traduisant ainsi une restauration importante de la variabilité génétique perdue au cours des étapes d'introduction et d'établissement des différents foyers envahissants. Les résultats de cette première étude ne fournissent cependant aucune information sur les différences de variabilité phénotypique et de valeur sélective entre la zone hybride et les régions parentales. Des études récentes d'invasion ont rapporté par exemple un effet positif de l'hybridation sur la valeur sélective et les capacités envahissantes de plusieurs espèces animales, comme la coccinelle asiatique *Harmonia axyridis* (Turgeon *et al.*, 2011) ou l'escargot d'eau douce *Melanoides tuberculata* (Facon *et al.*, 2005) et végétales (Ellstrand, Schierenbeck, 2000) tel que le Silène commun, *Silene vulgaris* (Keller, Taylor, 2010). Des études de génétique quantitative sur des caractères d'histoire de vie (comme la fécondité, la longévité, la survie et la dispersion) ont donc été menées pour nous permettre de déterminer si l'hybridation observée de *Dvv* en Italie du Nord confère un avantage aux individus hybrides et peut affecter la dynamique de son invasion en Europe.

Nous avons donc cherché à creuser et exploiter ces différentes pistes de recherche offertes par la détection de la zone hybride dans la suite de cette thèse.

VIII.2 – Contact secondaire, hybridation et estimation de la dispersion

La seconde étude que nous avons réalisée au cours de cette thèse (**article n°2**) avait pour but d'exploiter la zone hybride — précédemment détectée en Vénétie suite à un contact secondaire — entre les deux foyers européens envahissants de *DvV* afin d'en obtenir des estimations de distances de dispersion chez cette espèce. Les analyses conduites sur la variation spatiale des fréquences alléliques microsattellites et du déséquilibre de liaison ont généré deux estimations de l'écart-type de la distribution des distances parents-enfants, σ , comprises entre 18 et 22 km/generation^{1/2} avec des « support limits⁹ » associées approximativement comprises entre 16 et 26 km/generation^{1/2}.

Validité des estimations

Les méthodes d'estimation de la dispersion utilisées ici sont basées sur plusieurs hypothèses : la zone hybride est supposée neutre ou très faiblement soumise à la sélection, la migration des individus suit un processus de diffusion — *i.e.* une migration aléatoire, locale et isotropique qui résulte en un déplacement proche du mouvement brownien et pouvant être approximée par un noyau de dispersion gaussien — , les marqueurs génétiques utilisés sont supposés indépendants et non soumis à des zones du génome sous sélection, la taille de

⁹ Ensemble des valeurs de paramètres correspondant au maximum de vraisemblance moins 2 unités du log de la vraisemblance.

population est supposée rester constante dans l'espace et enfin la dérive est considérée comme faible. Quelle confiance peut-on *a priori* accorder à chacune de ces hypothèses ?

Tout d'abord on s'attend à ce que les marqueurs microsatellites utilisés (dans les différentes études de la thèse d'ailleurs) soient neutres [(Ellegren, 2004) mais voir (Sutherland, Richards, 1995; Watkins *et al.*, 1995)] mais il est possible que certains d'entre eux soient liés à des régions du génome soumises à la sélection (Charlesworth *et al.*, 1993; Zhong *et al.*, 1994). Cependant dans notre étude, plusieurs indices suggèrent fortement que la sélection n'est pas impliquée dans l'établissement des clines décrits. La sélection sur des régions spécifiques du génome peut affecter différemment l'ensemble des loci considérés et créer de l'hétérogénéité dans les centres et les pentes des clines. Ici, tous les clines sont coïncidents (même centre), concordants (même pente) et les fonctions sigmoïdes qui les décrivent sont moins raides que ce qui est attendu en cas de forte sélection sur de multiples loci (Barton, Gale, 1993). De plus, la contre-sélection des hybrides semble ici improbable parce que les deux foyers d'Italie NO et d'Europe CSE ont été fondés trop récemment (en 2010 et 1992, respectivement) pour que leur temps d'isolation (environ 20 générations) ne conduise à de l'isolement reproducteur (Edmands, 2002; Orr, Turelli, 2001).

Nos estimations de dispersion reposent aussi sur l'hypothèse d'un noyau de dispersion gaussien. Cette dernière suppose que les événements de migration longue distance sont très rares, ce qui n'est pas une généralité puisque la dispersion à longue distance est un processus fréquemment décrit chez plusieurs espèces (Lockwood *et al.*, 2007). Récemment plusieurs études ont montré que les insectes étaient capables de se propager sur de très grandes distances. Nous pouvons citer en exemple le charançon du coton, *Anthonomus grandis* (Kim, Sappington, 2004) ou même notre modèle biologique, *Dv* (Ciosi *et al.*, 2011; Miller *et al.*, 2005). Ainsi, les noyaux de dispersion ne semblent pas être souvent gaussiens mais plutôt

avoir des queues de distribution lourdes et des valeurs de kurtosis¹⁰ positives (Liebhold, Tobin, 2008; Lockwood *et al.*, 2007). Nous avons donc simulé des données génotypiques résultant de différents noyaux de dispersion avec des queues de distribution légères ou lourdes et avec des valeurs de kurtosis positives. Si ces valeurs de kurtosis ne sont pas trop importantes, les estimations à partir de données simulées ont montré que les valeurs estimées de σ sont proches des vraies valeurs. Ceci indique que les estimations de la dispersion ne sont pas très sensibles au type de noyau de dispersion et à la kurtosis. Ces résultats en corroborent d'autres, obtenus précédemment avec des méthodes comparables et montrant que la kurtosis a un faible effet sur l'estimation de σ (Rieux *et al.*, 2013).

Au sein des très grandes aires d'étude telle que celle étudiée ici (plus de 800 km de large), l'hétérogénéité environnementale induit probablement une hétérogénéité de densités de populations dans l'espace. De plus, les deux foyers envahissants de *Dvv* ont des âges différents : Italie NO est apparu 8 ans après Europe CSE (Miller *et al.*, 2005) et ceci peut aussi expliquer des différences de densité. Enfin, une différence de pression de sélection d'origine anthropique (traitement insecticides et rotation des cultures) pourrait également générer des différences de densités entre deux populations en expansion. Quand les densités sont égales, la pente des clines est une fonction de la dispersion. Des différences de densités de part et d'autre de la zone de contact devraient conduire à une pente de cline plus raide, en raison d'une plus forte contribution au flux de gène de la population la plus densément peuplée (Barton, 1979). Nous avons simulé un scénario dans lequel les différences de taille étaient constantes entre les deux populations simulées (une population Ouest et une population Est). Les simulations ont montré que ces différences avaient peu d'influence sur les estimations de la dispersion, puisque les biais observés étaient toujours inférieurs à 20% même pour un ratio de densité de 500%.

¹⁰ Défini comme le coefficient d'aplatissement d'une distribution.

Nous avons supposé un taux de recombinaison de 0.5 entre tous les loci microsattellites. Ceci semble raisonnable, puisque toutes les valeurs du déséquilibre de liaisons étaient égales à zéro en-dehors de la zone de contact. Cependant, cette hypothèse peut être fautive pour certaines paires de microsattellites pouvant afficher un taux de recombinaison légèrement en dessous de 0.5. Comme σ^2 est proportionnel à r — selon la méthode du déséquilibre de liaison (Barton, Gale, 1993) — cela peut conduire à une surestimation de σ lorsque l'on considère $r = 0.5$. Notons que pour une large gamme de taux de recombinaison, compris entre 0.3 et 0.5, une hypothèse erronée d'indépendance complète entre loci devrait résulter en un biais limité de l'estimation de σ de moins de 25%.

Les méthodes utilisées dans cette seconde étude pour estimer la dispersion (méthode du déséquilibre de liaison et la méthode des clines neutres, (Barton, Gale, 1993; Endler, 1977) sont toutes deux basées sur l'hypothèse que la dérive a un effet suffisamment limité pour ne pas créer d'hétérogénéité entre loci. *A priori* nous supposons que c'est le cas pour *Dvv* en Italie du Nord puisque de très grandes tailles de populations ont souvent été rapportées vis-à-vis des deux foyers envahissants (Borioni *et al.*, 2006; Szalai *et al.*, 2011). Par ailleurs, la dérive semble faible dans ces foyers puisque des études précédentes ont montré une différenciation génétique négligeable entre échantillons au sein de chacun des foyers (Bermond *et al.*, 2012; Ciosi *et al.*, 2008). Nous pensons aussi que la dérive a peu d'impact sur la structure génétique de *Dvv* dans la zone de contact car (i) nous avons trouvé des clines concordants et parallèles alors que la dérive devrait créer de l'hétérogénéité dans la forme des clines observés entre loci (centre, largeur..) et parce que (ii) 86% de la déviance a été expliquée par le modèle sélectionné pour ajuster les fréquences alléliques observées (le modèle M_1) alors qu'une dérive importante est supposée générer un bruit de fond important autour des fréquences alléliques attendues diminuant ainsi la qualité de l'ajustement (le pourcentage de déviance expliquée).

Comparaison avec d'autres estimations de la dispersion chez *DvV*

DvV est l'un des pires ravageurs du maïs à l'échelle mondiale. Les stratégies de lutte utilisées contre cette peste, y compris l'utilisation de maïs transgénique produisant des toxines de *Bacillus thuringiensis* et de zones refuges (voir **section II.3**) sont basées sur l'hypothèse d'une dispersion efficace de l'insecte. La combinaison des différentes pressions de sélection (*i.e.* divers pesticides) dans l'espace — afin d'accroître le degré d'hétérogénéité du traitement (REX, 2013) contre *DvV* — nécessite une connaissance approximative des distances sur lesquelles ce ravageur peut disperser. Actuellement, la chrysomèle des racines du maïs envahit l'Europe et les différents foyers envahissants se développent géographiquement à une vitesse qui dépend de ses capacités de dispersion. Les analyses de risques et les simulations des différents scénarios de défense de l'homme contre la propagation de *DvV* (Baufeld, Enzian, 2001; Carrasco *et al.*, 2012) exigent leur paramétrage. C'est pourquoi de nombreuses études antérieures ont mis l'accent sur la dispersion de *DvV* aux Etats-Unis (Coats *et al.*, 1986; Gray *et al.*, 2009) et en Europe (Baufeld, Enzian, 2001; Ciosi *et al.*, 2011). Carrasco *et al.* (2010) ont réalisé une expérimentation de capture-marquage-recapture et ont montré que la moitié des individus testés sur le terrain avait une distance de dispersion comprise entre 117 et 425 m/jour (avec une médiane de 117-188 m/jour). Étant donné que la durée de vie des adultes de *DvV* est de trois mois (Branson, Johnson, 1973; Hill, 1975), ces distances quotidiennes correspondent à des capacités de dispersion comprises entre 11 et 38 km/génération (ou une médiane de 11-17 km/génération) conformément à nos estimations. Par ailleurs, Baufeld et Enzian (2001) ont ajusté un modèle à l'expansion de la population d'Europe CSE depuis le début de cette invasion près de Belgrade en 1992 et ont estimé le taux d'expansion à 60-100 km/an. En supposant une approximation de diffusion, le modèle de la vague d'avance de Fisher suggère que la vitesse d'expansion d'une population colonisatrice avec un taux de croissance intrinsèque r est $C = 2\sigma r^{1/2}$ (Shigesada *et al.*, 1995). En 2011,

Szalai *et al.* (2011) ont montré que le taux de croissance annuel moyen du foyer Europe CSE était environ égal à quatre, ce qui donne une valeur de σ comprise entre 15 et 25 km/an^{1/2}. Nos estimations sont cohérentes avec ces résultats.

Les résultats obtenus au cours de cette seconde étude présentent un double intérêt. Premièrement, nous avons pu estimer assez précisément les capacités de dispersion de notre modèle biologique, ce qui était notre objectif principal. Deuxièmement, les analyses de clines réalisées et plus particulièrement l'observation de la forme des clines entre les différents loci suggèrent que les marqueurs utilisés ne sont pas liés à des zones du génome soumises à la sélection et que les hybrides ne sont pas contre-sélectionnés.

Cependant, nous avons voulu aller plus en avant dans le test de cette hypothèse. Il s'agissait de déterminer si la zone hybride identifiée est le théâtre d'une contre-sélection des hybrides, d'une zone hybride neutre ou même d'un avantage aux hybrides. C'est pourquoi nous avons réalisé deux études supplémentaires ayant pour objectifs de tester l'effet de l'hybridation au cours de l'invasion européenne de *Dv* en mesurant et étudiant des caractères phénotypiques positivement corrélés à la valeur sélective d'individus de type hybride et parental.

VIII.3 – Aucun effet majeur de l'hybridation au cours de l'invasion européenne de *Dv*

Dans la première de ces deux études (**article n°3**), nous avons testé si le statut hybride des individus influençait leur succès d'accouplement et leur survie à jeun. Nos résultats montrent que ce n'est probablement pas le cas. En outre, le choix des partenaires sexuels ne semble pas dépendre de la classe génotypique des adultes (hybride ou parentale).

Les individus hybrides sont fréquents au centre de la zone de contact

Sous l'hypothèse d'une contre-sélection des individus hybrides, les fréquences des génotypes parentaux sont classiquement plus élevées que celles des hybrides au cœur de la zone hybride (Jiggins, Mallet, 2000). En conséquence, on s'attend à ce que la distribution du coefficient Q soit bimodale avec de nombreuses valeurs de Q faibles (0-0.2) et hautes (0.8-1) et relativement peu de valeurs intermédiaires (0.2-0.8). Ce n'est pas ce que nous avons observé dans la première étude : (i) les résultats de l'analyse Structure montrent que les 3 échantillons contiennent une majorité d'individus hybrides et (ii) TV11 et Vi9 contiennent très peu d'individus appartenant à la classe 1 (0-0.2) probablement parce que ces échantillons ont été prélevés dans la partie ouest de la zone de contact plutôt qu'en son centre.

Le statut hybride n'influence probablement pas la valeur sélective individuelle

Dans cette étude, nous avons cherché à déterminer si les classes génotypiques (les classes de Structure) et le taux d'hybridation des individus ont influencé leur probabilité d'accouplement et leur capacité de survie dans des conditions stressantes. Pour ce faire, nous avons utilisé des tests statistiques (des tests exacts de Fisher, des GLMMs et des modèles de Cox mixtes) (i) pour comparer les génotypes des femelles accouplées avec ceux des femelles non accouplées ; (ii) pour comparer les génotypes de tous les mâles avec ceux des mâles ayant réussi à s'accoupler, par le biais du génotype du sperme contenu dans les femelles accouplées ; (iii) pour tester si l'accouplement est aléatoire au sein des différentes classes génotypiques, et (iv) pour tester si les classes génotypiques et le taux d'hybridation ont eu un impact sur le temps de survie à jeun des adultes.

Concernant les deux premiers points, aucun effet significatif n'a été détecté et nous pouvons ainsi conclure que de manière globale, le génotype des mâles n'a pas d'influence sur leur probabilité d'accouplement. La même conclusion tient aussi pour les femelles. Il n'y a donc aucune indication montrant que le statut hybride influence la probabilité d'accouplement.

Concernant l'hypothèse d'un accouplement au hasard, aucune association n'a été détectée entre les classes génotypiques des mâles et celles des femelles au sein des couples accouplés. Les classes génotypiques des femelles accouplées et celles du sperme au sein des femelles ne sont pas statistiquement liées. Les problèmes de puissance statistique ne sont probablement pas ici la cause de l'absence d'un effet puisque les fréquences des classes génotypiques étaient le plus souvent équilibrées entre les individus accouplés (le sperme et les femelles fécondées).

De la même manière, l'analyse des modèles de Cox montrent que les classes génotypiques (classes de Structure) et le taux d'hybridation n'influencent pas le temps de survie dans des conditions stressantes. Cependant, le sexe a une influence forte et significative : les mâles ont un risque de mortalité 1.3 fois plus élevé que celui des femelles. Ce résultat était inattendu car la littérature suggère que les mâles de *Dvv* vivent plus longtemps que les femelles dans la nature (102 jours pour les mâles contre 78 pour les femelles, (Hill, 1975). Deux hypothèses peuvent expliquer ce résultat : (i) les échantillons ont été prélevés après l'accouplement de la majorité des femelles, à un moment où elles se nourrissent alors que les mâles sont à la recherche de nouvelles partenaires sexuelles et ont probablement moins de ressources que les femelles (Hill, 1975; Spencer *et al.*, 2009). Étant donné que le test de survie a été réalisé dans des conditions de jeun, les mâles ont probablement subi un désavantage par rapport aux femelles ; (ii) au moment de l'accouplement les mâles transfèrent le sperme aux femelles par le spermatophore. Chez

certaines insectes, le spermatophore est associé au spermatophylax, une vésicule nutritive donnée par les mâles aux femelles comme cadeau nuptial qui a pour effet de prolonger le temps d'accouplement et d'assurer le transfert des spermatozoïdes (Arnaud, 1999; Vahed, 1998). Cette vésicule nutritive est présente chez *Dv* (Murphy, Krupke, 2011). Si les femelles utilisent aussi le spermatophylax comme ressource nutritive, elles peuvent bénéficier d'un bonus énergétique par rapport aux mâles, ce qui pourrait expliquer leur plus grande survie.

Critique de l'analyse

La présente étude est originale parce que la plupart des résultats sur la valeur sélective ont été obtenus sans manipulation ou expérimentation, directement sur des individus prélevés *in natura*. Les résultats que nous avons obtenus sont remarquables par leur simplicité : aucun effet de l'hybridation sur des caractères positivement liés à la valeur sélective n'a été détecté, quelle que soit la mesure du statut hybride (classes génotypiques ou taux d'hybridation) et quel que soit le caractère. Cependant, certaines faiblesses persistent. Tout d'abord, nous avons mesuré seulement deux caractères liés à la valeur sélective — le taux d'accouplement et le temps de survie à jeun des adultes — qui peuvent difficilement être généralisés à toutes ses composantes. Des analyses complémentaires sur d'autres caractères phénotypiques tels que la fertilité, la vitesse de développement ou le taux de survie des œufs sont donc nécessaires. D'autre part, l'analyse du temps de survie est semi-contrôlée. Nous ne disposons d'aucune information sur l'âge des individus ou sur la date de leur dernière prise alimentaire. Toutefois, l'hypothèse que l'âge et la satiété des individus sont répartis aléatoirement au sein des différentes classes hybrides semble raisonnable et la détection d'un effet du statut hybride ne devrait donc pas être liée à ces deux variables. Par ailleurs, très peu de femelles non-accouplées ont été observées. Les tests statistiques associés au succès d'accouplement des femelles peuvent donc ne pas être suffisamment puissants pour détecter un quelconque effet.

En ce qui concerne le succès d'accouplement des mâles, nous sommes obligés de comparer les génotypes des mâles aux génotypes des spermatozoïdes contenus dans les femelles car nous n'avons pas accès directement aux mâles accouplés et non accouplés. Nous ne savons donc pas s'il existe une proportion significative de mâles ne s'accouplant pas, bien que nous fassions l'hypothèse de l'existence d'une telle fraction. Dans le cas limite où tous les mâles s'accouplent de façon égale, il est normal de ne pas détecter de différence entre génotypes des mâles et génotypes des spermatozoïdes.

L'absence d'effet de l'hybridation peut être expliquée par la situation temporelle de l'étude au cours du processus envahissant. En effet, la plupart des études qui ont révélé un impact positif de l'hybridation sur l'invasion ont été menées dans des situations où l'hybridation s'était déroulée lors des étapes précoces de l'invasion, peu après l'introduction (Drake, 2006; Ellstrand, Schierenbeck, 2000; Keller, Taylor, 2010). Dans la présente étude, notre échantillonnage et les tests conduits sont réalisés longtemps après les premières introductions de *Dvv* en Europe. Les impacts de l'hybridation pourraient différer considérablement entre des contacts qui se produisent peu de temps après les différentes introductions de populations envahissantes et des contacts secondaires se produisant suite à leur expansion géographique. En supposant que le succès de l'invasion dépend en partie de la capacité de la population à répondre à de nouvelles pressions sélectives, dans le cas d'un contact secondaire comme observé en Vénétie, les deux populations ont déjà pu surmonter les nouvelles pressions de sélection et évoluer vers des niveaux supérieurs de valeur sélective. Par conséquent, il peut être plus difficile de détecter des effets positifs de l'hybridation dans ce contexte.

Confirmation des résultats obtenus sur d'autres caractères mesurés

Dans une seconde étude (**Chapitre IV**), nous avons cherché à déterminer si l'hybridation entre les foyers d'Italie NO et d'Europe CSE se traduisait par une augmentation de la variance phénotypique ou par un effet hétérosis.

Différents croisements génétiques ont donc été réalisés au laboratoire à partir de ces deux foyers afin d'obtenir des individus hybrides et parentaux sur lesquels 10 caractères phénotypiques ont été mesurés. Les caractères ont été choisis en raison de leur lien direct à la valeur sélective (fécondité réalisée, survie des œufs pendant l'hiver, survie des œufs jusqu'à l'âge adulte, valeur sélective globale, durée de vie adulte) ou indirect (masse corporelle fraîche, longueur des élytres, largeur des élytres) ou parce qu'ils reflètent la capacité de *Dv* à disperser et envahir (proportion d'adultes volants, temps moyen d'attente jusqu'au décollage de l'insecte) (Li *et al.*, 2010).

Globalement, les résultats obtenus indiquent que les individus issus de croisements hybrides ont des performances équivalentes à celles des individus issus de l'un des deux croisements parentaux (généralement le plus mauvais des deux). Aucun caractère phénotypique mesuré n'a permis de mettre en évidence une augmentation de la valeur moyenne chez les hybrides par rapport aux parents. Seul un caractère « la survie des œufs pendant l'hiver » a révélé que les performances hybrides étaient moins bonnes que celles des parents. Par ailleurs, aucune différence significative de variance n'a été détectée sur les 10 caractères mesurés entre croisements hybrides et parentaux. Les efforts fournis pour reconstruire les familles au sein de chaque type de croisement, pour en tenir compte au sein des analyses statistiques et ainsi estimer un proxy de la variance génétique n'ont donc pas été à la hauteur de nos attentes. En effet, la prise en compte de la structure en famille n'a rien

apporté, si ce n'est qu'une variance inter familles a été détectée pour certains caractères phénotypiques mesurés.

Les résultats de cette dernière étude corroborent donc ceux de l'étude précédente (**article n°3**). Ils n'indiquent aucun effet important de l'hybridation entre les deux principaux foyers envahissants au cours de l'invasion européenne de *Dv*v et suggèrent sur cette base qu'aucun changement majeur de la dynamique d'invasion de cette peste n'est à prévoir en Europe.

VIII.4 – Synthèse

De récentes études ont montré que les introductions multiples de *Dv*v en Europe en provenance d'Amérique du Nord ont conduit à la formation de plusieurs foyers dont deux principaux : le foyer d'Italie du Nord-Ouest (fondé en 2000) et le foyer d'Europe Centrale et du Sud-Est (fondé en 1992) (Ciosi *et al.*, 2008; Miller *et al.*, 2005). Les goulots d'étranglement associés à leur introduction respective ont induit une importante baisse de leur variabilité génétique par rapport à la population source (30 à 40%) et ont conduit à leur forte différenciation génétique (Ciosi *et al.*, 2008). L'observation de leur aire de répartition respective au cours du temps a montré que ces deux foyers ont connu une expansion géographique, qu'ils ont convergé l'un vers l'autre et sont probablement entrés en contact aux alentours de 2008 en Italie du Nord dans la région de la Vénétie. L'observation d'un tel évènement a naturellement suscité de nombreuses questions chez les chercheurs étudiant la biologie de l'invasion de *Dv*v en Europe (Ciosi, 2009) : la zone de contact détectée en Vénétie est-elle une zone hybride ? Si tel est le cas, est-ce que l'hybridation permettra de restaurer une partie de la variabilité génétique de la population source américaine ? Les populations hybrides auront-elles une variabilité génétique neutre ou autre plus élevée que celle des populations parentales, comme attendu ? Quelles en seront les conséquences au niveau

phénotypique ? Est-ce que l'hybridation aura des conséquences sur la dynamique d'invasion de *Dvv* en Europe ?

Ma thèse s'est donc inscrite dans ce contexte et avait pour but de répondre à bon nombre de ces questions. Je résume ici les principaux résultats obtenus à l'issue de ces trois années de recherche et les différentes réponses apportées aux questions posées.

Dans un premier temps (voir **article n°1**), nous avons prouvé grâce à une étude de génétique des populations que la zone de contact située en Vénétie résultait bien en une zone hybride, c'est-à-dire une zone dans laquelle des individus issus de populations différenciées génétiquement se croisaient, s'accouplaient et produisaient des individus hybrides.

=> La Vénétie est bien une zone hybride.

Dans un second temps (voir **article n°2**), nous avons profité de l'opportunité que nous offrait cette zone hybride pour estimer un paramètre clef de l'invasion : la dispersion. Pour cela nous avons réalisé une analyse de clines approfondie et une analyse du déséquilibre de liaison. Les deux méthodes ont fourni des estimations de la dispersion de *Dvv* entre les deux foyers européens congruentes et égales à environ 20 km/génération^{1/2}.

=> *Dvv* a une capacité de dispersion entre les deux foyers européens estimée à environ 20 km/génération^{1/2}.

Le type de zone hybride en présence peut fournir des informations sur les conséquences évolutives de l'hybridation (voir **sections I.2.1.b et I.2.2**) ce qui représente un des objectifs de cette thèse. Notre estimation de la dispersion *via* la théorie des clines repose sur l'hypothèse d'une zone hybride neutre (caractérisée par l'absence de sélection et aucune différence de valeur sélective entre individus de type parentaux et hybrides) mais est-ce

réellement le cas ? *A priori*, nous pensions que le temps d'isolation entre les deux foyers précédemment décrits n'était pas suffisant pour induire un isolement reproducteur et se traduire par une contre-sélection des hybrides (zone de tension). *A posteriori*, l'absence de pentes de clines raides et d'hétérogénéité des ajustements entre loci semble indiquer que les marqueurs utilisés ne sont pas liés avec des zones du génome soumises à la sélection. L'hypothèse d'une zone hybride neutre apparaissait donc à ce stade comme probable mais nécessitait encore d'être testée. Y a-t-il une différence significative de valeur sélective entre des individus hybrides et parentaux? C'est ce que nous avons regardé au travers des deux dernières études....

=> La zone hybride détectée en Vénétie correspond probablement à une zone hybride neutre mais cette hypothèse reste encore à être testée.

Les deux dernières études (**article n°3 et Chapitre IV**) réalisées *in natura* et au laboratoire ont montré que les performances des individus parentaux et hybrides (mesurées sur différents caractères phénotypiques liés directement ou indirectement à la valeur sélective mais aussi aux capacités de dispersion de *Dvν*) ne sont pas significativement différentes. Les conséquences évolutives de l'hybridation dans le cas de la biologie de l'invasion de *Dvν* en Europe apparaissent donc comme neutres sur la base des études que nous avons menées au cours de cette thèse et ne suggère donc pas une nouvelle dynamique d'invasion.

=>L'hybridation entre les deux foyers envahissants de *Dvν* en Europe paraît neutre d'un point de vue évolutif : les hybrides ont une valeur sélective équivalente à celle des parents.

Même si aucun effet important de l'hybridation n'a été détecté au cours de cette thèse, nous pensons que le contexte biologique de l'invasion de *Dvν* en Europe était adéquat pour

étudier les conséquences évolutives de l'hybridation. En effet, nous l'avons déjà évoqué plusieurs fois dans ce manuscrit, les effets forts des goulots d'étranglement lors de la phase d'introduction des différentes populations introduites ont provoqué une forte diminution de la variabilité génétique en leur sein par rapport à la population source et une différenciation génétique importante entre foyers indépendamment fondés (Ciosi *et al.*, 2008). L'hybridation ultérieure de ces foyers génétiquement différenciés peut alors se traduire chez les individus par une augmentation ou une diminution de leur valeur sélective (voir **sections I.3.2 et I.3.3**). Dans notre cas, la détection d'un effet hétérosis a certainement été rendue plus difficile en raison des capacités d'invasion importantes de *Dv* au sein de chacun des foyers parentaux (Italie du NO et Europe CSE) traduisant ainsi des niveaux de valeur sélective conséquents et suffisants pour surmonter la réduction de potentiel adaptatif (entraînée par la réduction de la variabilité génétique) liée aux étapes précoces de l'invasion. Notons que l'observation d'une dépression hybride aurait pu dans ce cas être détectée. D'autres situations plus favorables, par exemple avec des introductions plus récentes (Drake, 2006; Ellstrand, Schierenbeck, 2000) auraient peut-être permis de constater plus facilement des effets forts tels que l'hétérosis ou la dépression hybride.

VIII.5 – La zone hybride temporaire comme opportunité pour estimer la dispersion

Les zones hybrides temporaires qui disparaissent au cours du temps et qui résultent de contacts secondaires neutres entre populations sont rarement utilisées pour estimer la dispersion des espèces. Une seule étude a récemment fait usage de ce contexte et de la relation trouvée par Endler (1977) entre σ et la largeur des clines neutres. Il s'agit d'une étude des

distances de dispersion d'un champignon phyto-pathogène du bananier *Mycosphaerella fijiensis* en Afrique (Rieux *et al.*, 2013). Ce constat apparaît comme surprenant car la méthodologie utilisée dans un tel contexte présente de nombreux avantages : (i) elle fournit des estimations de distances de dispersion sujette à de faibles biais ; (ii) elle fournit des estimations qui sont robustes lorsqu'on s'écarte des hypothèses de départ ; (iii) elle fournit des « supports limits » étroites, indiquant un faible niveau d'incertitude sur les estimations ; (iv) elle est relativement peu coûteuse en termes de temps et de moyens car elle nécessite uniquement le génotypage de quelques marqueurs génétiques polymorphes pour quelques centaines d'individus collectés lors d'un échantillonnage unique et (v) il n'est pas nécessaire de connaître précisément le temps depuis le contact entre les populations car ce facteur peut être estimé. La principale contrainte associée à cette méthode est que l'emplacement de la zone de contact secondaire doit être à peu près connu avant le prélèvement, afin de garantir des estimations précises de la largeur du cline.

Cette méthode semble être particulièrement appropriée à l'étude de la dispersion dans le cadre des invasions biologiques. En effet, plusieurs invasions proviennent d'introductions multiples dans des zones déconnectées (Barney, 2006; Dlugosch, Parker, 2008; Hufbauer, Sforza, 2008; Lombaert *et al.*, 2010; Urbanelli *et al.*, 2000) et beaucoup d'entre elles conduisent à des contacts secondaires (*e.g.* Lavergne, Molofsky, 2007; Lombaert *et al.*, 2010; Lucek *et al.*, 2010; Rieux *et al.*, 2013) dans lesquels la dispersion pourrait être estimée *via* des analyses de clines. Mais comme nous l'avons déjà évoqué précédemment, en dépit du nombre important de situations adaptées, peu d'exemples de telles études sont disponibles dans la littérature. Ce constat peut s'expliquer par une simple difficulté : l'expansion spatiale rapide des populations envahissantes, leur observation et leur échantillonnage tardif résultent probablement en l'observation d'une seule population génétiquement homogène ne présentant aucun cline de fréquences alléliques (Ciosi, 2009).

Les clines de fréquences alléliques temporaires sont donc probablement très fréquents *in natura*, en particulier dans le contexte des invasions biologiques et pourraient être utilisés pour estimer les distances de dispersion de nombreuses espèces avec précision si le temps depuis le contact n'est pas trop grand. Par ailleurs, les observations et études répétées de contacts secondaires résultant en évènement d'hybridation pourraient également nous informer sur l'importance de ce processus dans le cadre des invasions biologiques. Dans quels cas peut-on s'attendre à observer de l'hétérosis (nécessité de populations introduites de très faible taille ayant fixé les mutations délétères ?), de la dépression hybride (nécessité d'adaptation locale et de sous dominance ?), une augmentation de la variance génétique (nécessité d'une multiplicité de populations sources) ?

RÉFÉRENCES BIBLIOGRAPHIQUES

- Allee WC, Park O, Emerson AE, Park T, Schmidt KP (1949) *Principles of animal ecology* WB Saundere Co. Ltd.
- Anderson E, Thompson E (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217-1229.
- Arnaud L (1999) La compétition spermatique chez les insectes: les stratégies d'assurance de la paternité et la préséance du sperme. *Biotechnol. Agron. Soc. Environ* **3**, 86-103.
- Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics* **23**, 237-261.
- Barney J (2006) North American History of Two Invasive Plant Species: Phylogeographic Distribution, Dispersal Vectors, and Multiple Introductions. *Biological Invasions* **8**, 703-717.
- Barton NH (1979) The dynamics of hybrid zones. *Heredity* **43**, 341-359.
- Barton NH (1982) The Structure of the Hybrid Zone in *Uroderma bilobatum* (Chiroptera, Phyllostomatidae). *Evolution* **36**, 863-866.
- Barton NH (2000) Estimating multilocus linkage disequilibria. *Heredity* **84**, 373-389.
- Barton NH, Gale KS (1993) Genetic analysis of hybrid zone. In: *Hybrid zones and the evolutionary process* (ed. Harison RG), pp. 13-45. Oxford University Press, New York.
- Barton NH, Hewitt GM (1981) The genetic basis of hybrid inviability in the grasshopper *Podisma pedestris*. *Heredity* **47**, 367-383.
- Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113-148.
- Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature* **341**, 497-503.
- Baufeld P, Enzian S (2001) Simulations model for spreading scenarios of western corn rootworm (*Diabrotica virgifera virgifera*) in case of Germany. *IWGO Newsletter* **22**, 14-15.
- Beaumont MA, Zhang WY, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025-2035.
- Begon M, Harper JL, Townsend CR (1990) *Ecology: individuals, populations and communities*.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* **57**, 289-300.
- Benvenuto C, Cheyppé-Buchmann S, Bermond G, Ris N, Fauvergue X (2012) Intraspecific hybridization, life history strategies and potential invasion success in a parasitoid wasp. *Evolutionary Ecology* **26**, 1311-1329.
- Bermond G, Blin A, Vercken E, *et al.* (2013) Estimation of the dispersal of a major pest of maize by cline analysis of a temporary contact zone between two invasive outbreaks. *Molecular Ecology*.

- Bermond G, Ciosi M, Lombaert E, *et al.* (2012) Secondary Contact and Admixture between Independently Invading Populations of the Western Corn Rootworm, *Diabrotica virgifera virgifera* in Europe. *PloS one* **7**, e50129.
- Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology* **19**, 2609-2625.
- Boriani M, Agost M, Kiss J, Edwards CR (2006) Sustainable management of the western corn rootworm, *Diabrotica virgifera virgifera* LeConte (Coleoptera: Chrysomelidae), in infested areas: experiences in Italy, Hungary and the USA. *EPPO Bulletin* **36**, 531-537.
- Born W, Rauschmayer F, Brauer I (2005) Economic evaluation of biological invasions - a survey. *Ecological Economics* **55**, 321-336.
- Bossdorf O, Auge H, Lafuma L, *et al.* (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**, 1-11.
- Branson TF (1976) Viability and hatching pattern of eggs of the western corn rootworm exposed to chill periods of different durations. *Entomologia Experimentalis et Applicata* **19**, 77-81.
- Branson TF, Guss PL, Krysan JL, Sutter GR (1975) Corn rootworms: laboratory rearing and manipulation. In: *Corn rootworms: laboratory rearing and manipulation.*, p. [2 +] 18 pp.
- Branson TF, Jackson JJ, Sutter GR (1988) Improved method for rearing *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae). *Journal of Economic Entomology* **81**, 410-414.
- Branson TF, Johnson RD (1973) Adult western corn rootworms: oviposition, fecundity, and longevity in the laboratory. *Journal of Economic Entomology* **66**, 417-418.
- Branson TF, Krysan JL (1981) Feeding and oviposition behavior and life cycle strategies of *Diabrotica*: an evolutionary view with implications for pest management. *Environmental Entomology* **10**, 826-831.
- Burke JM, Arnold ML (2001) Genetics and the fitness of hybrids. *Annual Review of Genetics* **35**, 31-52.
- Carlton JT (1996) Biological invasions and cryptogenic species. *Ecology* **77**, 1653-1655.
- Carrasco LR, Cook D, Baker R, *et al.* (2012) Towards the integration of spread and economic impacts of biological invasions in a landscape of learning and imitating agents. *Ecological Economics* **76**, 95-103.
- Carrasco LR, Harwood TD, Toepfer S, *et al.* (2010) Dispersal kernels of the invasive alien western corn rootworm and the effectiveness of buffer zones in eradication programmes in Europe. *Annals of Applied Biology* **156**, 63-77.
- Chapple DG, Miller KA, Kraus F, Thompson MB (2013) Divergent introduction histories among invasive populations of the delicate skink (*Lampropholis delicata*): has the importance of genetic admixture in the success of biological invasions been overemphasized? *Diversity and Distributions* **19**, 134-146.
- Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289-1303.
- Ciosi M (2009) Processus Démographiques et Génétiques à l'oeuvre au cours d'une Invasion Biologique: le cas de la Chrysomèle des racines du Maïs *Diabrotica virgifera virgifera* Thèse de Doctorat en Biologie, Université de Nice Sophia-Antipolis

- Ciosi M, Miller NJ, Kim KS, *et al.* (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology* **17**, 3614-3627.
- Ciosi M, Miller NJ, Toepfer S, Estoup A, Guillemaud T (2011) Stratified dispersal and increasing genetic variation during the invasion of Central Europe by the western corn rootworm, *Diabrotica virgifera virgifera*. *Evolutionary Applications* **4**, 54-70.
- Ciosi M, Toepfer S, Li H, *et al.* (2009) European populations of *Diabrotica virgifera virgifera* are resistant to aldrin, but not to methyl parathion. *Journal of Applied Entomology* **133**, 307-314.
- Coats SA, Tollefson JJ, Mutchmor JA (1986) Study of migratory flight in the western corn rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology* **15**, 620-625.
- Cohen AN, Carlton JT (1998) Accelerating invasion rate in a highly invaded estuary. *Science* **279**, 555-558.
- Colautti RI, MacIsaac HJ (2004) A neutral terminology to define 'invasive' species. *Diversity and Distributions* **10**, 135-141.
- Courchamp F, Chapuis JL, Pascal M (2003) Mammal invaders on islands: impact, control and control impact. *Biological Reviews* **78**, 347-383.
- Coyne JA (1989) Genetics of sexual isolation between 2 sibling species, *Drosophila simulans* and *Drosophila mauritiana*. Proceedings of the National Academy of Sciences of the United States of America **86**, 5464-5468.
- Csillery K, Blum MGB, Gaggiotti OE, Francois O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* **25**, 410-418.
- Darling JA, Bagley MJ, Roman J, Tepolt CK, Geller JB (2008) Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Molecular Ecology* **17**, 4992-5007.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* **17**, 431-449.
- Dobzhansky T, Gowen J (1952) Nature and origin of heterosis. *Heterosis.*, 218-223.
- Drake JM (2006) Heterosis, the catapult effect and establishment success of a colonizing bird. *Biology Letters* **2**, 304-307.
- Duan JJ, Weber DC, Dorn S (1998) Flight behavior of pre- and postdiapause apple blossom weevils in relation to ambient temperature. *Entomologia Experimentalis et Applicata* **88**, 97-99.
- Duncan RP, Blackburn TM, Sol D (2003) The ecology of bird introductions. *Annual Review of Ecology Evolution and Systematics* **34**, 71-98.
- Edmunds S (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* **53**, 1757-1768.
- Edmunds S (2002) Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution* **17**, 520-527.
- Edwards CR, Kiss J (2005) New WCR 2004 General Spread Map for Europe. <http://extension.entm.purdue.edu/wcr/>.
- Edwards CR, Kiss J (2008) New WCR 2007 General Spread Map for Europe. <http://extension.entm.purdue.edu/wcr/>.
- Edwards CR, Kiss J (2009) New WCR 2008 General Spread Map for Europe. *IWGO Newsletter* **29**.

- Edwards CR, Kiss J (2010) New WCR 2009 General Spread Map for Europe. *IWGO Newsletter* **30**.
- Eilenberg J, Hajek A, Lomer C (2001) Suggestions for unifying the terminology in biological control. *Biocontrol* **46**, 387-400.
- Ellegren H (2004) Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics* **5**, 435-445.
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7043-7050.
- Elton CS (1958) *The Ecology of Invasions by Animals and Plants* Methuen & Co LTD, London.
- Endler JA (1977) Geographic variation, speciation and clines. In: *Monographs in population biology*, pp. 30-96. Princeton University Press, New Jersey.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology* **19**, 4113-4130.
- Facon B, Jarne P, Pointier JP, David P (2005) Hybridization and invasiveness in the freshwater snail *Melanoides tuberculata*: hybrid vigour is more important than increase in genetic variance. *Journal of Evolutionary Biology* **18**, 524-535.
- Facon B, Pointier J-P, Glaubrecht M, *et al.* (2003) A molecular phylogeography approach to biological invasions of the New World by parthenogenetic Thiarid snails. *Molecular Ecology* **12**, 3027-3039.
- 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. *Current Biology* **18**, 363-367.
- Fauvergue X, Vercken E, Malausa T, Hufbauer RA (2012) The biology of small, introduced populations, with special reference to biological control. *Evolutionary applications* **5**, 424-443.
- Fitzpatrick BM (2008) Hybrid dysfunction: Population genetic and quantitative genetic perspectives. *The American Naturalist* **171**, 491-498.
- Fox CW, Czesak ME (2006) Selection on body size and sexual size dimorphism differs between host species in a seed-feeding beetle. *Journal of Evolutionary Biology* **19**, 1167-1174.
- Furlan L, Italy GdLD (2008) Update on the presence of western corn rootworm in Italy in 2007. Aggiornamento sulla presenza di *Diabrotica* in Italia nel 2007. *L'Informatore Agrario*, 44-46.
- Gassmann AJ, Petzold-Maxwell JL, Keweshan RS, Dunbar MW (2011) Field-Evolved Resistance to Bt Maize by Western Corn Rootworm. *PLoS ONE* **6**.
- Gay L, Crochet PA, Bell DA, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution* **62**, 2789-2806.
- Genton BJ, Shykoff JA, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Molecular Ecology* **14**, 4275-4285.
- George BW, and E.E. Ortman (1965) Rearing the western corn rootworm in the laboratory. *Journal of Economic Entomology* **58**, 375-377.

- Gillette CP (1912) *Diabrotica virgifera* Le Conte as a corn rootworm. . *Journal of Economic Entomology* **5**, 364-366.
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995).
- Grant PR, Grant BR (1994) Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution*, 297-316.
- Gray ME, Sappington TW, Miller NJ, Moeser J, Bohn MO (2009) Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annual Review of Entomology* **54**, 303-321.
- Haanes H, Roed KH, Solberg EJ, Herfindal I, Saether BE (2011) Genetic discontinuities in a continuously distributed and highly mobile ungulate, the Norwegian moose. *Conservation Genetics* **12**, 1131-1143.
- Hammack L, French BW (2007) Sexual dimorphism of basitarsi in pest species of *Diabrotica* and *Cerotoma* (Coleoptera : Chrysomelidae). *Annals of the Entomological Society of America* **100**, 59-63.
- Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology* **7**, 69-128.
- Harrison RG (1993) Hybrids and hybrid zones: historical perspective. *Hybrid zones and the evolutionary process*, 3-12.
- Hewitt GM (1988) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution* **3**, 158-167.
- Hill R, Hixson E, Muma MH (1948) Corn rootworm control with benzene hexachloride, DDT, nitrogen fertilizers and crop rotations. *Journal of Economic Entomology* **41**, 392-401.
- Hill RE (1975) Mating, oviposition patterns, fecundity and longevity of the western corn rootworm. *Journal of Economic Entomology* **68**, 311-315.
- Hufbauer RA, Sforza R (2008) Multiple introductions of two invasive *Centaurea* taxa inferred from cpDNA haplotypes. *Diversity and Distributions* **14**, 252-261.
- Huxley JS (1938) Clines: an auxiliary taxonomic principle. *Nature* **142**, 219-220.
- IUCN SSC Invasive Species Specialist Group (ISSG) (2009) About Invasive Species.
- James C (2009) Brief 41: Global status of commercialized biotech/GM crops: 2009. ISAAA Brief. Ithaca, New York: ISAAA.
- Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution* **15**, 250-255.
- Johnston JA, Donovan LA, Arnold ML (2004) Novel phenotypes among early generation hybrids of two Louisiana iris species: flooding experiments. *Journal of Ecology* **92**, 967-976.
- Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters* **11**, 852-866.
- Keller SR, Taylor DR (2010) Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology* **23**, 1720-1731.
- Kim KS, Sappington TW (2004) Boll weevil (*Anthonomus grandis* Boheman)(Coleoptera: Curculionidae) dispersal in the southern United States: evidence from mitochondrial DNA variation. *Environmental Entomology* **33**, 457-470.

- Kiss J, Edwards CR, Berger HK, *et al.* (2005) Monitoring of western corn rootworm (*Diabrotica virgifera virgifera* LeConte) in Europe 1992-2003. In: *Western Corn Rootworm: Ecology and Management* (eds. Vidal S, Kuhlmann U, Edwards CR), pp. 29-39. CABI Publishing, Cambridge, MA USA.
- Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* **16**, 199-204.
- Kolbe JJ, Glor RE, Schettino LRG, *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature* **431**, 177-181.
- Kolbe JJ, Larson A, Losos JB, de Queiroz K (2008) Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biology Letters* **4**, 434-437.
- Krysan J (1976) The early embryology of *Diabrotica undecimpunctata* Howardi (Coleoptera: Chrysomelidae). *Journal of Morphology* **149**, 121-138.
- Krysan JL (1972) Embryonic stages of *Diabrotica virgifera* (Coleoptera: Chrysomelidae) at diapause. *Annals of the Entomological Society of America* **65**, 768-769.
- Krysan JL (1982) Diapause in the Nearctic species of the virgifera group of *Diabrotica*: evidence for tropical origin and temperate adaptations. *Annals of the Entomological Society of America* **75**, 136-142.
- Krysan JL, Smith RF (1987) Systematics of the virgifera species group of *Diabrotica* (Coleoptera: Chrysomelidae: Galerucinae). *Entomography* **5**, 375-484.
- Krysan JLM, T. A [eds.] (1986) *Methods for the study of pest Diabrotica* Springer, New York.
- Kuhar TP, Youngman RR (1995) Sex ratio and sexual dimorphism in western corn rootworm (Coleoptera: Chrysomelidae) adults on yellow sticky traps in corn. *Environmental Entomology* **24**, 1408-1413.
- Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist* **142**, 911-927.
- Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3883-3888.
- Leung B, Lodge DM, Finnoff D, *et al.* (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**, 2407-2413.
- Levine E, Oloumi-Sadeghi H (1991) Management of diabroticite rootworms in corn. *Annual Review of Entomology* **36**, 229-255.
- Levine E, OloumiSadeghi H (1996) Western corn rootworm (Coleoptera: Chrysomelidae) larval injury to corn grown for seed production following soybeans grown for seed production. *Journal of Economic Entomology* **89**, 1010-1016.
- Levine E, Spencer JL, Isard SA, Onstad DW, Gray ME (2002) Adaptation of the Western corn rootworm to crop rotation: evolution of a new strain in response to a management practice. *American Entomologist* **48**, 94-107.
- Levine JM, Adler PB, Yelenik SG (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters* **7**, 975-989.
- Li H, Toepfer S, Kuhlmann U (2009) Relationship between phenotypic traits and selected fitness components of *Diabrotica virgifera virgifera*. *Entomologia Experimentalis et Applicata* **131**, 254-263.

- Li H, Toepfer S, Kuhlmann U (2010) Flight and crawling activities of *Diabrotica virgifera virgifera* (Coleoptera: Chrysomelidae) in relation to morphometric traits. *Journal of Applied Entomology* **134**, 449-461.
- Liebholt AM, Tobin PC (2008) Population Ecology of Insect Invasions and Their Management*. *Annu. Rev. Entomol.* **53**, 387-408.
- Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. *Trends in Genetics* **23**, 60-66.
- Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution* **20**, 223-228.
- Lockwood JL, Hoopes MF, Marchetti MP (2007) *Invasion ecology* Blackwell, Malden, MA.
- Lombaert E, Guillemaud T, Cornuet JM, *et al.* (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS ONE* **5**.
- Lombaert E, Malausa T, Devred R, Estoup A (2008) Phenotypic variation in invasive and biocontrol populations of the harlequin ladybird, *Harmonia axyridis*. *Biocontrol* **53**, 89-102.
- Lucek KAY, Roy D, Bezault E, Sivasundar A, Seehausen OLE (2010) Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Molecular Ecology* **19**, 3995-4011.
- Lynch M (1991) The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**, 622-629.
- Mabry TR, Spencer JL, Levine E, Isard SA (2004) Western corn rootworm (Coleoptera: Chrysomelidae) behavior is affected by alternating diets of maize and soybean. *Environmental Entomology* **33**, 860-871.
- Mack RN, Simberloff D, Lonsdale WM, *et al.* (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689-710.
- Marrs RA, Sforza R, Hufbauer RA (2008) When invasion increases population genetic structure: a study with *Centaurea diffusa*. *Biological Invasions* **10**, 561-572.
- Marshall TC, Spalton JA (2000) Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. *Animal Conservation* **3**, 241-248.
- Mayr E (1963) Animal species and evolution. In: *Animal species and their evolution*. MA: Harvard University Press, Cambridge, UK.
- McKinney ML, Lockwood JL (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution* **14**, 450-453.
- Meinke LJ, Siegfried BD, Wright RJ, Chandler LD (1998) Adult susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to selected insecticides. *Journal of Economic Entomology* **91**, 594-600.
- Miller N, Estoup A, Toepfer S, *et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**, 992.
- Modic S, Shroers H-J, Urek G (2008) Survival of eggs of *Diabrotica virgifera virgifera* during winter and isolation of some of their fungal enemies. In: *Closing EU-project meeting DIABR-ACT, 2008* Gottingen.
- Monecke S, Monecke H, Monecke J (2009) Modelling the black death. A historical case study and implications for the epidemiology of bubonic plague. *International Journal of Medical Microbiology* **299**, 582-593.
- Moody ML, Les DH (2002) Evidence of hybridity in invasive watermilfoil (*Myriophyllum*) populations. *Proceedings of the National Academy of Sciences* **99**, 14867-14871.

- Moyle PB, Light T (1996) Fish invasions in California: do abiotic factors determine success? *Ecology* **77**, 1666-1670.
- Murphy AF, Krupke CH (2011) Mating Success and Spermatophore Composition in Western Corn Rootworm (Coleoptera: Chrysomelidae). *Environmental Entomology* **40**, 1585-1594.
- Naranjo SE (1991) Influence of temperature and larval density on flight performance of *Diabrotica virgifera virgifera* Leconte (Coleoptera: Chrysomelidae). *Canadian Entomologist* **123**, 187-196.
- Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.
- Ogutu-Ohwayo R (2001) Nile perch in Lake Victoria: balancing the costs and benefits of aliens. In: *Invasive Species and Biodiversity Management*, pp. 47-63. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution* **19**, 18-24.
- Orians CM (2000) The effects of hybridization in plants on secondary chemistry: implications for the ecology and evolution of plant-herbivore interactions. *American Journal of Botany* **87**, 1749-1756.
- Orr HA, Turelli M (2001) The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085-1094.
- Petit RJ, El Mousadik A, Pons O (1998) Identifying Populations for Conservation on the Basis of Genetic Markers. *Conservation Biology* **12**, 844-855.
- Phillips BL, Brown GP, Shine R (2010) Life-history evolution in range-shifting populations. *Ecology* **91**, 1617-1627.
- Pimentel D, McNair S, Janecka J, *et al.* (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture Ecosystems & Environment* **84**, 1-20.
- Pimentel D, Wilson C, McCullum C, *et al.* (1997) Economic and environmental benefits of biodiversity. *BioScience* **47**, 747-757.
- Piry S, Alapetite A, Cornuet JM, *et al.* (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536-539.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Pysek P, Richardson DM (2006) The biogeography of naturalization in alien plants. *Journal of Biogeography* **33**, 2040-2050.
- Pysek P, Sadlo J, Mandak B, Jarosik V (2003) Czech alien flora and the historical pattern of its formation: what came first to Central Europe? *Oecologia* **135**, 122-130.
- R Development Core Team (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9197-9201.
- Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution* **49**, 1280-1283.
- Raymond M, Rousset F (1995b) Genepop (version. 1.2), a population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

- REX (2013) Heterogeneity of selection and the evolution of resistance. *Trends in Ecology & Evolution* **28**, 110-118.
- Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112**, 183-198.
- Richardson DM, Pysek P, Rejmanek M, *et al.* (2000) Naturalization and invasion of alien plants: Concepts and definitions. *Diversity and Distributions* **6**, 93-107.
- Ricklefs RE, Bermingham E (2002) The concept of the taxon cycle in biogeography. *Global Ecology and Biogeography* **11**, 353-361.
- Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. *Heredity* **83**, 363-372.
- Rieseberg LH, Raymond O, Rosenthal DM, *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211-1216.
- Rieux A, Lenormand T, Carlier J, Lapeyre de Bellaire L, Ravigné V (2013) Using neutral cline decay to estimate contemporary dispersal: a generic tool and its application to a major crop pathogen. *Ecology letters*.
- Rollins LA, Woolnough AP, Wilton AN, Sinclair R, Sherwin WB (2009) Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Molecular Ecology* **18**, 1560-1573.
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution* **22**, 454-464.
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.
- Ruiz GM, Rawlings TK, Dobbs FC, *et al.* (2000) Global spread of microorganisms by ships - Ballast water discharged from vessels harbours a cocktail of potential pathogens. *Nature* **408**, 49-50.
- Sakai AK, Allendorf FW, Holt JS, *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics* **32**, 305-332.
- Sappington W, Siegfried BD, Guillemaud T (2006) Coordinated *Diabrotica* Genetics Research: Accelerating Progress on an Urgent Insect Pest Problem. *American Entomologist*, 92-99.
- Sax DF, Stachowicz JJ, Brown JH, *et al.* (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution* **22**, 465-471.
- Schierup MH, Christiansen FB (1996) Inbreeding depression and outbreeding depression in plants. *Heredity* **77**, 461-468.
- Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. *American Naturalist* **146**, 229-251.
- Shull GH (1914) Duplicate genes for capsule-form in *Bursa bursa pastoris*. *Molecular and General Genetics MGG* **12**, 97-149.
- Singh P, Moore RF (1999) *Handbook of insect rearing*. Elsevier Amsterdam, The Netherlands.
- Sivcev I, Kljajic P, Kostic M, Sivcev L, Stankovic S (2012) Management of Western corn rootworm (*Diabrotica virgifera virgifera*). *Pesticidi i Fitomedicina* **27**, 189-201.
- Sivcev I, Manojlovic B, Krnjajic S, *et al.* (1994) Distribution and harmfulness of *Diabrotica virgifera* LeConte (Coleoptera, Chrysomelidae), a new maize pest in Yugoslavia. *Zastita Bilja* **45**, 19-26.

- Smith RF (1966) Distributional patterns of selected western north american insects: the distribution of diabroticites in western north America. *Bulletin of Entomology of society of America* **12**, 108-110.
- Spencer JL, Hibbard BE, Moeser J, Onstad DW (2009) Behaviour and ecology of the western corn rootworm (*Diabrotica virgifera virgifera* LeConte). *Agricultural and Forest Entomology* **11**, 9-27.
- Spencer JL, Levine E, Isard SA (1997) Corn rootworm injury to first-year corn: New research findings, 73-81.
- Staetz CA, Ball HJ, Carlson SD (1976) Antennal morphology of *Diabrotica virgifera* adults (Coleoptera: Chrysomelidae). *Annals of the Entomological Society of America* **69**, 695-698.
- Suarez AV, Holway DA, Case TJ (2001) Patterns of spread in biological invasions dominated by long-distance jump dispersal: Insights from Argentine ants. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 1095-1100.
- Suarez AV, Holway DA, Ward PS (2005) The role of opportunity in the unintentional introduction of nonnative ants 10.1073/pnas.0506119102. *PNAS* **102**, 17032-17035.
- Sutherland GR, Richards RI (1995) Simple tandem DNA repeats and human genetic-disease. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3636-3641.
- Szalai M, Papp Komáromi J, Bažok R, *et al.* (2011) Generational growth rate estimates of *Diabrotica virgifera virgifera* populations (Coleoptera: Chrysomelidae). *Journal of Pest Science* **84**, 133-142.
- Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution* **19**, 489-496.
- Tatem AJ, Hay SI (2007) Climatic similarity and biological exchange in the worldwide airline transportation network. *Proceedings of the Royal Society B-Biological Sciences* **274**, 1489-1496.
- Tatem AJ, Hay SI, Rogers DJ (2006) Global traffic and disease vector dispersal. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6242-6247.
- Templeton AR (2009) Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. *Molecular Ecology* **18**, 319-331.
- Templeton AR (2010) Coherent and incoherent inference in phylogeography and human evolution. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6376-6381.
- Toepfer S, Kuhlmann U (2006) Constructing life-tables for the invasive maize pest *Diabrotica virgifera virgifera* (Col.: Chrysomelidae) in Europe. *Journal of Applied Entomology* **130**, 193-205.
- Toepfer S, Levay N, Kiss J (2005) Suitability of different fluorescent powders for mass-marking the Chrysomelid, *Diabrotica virgifera virgifera* LeConte. *Journal of Applied Entomology* **129**, 456-464.
- Turelli M, Orr HA (2000) Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* **154**, 1663-1679.
- Turgeon J, Tayeh A, Facon B, *et al.* (2011) Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *Journal of Evolutionary Biology* **24**, 1044-1052.

- Urbanelli S, Bellini R, Carrieri M, Sallicandro P, Celli G (2000) Population structure of *Aedes albopictus* (Skuse): the mosquito which is colonizing Mediterranean countries. *Heredity* **84**, 331-337.
- Vahed K (1998) The function of nuptial feeding in insects: a review of empirical studies. *Biological Reviews* **73**, 43-78.
- Verhoeven KJF, Macel M, Wolfe LM, Biere A (2011) Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B-Biological Sciences* **278**, 2-8.
- Vitousek PM, Dantonio CM, Loope LL, Rejmanek M, Westbrooks R (1997) Introduced species: A significant component of human-caused global change. *New Zealand Journal of Ecology* **21**, 1-16.
- Wang H, McArthur ED, Sanderson SC, Graham JH, Freeman DC (1997) Narrow hybrid zone between two subspecies of big sagebrush (*Artemisia tridentata*: Asteraceae) .IV. Reciprocal transplant experiments. *Evolution* **51**, 95-102.
- Wang J (2004) Sibship Reconstruction From Genetic Data With Typing Errors. *Genetics* **166**, 1963-1979.
- Wares JP, Hughes AR, R.K. G (2005) Mechanisms that drive evolutionary change: insights from species introductions and invasions. In: *Species Invasions: Insights into Ecology, Evolution and Biogeography* (eds. Sax DF, Stachowicz JJ, Gaines SD), pp. 229-257. Sinauer Associates Inc, Sunderland, MA USA.
- Watkins WS, Bamshad M, Jorde LB (1995) Population-genetics of trinucleotide repeat polymorphisms. *Human Molecular Genetics* **4**, 1485-1491.
- Weir BS, Cockerham C (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
- Wesseler J, Fall EH (2010) Potential damage costs of *Diabrotica virgifera virgifera* infestation in Europe - the 'no control' scenario. *Journal of Applied Entomology* **134**, 385-394.
- Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, 433-458.
- Williamson M (1996) *Biological invasions* Chapman and Hall, London, UK.
- Williamson M (2006) Explaining and predicting the success of invading species at different stages of invasion. *Biological Invasions* **8**, 1561-1568.
- Williamson M, Fitter A (1996) The varying success of invaders. *Ecology* **77**, 1661-1666.
- Wright RJ, Meinke LJ, Siegfried BD (1996) Corn rootworm management and insecticide resistance management, 45-53.
- Zhong N, Ye LL, Dobkin C, Brown WT (1994) Fragile-X founder chromosome effects-linkage disequilibrium or microsatellite heterogeneity. *American Journal of Medical Genetics* **51**, 405-411.