THESE DE DOCTORAT

DE L'UNIVERSITE DE PARIS XI - ORSAY

Ecole Doctorale : Gènes, Génomes, Cellules

BIOLOGIE

Spécialité : Génétique des Populations & Phylogéographie

Michel SEZONLIN

PHYLOGEOGRAPHIE ET GENETIQUE DES POPULATIONS DU FOREUR DE TIGES DE CEREALES *BUSSEOLA FUSCA* (FULLER) (LEPIDOPTERA, NOCTUIDAE) EN AFRIQUE SUBSAHARIENNE, IMPLICATIONS POUR LA LUTTE BIOLOGIQUE CONTRE CET INSECTE

Soutenue le 04 décembre 2006

JURY

Pierre CAPY	Président	Professeur, Université Paris XI
Frédéric FLEURY	Rapporteur	Professeur, Université Lyon I
Jane LECOMTE	Rapporteur	Maître de Conférences, Université Paris XI
Stéphane DUPAS	Examinateur	Chargé de Recherche, IRD
Marie-Louise CARIOU	Directrice de thèse	Directrice de Recherche, CNRS
Jean-François SILVAIN	Directeur de thèse	Directeur de Recherche, IRD

Résumé : *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae) est un ravageur majeur du maïs et du sorgho cultivé en Afrique Subsaharienne. Des différences observées dans les traits de vie et l'écologie de l'espèce entre les populations Est et Ouest africaines suggèrent l'existence de populations génétiquement différenciées selon la géographie ou les types de biotopes. Le rôle des plantes hôtes, et notamment celui des deux principales plantes cultivées consommées par l'insecte, dans l'éventuelle structuration génétique de ses populations est à déterminer. Une vaste étude phylogéographique basée sur l'utilisation d'un marqueur mitochondrial, le cytochrome *b* et sur l'analyse de nombreux échantillons (590 individus provenant de 112 localités incluant la majeure partie de l'aire de distribution de foreur), a été réalisée. Cette étude a révélé l'existence de trois clades principaux d'haplotypes (*W*, *KI*, *KII*) correspondant à des populations isolées à l'Ouest et à l'Est de l'Afrique au Pléistocène, il y a environ un million d'années. Le clade *W* localisé en Afrique de l'Ouest est isolé géographiquement des clades *KI* et *KII* qui sont partiellement sympatriques. Le clade *KI* voit sa distribution limitée à une partie de l'Afrique de l'Est alors que le clade *KII* présente une large distribution géographique couvrant l'Est, le Centre et le Sud de l'Afrique Subsaharienne. Les résultats des analyses phylogénétiques, les paramètres démographiques calculés, les statistiques de Wright ainsi que les analyses des clades emboîtés confirment que ces trois populations, après avoir été isolées dans des aires refuges différentes, ont connu une expansion démographique et géographique, même si les populations locales sont caractérisées par des phénomènes de restriction des flux géniques avec isolement par la distance. L'analyse de mismatch distribution et les valeurs négative de l'indice D de Tajima sont en accord avec l'hypothèse d'une expansion démographique des trois clades. Des différenciations génétiques significatives ont été mises en évidence à différents niveaux hiérarchiques par l'analyse moléculaire de la variance (AMOVA). La Ligne Volcanique du Cameroun et la Rift Valley seraient deux des facteurs qui ont modelé la structure génétique des populations de *B. fusca*. La plus grande diversité haplotypique et nucléotidique de *B. fusca* dans les régions ghanéenne (clade *W*), érythréenne (clade *KI*) et kenyane (clade *KII*) a permis de les identifier comme les centres probables d'origine géographique de chaque clade. L'histoire génétique des populations de *B. fusca*, telle que révélée par l'analyse du génome mitochondrial, apparaît indépendante de la domestication du sorgho et de l'introduction et de l'expansion des cultures de maïs. L'ancienne structure génétique est maintenue à travers les différentes époques avec un passage récurrent et local des individus de *B. fusca* des plantes hôtes sauvages aux cultivées. Un parallèle original a pu être fait entre les résultats de l'étude phylogéographique de cet insecte graminivore et ceux de plusieurs études portant sur des mammifères herbivores africains. Cette analyse parallèle montre que des facteurs paléo-climatiques similaires ont probablement modelé les populations des animaux de groupes éloignés associés aux milieux graminéens africains.

Mots clés : *Busseola fusca*, cytochrome *b*, génétique des populations, phylogéographie, Pléistocène, histoire évolutive régionale, biogéographie africaine.

Abstract: *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae) is a major pest of maize and cultivated sorghum in sub-Saharan Africa. The observed difference in the life features and the ecology of species among East and West African populations suggest the existence of populations genetically differentiated in accordance with the geography or biotope types. The role of host plant, particularly that related to two major cultivated plant consumed by insect, in possible of genetic structure of its populations is being established. A large phylogeographic study based on use of one mitochondrial marker, the cytochrome *b* and on analysis of numerous samplings (590 individuals from 112 localities including the major part of the spatial distribution of borer) has been performed. This study has showed the presence of three main haplotype clades (*W*, *KI*, *KII*) corresponding to populations isolated in West and East Africa in Pleistocene, around one million years ago. Clade *W* localized in West Africa split geographically from clade *KI* and clade *KII* that are partially sympatric. Clade *KI* is limited to one part of East Africa whereas clade *KII* shows a large geographical distribution covering well East, Central and Southern sub-Saharan Africa. Phylogenetic, F-statistics, calculated demographic parameters and nested clade phylogeographic analyses results confirmed that the clades, after their isolation in three different refuge areas, experienced geographic and demographic expansion even if local populations were characterized by the phenomena of restricted gene flow with isolation by distance. Mismatch distribution analysis and the negative values of Tajima D index are consistent with a demographic expansion hypothesis of three clades. Significant genetic differentiations have been highlighted at various hierarchical levels by analysis of molecular variance (AMOVA). The Cameroon Volcanic Line and the Rift Valley appear to be two factors contributing to the genetic structure of *B. fusca* populations. The highest haplotype and nucleotide diversity in Ghanaian (clade *W*), Eritrean (clade *KI*) and Kenyan (clade *KII*) regions has allowed identifying them as the likelihood geographic centres of origin of each clade. The population genetic history of *B. fusca* as revealed by mitochondrial genome analysis appears independent to sorghum domestication and introduction and expansion of maize. The ancient genetic structure is maintained through different periods with recurrent and local shift of *B. fusca* individuals from wild host plants to cultivated. An original parallel has been able to be performed with the results of phylogeographic study of this graminaceous insect and all those of many studies related on African mammalian herbivorous. This parallel analysis indicates that similar paleoclimatic factors have likely shaped animal populations from distant groups associated with African graminaceous environments.

Keywords: *Busseola fusca*, cytochrome *b*, population genetics, phylogeography, Pleistocene, regional evolutionary history, African biogeography.

Remerciements

Je voudrais exprimer ici ma gracieuse reconnaissance à cette foule de personnes enthousiastes qui ont participé plus ou moins directement au déroulement et à la concrétisation de ce travail de thèse.

Mes profondes reconnaissances vont en premier à mes directeurs de thèse, Madame Marie-Louise CARIOU et Monsieur Jean-François SILVAIN, pour les nombreux efforts qu'ils ont consentis : ils ont assuré tout au long de ma formation doctorale un encadrement indéfectible tant sur place ici en France comme au Bénin. Ils ont su me proposer un sujet intéressant et très moderne. J'ai pu bénéficier de leurs nombreuses expériences professionnelles et fructueuses.

Mes gratitudes vont à Monsieur Pierre CAPY pour ses nombreuses aides aussi bien administratives que scientifiques.

Monsieur Hervé LE GUYADER, responsable du DEA Biodiversité : Génétique, Histoire, Mécanismes de l'Evolution, a accepté sans condition mon inscription. Il a ainsi facilité mes premiers accueils en France dont l'aboutissement a conduit à cette soutenance de thèse. Qu'il reçoive toutes mes reconnaissances.

J'exprime mes multitudes gratitudes aux membres de jury qui ont aimablement accepté de juger ce travail de thèse.

Je remercie le Service de la Coopération et d'Action Culturelle (SCAC) de l'Ambassade de France à Cotonou (République du Bénin) et le Département de Soutien et Formation (DSF) des Communautés Scientifiques du Sud de l'Institut de Recherche et de Développement (IRD) pour avoir accepté de financer cette formation doctorale.

Une mention spéciale à mes amis Nathalie FAURE et Gaël KERGOAT pour toutes les affections et les contributions scientifiques dont ils ont fait preuve au cours de mes différents séjours en France.

Je dis mon merci à Stéphane DUPAS pour ses remarques, ses commentaires pertinents au cours des différentes rédactions de manuscrit d'articles scientifiques et de thèse et son soutien moral infaillible. Il m'a initié à l'approche phylogéographique de la biologie évolutive.

Messieurs Bruno LE RU, Paul-André CALATAYUD, Philippe LE GALL et Pascal MOYAL m'ont beaucoup aidé à disposer de matériels biologiques et à les identifier. Je les remercie pour leur contribution scientifique hautement louable.

Jean-Luc DA LAGE, Frédérique MACZKOWIAK et Emmanuelle RENARD m'avaient initié aux différentes techniques de biologie moléculaire. Je les remercie de tout mon cœur.

Un merci particulier à David OGEREAU toujours souriant avec une humeur très agréable.

Jean-Bernard EMOND et Olivier LANGELLA m'ont apporté beaucoup d'aides bioinformatiques : infiniment merci.

Je remercie très sincèrement Mesdames Françoise LEMEUNIER, Sylvie AULARD-WIDEMANN, Evelyne SIMONEAU et Messieurs Jean Roy DAVID, Claude MAISONHAUTE pour leurs sympathies et leurs conseils de tous les jours. J'ai vécu les 5000 Kms entre le Bénin et la France avec une passion agréable.

Mes remerciements vont également aux différents étudiants et à tout le personnel du Laboratoire Evolution, Génomes et Spéciation, CNRS de Gif-sur-Yvette en France.

Mes stages en alternance se sont également déroulés à International Institut of Tropical Agriculture (IITA), Station du Bénin.

Je tiens à exprimer mes gratitudes particulières à Messieurs Peter NEUENSCHWANDER et James BRAIMA, respectivement ancien et nouveau directeur de la station de l'IITA - Bénin, pour m'avoir permis de réaliser une partie de mes stages à l'IITA - Bénin.

Mes remerciements vont aussi à Monsieur Georg GOERGEN, responsable du Centre de Biodiversité de l'IITA - Bénin, pour ses nombreux soutiens logistiques, matériels et surtout ses précieux conseils. Il m'a surtout fourni, au cours de mes séjours dans l'Institut, un cadre agréable de travail me permettant de parachever la rédaction de ce manuscrit et d'échanger régulièrement les correspondances avec mes directeurs de thèses.

Merci à Monsieur Manuel TAMO pour ses nombreux apports scientifiques et ses nombreux conseils constructifs.

Un grand merci à mes amis Prosper AMOUSSOU-GUENOU, Francis TCHIBOZO et William Pamphile IDOHOU du Centre de Biodiversité IITA pour avoir réussi à créer une ambiance de travail hors norme.

Cette formation doctorale a été l'initiative du Laboratoire de Génétique de la Faculté des Sciences et Techniques (FAST) de l'Université d'Abomey-Calavi (ex-UNB) de la République du Bénin.

Mes gratitudes vont à Messieurs Ogoubi DAINOU et Serge TOSTAIN qui ont guidé mes premiers pas pour que je me familiarise avec cette discipline en plein essor qu'est la Génétique. Ils n'ont ménagé aucun effort pour me trouver un financement afin que je puisse poursuivre les études doctorales.

Je remercie tout le personnel du Laboratoire de Génétique et en particulier son directeur Monsieur Clément AGBANGLA pour leur compagnie.

Messieurs Nestor SAKITI, Martin AKOGBETO, respectivement ancien et nouveau chef du département de Zoologie et de Génétique, et Monsieur Cyprien GNANVO, Doyen de la Faculté des Sciences et techniques de l'Université d'Abomey-Calavi, m'ont facilité les formalités administratives. Je leur exprime ici ma reconnaissance.

Merci à toutes les personnes de bonne volonté de la Faculté des Sciences et Techniques de l'Université d'Abomey-Calavi qui m'ont apporté leur contribution de diverses manières au cours de cette formation.

Ma famille pour qui la séparation a été très difficile, mais, qui a su concilier la passion et la raison, le sentiment et l'avenir.

Je dis à ma grande sœur Marguerite un merci particulier pour toutes ses affections et ses assistances.

Je remercie infiniment mes bien-aimés parents, ma mère Eugénie WATCHINOU et surtout mon père Pierre Ahissou SEZONLIN, à qui je dois beaucoup mon éducation : ils ont toujours cru en moi et ont réussi à faire de moi tout ce que je suis de bon. Ce mémoire est le couronnement de succès de tous les efforts et les nombreux sacrifices qu'ils ont consentis.

Et enfin, un énorme et spécial merci à mon épouse Christine OUINSAVI qui partage ma vie ; me soutient et m'encourage depuis toujours… et surtout qui me rend tellement heureux devant les épreuves de la vie.

SOMMAIRE

Introduction générale 01

Chapitre 1 : Les outils moléculaires en entomologie appliquée 07

Section I : Les principaux types de marqueurs moléculaires 08

1.1. Historique 08

1.2. Les enzymes protéiques 08

1.3. Les marqueurs moléculaires nucléotidiques 12

1.3.1. Les fragments d'ADN 12

1.3.2. Les séquences nucléotidiques 18

1.4. Intérêts 25

Section II : Application et utilisation des outils moléculaires en biologie évolutive des populations et des espèces d'insectes 26

2.1. Etude des capacités de dispersion 26

2.2. Histoire des populations et des espèces 28

2.3. Evolution de la phytophagie : races d'hôtes, résistance aux insecticides 28

2.4. Espèces cryptiques 30

2.5. Rôle des espèces modèles (I) : patrons biogéographiques 31

2.6. Rôle des espèces modèles (II) : mécanismes de changement d'hôte et origine des insectes ravageurs 31

2.7. Quelques travaux de biologie moléculaires sur les espèces d'insectes ravageurs de cultures 34

Chapitre 2 : Le modèle *Busseola fusca* 36

Section I : Biologie et écologie de *B. fusca* 37

1.1. Biologie de *B. fusca* 37

1.2. Ecologie de *B. fusca* 39

Section II : Problématique générale de la thèse 45

Chapitre 3 : Matériel et méthodes 50

Section I : Les données biologiques 51

1.1. Echantillonnage des individus de *B. fusca* 51

1.2. Elevage et conservation des papillons 56

Section II : Choix et mise et au point des marqueurs et obtention des données moléculaires 57

2.1. Choix des marqueurs moléculaires 57

2.2. Extraction d'ADN génomique 58

2.3. Amplification et séquençage d'un fragment du gène de cytochrome *b* 58

2.3.1. Amplification à la PCR et purification des produits 58

2.3.2. Réaction de séquence et précipitation 59

2.3.3. Séquençage et alignement des séquences nucléotidiques 59

Section III : Les méthodes d'études en génétique des populations et en phylogéographie / Analyse des données moléculaires du cytochrome *b* 62

3.1. Reconstruction de la phylogénie des haplotypes 62

3.2. Structure et diversité génétique des populations de *B. fusca* 63

3.3. Démographie des populations de *B. fusca* 64

3.4. Les méthodes synthétiques : la NCPA 64

Chapitre 4 : Résultats et discussion 66

Section I : Les articles publiés 67

1.1. Article 1 : **Phylogeography and population genetics of the maize stalk borer *Bussseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa** 67

M. SEZONLIN, S. DUPAS, B. LE RÜ, P. LE GALL, P. MOYAL, P-A. CALATAYUD, I. GIFFARD, N. FAURE and J-F. SILVAIN**.** ***Molecular Ecology (2006) 15, 407-420***

(Phylogéographie et génétique des populations du foreur de tige de maïs Busseola fusca (Lepidoptera, Noctuidae) en Afrique Subsaharienne)

Résumé 68

Abstract 70

Introduction 70

Material and methods 71

Results 75

Discussion 77

Acknowledgments 81

References 81

1.2. Article 2 : **Phylogeographic pattern and regional evolutionary history of the maize stalk borer *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae) in sub-Saharan Africa** 84

Michel SEZONLIN, Stéphane DUPAS, Bruno LE RÜ, Nathalie FAURE, Philippe LE GALL and Jean-François SILVAIN, ***Sous presse dans les Annales de la Société Entomologique de la France***

(Patron phylogéographique et histoire évolutive régionale du foreur de tige de maïs Busseola fusca (Lepidoptera, Noctuidae) en Afrique Subsaharienne)

Résumé 85
Abstract 87
Introduction 89
Material and methods 90
Results 96
Discussion 102
Acknowledgments 107
References 107

Section II : Etudes complémentaires des populations de *B. fusca* du Cameroun 113

2.1. La diversité moléculaire 113

2.2. Structure génétique fine des populations de *B. fusca* 115

2.3. Analyse phylogéographique des clades emboîtés de *B. fusca* 116

Chapitre 5 : Discussion générale et perspectives 118

Bibliographie (hors articles) 135

Introduction générale

Les insectes (Hexapoda, Insecta) constituent l'un des groupes d'organismes vivants les plus importants et les plus diversifiés de la planète (Borror *et al*. 1989). Ils ont été depuis des centaines de millions d'années, et sont toujours, les animaux terrestres et aquatiques s'adaptant le mieux à des changements de conditions alimentaires, climatiques et à la compétition avec les autres animaux (Wigglesworth 1948). Le nombre d'espèces d'hexapodes connus est estimé à environ un million, soit au moins 55 % de toutes les espèces décrites jusqu'à ce jour (Kumar 1991). Ils représentent aujourd'hui environ 5/6 des animaux connus selon Bland & Jaques (1978).

En tant qu'indicateurs de biodiversité, les insectes permettent de comprendre la biogéographie d'une région donnée. Ils ont plus directement un impact considérable sur l'agriculture (les insectes ravageurs, pollinisateurs, et parasitoïdes des ravageurs de cultures) et la santé humaine (les agents vecteurs de maladie). Ainsi, étudier les patrons de diversité et comprendre les mécanismes de diversification des insectes peut contribuer à résoudre des problèmes agricoles et médicaux.

Jusqu'à une époque récente, les études en systématique des insectes étaient basées sur la description des caractères morphologiques. Celles-ci ont permis de décrire des taxons qui pour certains insectes sont toujours valides. D'autres ont été redéfinis, et parfois entièrement remis en cause, suite à l'intégration des analyses moléculaires récentes.

Les progrès en biologie moléculaire ont révolutionné l'étude de ce groupe animal, dans les domaines de la génétique des populations, de l'écologie moléculaire et chimique, de la biosystématique et de la physiologie moléculaire.

L'utilisation de marqueurs moléculaires se généralise chez les insectes, notamment chez les insectes phytophages, c'est à dire les hexapodes consommateurs de végétaux (Thomas *et al*. 2003 ; Prowell *et al*. 2004, Alvarez *et al*. 2005). On note cependant que les taxons les plus étudiés sont européens ou nord-américains. A ce jour, très peu d'analyses moléculaires en génétique des populations ont concerné des insectes africains en général (*Dorylus molestus*, Kronauer *et al*. 2004) et en particulier les ravageurs des cultures (*Eldana saccharina*, King *et al*. 2002 ; *Ceratitis rosa* et *Ceratitis fasciventris*, Baliraine *et al*. 2004). Or, la connaissance de la diversité génétique d'un insecte ravageur d'importance économique majeure est essentielle pour la définition de stratégies de lutte adaptées. L'identification des populations écologiquement et génétiquement différenciées (races ou lignées géographiques, races d'hôtes) chez les insectes ravageurs de cultures et leur caractérisation à l'aide de marqueurs moléculaires est un moyen efficace pour définir ces stratégies. Cette caractérisation moléculaire permettra d'appréhender l'histoire évolutive et phylogéographique

des insectes ravageurs de plantes cultivées en vue de comprendre les causes géographiques ou écologiques qui ont conduit à la différenciation génétique de leurs populations. En outre, la connaissance de la phylogéographie d'un insecte africain peut aussi apporter des enseignements nouveaux dans le domaine de la compréhension des facteurs à l'origine des patrons biogéographiques en Afrique.

Les foreurs de tiges de céréales cultivées constituent un problème agro-économique majeur en Afrique Subsaharienne. Si les différents programmes de lutte contre ces foreurs de tiges céréalières ont connu très peu de succès, cela est peut-être en partie dû à la méconnaissance de la structure génétique de leurs populations, de leur diversité et de leur histoire évolutive. A cela, s'ajoute l'absence de parasitoïdes spécifiques à chaque espèce d'insectes ravageurs.

Plusieurs facteurs influencent les pertes de rendement dues aux foreurs de tiges. La perte de la production du sorgho est étroitement corrélée avec la date d'infestation. En Afrique du Sud, la perte de la production de maïs liée aux activités de *B. fusca* est estimée à 10 % de la production totale (Barrow 1987 ; Mally 1920 ; Matthee 1974 ; Van Rensburg 1987). Dans ce cas, et contrairement aux travaux réalisés par Gounou *et al*. (1994) en Afrique de l'Ouest, les estimations de la perte de production de maïs due à *B. fusca* sont non seulement corrélées significativement avec le dégât foliaire, mais aussi avec le pourcentage de tiges forées. Au Lesotho, la variation saisonnière de la perte de rendement du maïs liée aux dégâts causés par *B. fusca* varie de 0,4 et 36,6 % (Ebenebe 1999). En Tanzanie (Walker 1960) et au Kenya (Walker & Hodson 1976), un taux d'infestation de 10 % des plantes par *B. fusca* entraîne une perte de 12 % de la production du grain de maïs. Au Burundi, on note une réduction de la production de maïs allant de 30 à 50 % liée aux dégâts causés par *B. fusca* (Muyango 1987). Dans la région nord du Nigeria, en zone de savane guinéenne où *B. fusca* est abondant, on enregistre une perte de la production du sorgho de l'ordre de 49 % (Ajayi 1998).

Seshu *et al*. (1989), travaillant sur sorgho cultivé, ont noté qu'une perte élevée de l'ordre de 80 % de la production des grains est associée à une infestation de huit chenilles de *Chilo partellus* par plante âgée de 21 jours alors qu'une telle densité d'infestation n'a significativement pas d'effet sur les plantes âgées de soixante jours après la germination. Ils conclurent que la perte de production de grains de sorgho liée à l'infestation par *C. partellus* était fortement influencée par les cultivars, l'âge de la culture au moment de l'infestation et aussi la densité de foreurs par tige. Parallèlement, McFarane (1990) observe que les plantes de sorgho infestées par *B. fusca* entre la germination et le bouturage subissent une perte

significativement importante. Les bons rendements des produits céréaliers sont obtenus seulement si les cultures sont traitées avec des insecticides entre 15ème et 30ème jour après la germination des semis : ils sont protégés de l'attaque des insectes foreurs de tiges. Une infestation artificielle avec les ravageurs de tiges au 15è jour après la germination des semis entraîne une perte cruciale des céréales : elle est de 60 à 62 % pour les parcelles du sorgho cultivé non traitées avec des insecticides (Taneja & Nwanze 1989). Il a été montré au Zimbabwe que les foreurs de tiges sont les insectes économiquement nuisibles du maïs. Ils peuvent parfois entraîner une destruction totale de cette culture céréalière. Le taux d'infestation de ces ravageurs de tiges varie de 30 à 70 % dans les champs des agriculteurs à faibles ressources alors qu'il est moins de 30 % dans les fermes commerciales où les insecticides sont utilisés comme le moyen de la lutte chimique (Sithole 1978b).

Le maïs (*Zea mays* L.), plante introduite en Afrique, et le sorgho cultivé (*Sorghum bicolor* L.) sont les céréales les plus largement cultivées en Afrique de l'Est (Polaszek 1998). Le rendement de la production de ces cultures chez les paysans aux ressources financières limitées est faible, en moyenne de 1 à 1,5 tonnes par hectare (ICIPE 2004). Ce faible rendement agricole céréalier est lié en grande partie aux contraintes causées par les insectes ravageurs de cultures parmi lesquels figurent en bonne place les Lépidoptères foreurs de tiges. Ces derniers comportent trois familles (Noctuidae, Crambidae et Pyralidae) qui sont les plus nuisibles économiquement (Polaszek 1998). Au Kenya où des études écologiques sont conduites sur le complexe des foreurs de tiges depuis plusieurs décennies, la perte annuelle de la production céréalière est de l'ordre de 5 à 73 % de la production totale (Seshu Reddy & Walker 1990 ; Mulaa 1995 ; Songa *et al.* 2002). Dans le cas particulier de la Coast Province (Kenya), une étude similaire indique que 18 % des pertes de la production de maïs sont dues à *C. partellus* et *C. orichalcociellus* (Warui & Kuria 1983). Jepson (1954) évalue à 40 à 100 % le pourcentage des plantes de sorgho infestées par *B. fusca* sur quelques localités tanzaniennes alors qu'en Ethiopie les dégâts causés par les chenilles de ce même ravageur ne sont que de l'ordre de 15 %. En Ouganda, la perte de la production de sorgho attribuée à l'activité de *C. partellus* est moyenne de 56 % (Starks 1969) tandis qu'au Kenya cette perte céréalière peut varier de 2 à 88 % (Kfir *et al*. 2002).

Face à cette importance économique des foreurs, des méthodes de lutte ont été proposées et utilisées contre ces ravageurs. On citera : l'utilisation des phéromones sexuelles synthétiques et des insecticides, l'association de plusieurs espèces cultivées sur une même parcelle, la gestion du milieu et des résidus des plantes cultivées, les dates de semis différées

dans le temps avec des densités variables, l'utilisation des variétés résistantes aux attaques des insectes phytophages, les lâchers de parasitoïdes (Kfir *et al*. 2002).

La génétique des populations apparaît de nos jours comme la discipline permettant de faire le lien entre les données écologiques et la diversité génétique existant entre les populations d'une espèce. Selon Avise (2000), la phylogéographie est définie comme la science qui étudie les principes et processus qui gouvernent la distribution géographique des lignées généalogiques, et en particulier, ceux mis en évidence au sein d'une espèce ou entre espèces étroitement apparentées.

Les espèces polyphages ou à large distribution géographique ne semblent pas avoir les mêmes traits de vie (biologie, écologie, incidences sur les pertes de récoltes) selon les plantes, les régions. Cela suggère que les populations écologiquement différenciées de telles espèces phytophages sont aussi génétiquement différenciées, une situation qui peut conduire à la formation des races d'hôtes voire éventuellement d'espèces, souvent cryptiques. On va donc chercher avec des marqueurs moléculaires neutres à mettre en évidence ces différences génétiques intra et interpopulationnelles. Mieux encore, on va essayer de retracer l'histoire de ces populations via une étude phylogéographique pour déterminer si la différenciation éventuelle d'une espèce de ravageur de cultures agricoles a quelque chose à voir avec l'histoire de la domestication de ses plantes hôtes cultivées. Si on parvient à montrer avec les analyses moléculaires qu'il existe des populations génétiquement et géographiquement différenciées, l'objectif serait alors de mieux adapter les méthodes de lutte chimiques et biologiques et de surveillance à la réalité de l'écologie et de la génétique des populations ainsi mises en évidence. Ainsi, lorsque les analyses moléculaires indiquent qu'on est face des populations différenciées d'une espèce de phytophage, on aura besoin d'adapter ces méthodes de lutte en fonction des régions ou des pays : le choix des parasitoïdes, par exemple, va dépendre de la nature de la population cible. Par contre, lorsque ces analyses montrent que les populations géographiques sont génétiquement homogènes, on ne tiendra pas compte des réalités géographiques. Dans tous les cas, la réalité écologique de chaque zone géographique interviendra pour permettre de mieux élaborer les stratégies de lutte.

Busseola fusca (Lepidoptera, Noctuidae), une espèce présente dans toute l'Afrique Subsaharienne, constitue un intéressant modèle d'étude au sein des Lépidoptères ravageurs de cultures de céréales sur ce continent. Nos travaux sur ce modèle de ravageurs de cultures sont présentés dans ce mémoire qui comporte cinq chapitres. Le **chapitre 1** fait un bilan synthétique des connaissances relatives à utilisation des outils moléculaires en entomologie appliquée. Dans cette partie, l'accent est beaucoup mis sur les espèces d'insectes

phytophages. Le **chapitre 2** présente le modèle *Busseola fusca*. Il précise sa biologie et son écologie tout en soulignant les nuisances économiques liées à cette dernière. Il introduit la problématique générale de notre étude sur ce ravageur. Dans le **chapitre 3**, nous présentons où et comment ce matériel biologique a été échantillonné et les différentes méthodes utilisées pour étudier la génétique des populations et la phylogéographie de ce foreur de tiges de céréales. Les résultats de l'ensemble des données moléculaires sont présentés dans le **chapitre 4** à travers les différents articles parus ou acceptés dans les revues scientifiques. Ce chapitre comporte aussi une partie qui présente les résultats des analyses complémentaires qui ne sont pas intégrées dans les articles scientifiques. Le **chapitre 5** s'articule autour de la discussion générale qui porte sur l'ensemble des données moléculaires obtenues sur cet insecte. Il se terminera avec les perspectives pour les futures recherches sur *B. fusca*.

Chapitre 1
Les outils moléculaires en entomologie appliquée

Section I : Les principaux types de marqueurs moléculaires

1.1. Historique

La connaissance de la nature de l'information génétique (ADN pour les organismes complexes et ARN pour certains virus) a eu un impact majeur sur les études en écologie et en évolution (Avise 1996). L'usage de plusieurs types de marqueurs moléculaires, associés à des approches méthodologiques de plus en plus élaborées, a permis de rendre plus accessible cette information génétique au sein des différents organismes. Ainsi, les marqueurs protéiques et nucléotidiques constituent des outils moléculaires puissants qui ont facilité les investigations en phylogénie, en évolution, en écologie et en dynamique des populations (Loxdale & den Hollander 1989 ; Smith & Wayne 1996 ; Symondson & Liddell 1996). Dans le cas de l'entomologie, la science de l'étude des insectes, le champ d'utilisation des outils de la biologie moléculaire s'est nettement élargi ces dix dernières années et nombreux sont les entomologistes qui utilisent actuellement ces techniques. Elles ont conduit à un nouveau degré de résolution dans l'étude de la diversité écologique et taxonomique des insectes (Loxdale & Lushai 1998). Cette partie fera le point sur l'évolution historique de ces marqueurs moléculaires et de leurs applications tout en insistant plus particulièrement sur le cas des insectes phytophages.

1.2. Les enzymes protéiques

L'électrophorèse des protéines, c'est-à-dire la migration des protéines sous l'influence d'un champ électrique, est l'une des méthodes les moins coûteuses pour l'obtention de données génétiques au niveau moléculaire. Elle a été aussi historiquement la première technique moléculaire utilisable. Depuis l'apparition du gel d'amidon (Smithies 1956) avec la visualisation histochimique des enzymes sur gels (Hunter & Markert 1957) suivies des études modèles de Harris (1966), une révolution majeure dans la compréhension des processus micro et macroévolutifs a été possible. L'électrophorèse sur gel d'amidon ou d'acrylamide, dont l'usage s'est répandu à la fin des années soixante, a été pour bon nombre de thématique de recherche la méthode d'analyse du polymorphisme génétique dans les populations et les espèces (Lewontin & Hubby 1966). Le principe en est le suivant : on prélève chez chaque individu une petite quantité d'un tissu particulier que l'on broie. La goutte d'extrait protéique obtenue est déposée sur une plaque de gel imbibée d'une solution

d'électrolytes. On établit ensuite un champ électrique. Chaque protéine se déplace dans ce champ avec une vitesse qui est principalement fonction du rapport entre sa charge et sa taille moléculaires et prend, après un temps défini de migration, une position caractéristique par rapport au front de migration. A la fin de la migration électrophorétique, le gel est mis en contact avec un produit pouvant servir de substrat à une réaction enzymatique : il donne un produit coloré. On observera une bande à l'endroit où se sont arrêtées les molécules de l'enzyme correspondant. Une bande est observée pour les individus homozygotes et plusieurs bandes pour les individus hétérozygotes. On peut alors identifier les isozymes, enzymes ayant des fonctions cellulaires voisines, puisqu'elles sont actives sur le même substrat artificiel. Cette technique a permis d'estimer la diversité génétique dans les populations animales (Pasteur & Stordeur 1976) et végétales.

Les différents allèles d'un locus donné sont issus des substitutions nucléotidiques ayant conduit au remplacement d'acides aminés (Avise 1996). Les alloenzymes montrent des patrons de bandes protéiques variés révélateurs, chez les individus hétérozygotes, de la structure de la forme active de l'enzyme : monomérique, dimérique, trimérique et tétramérique (figure 1-1).

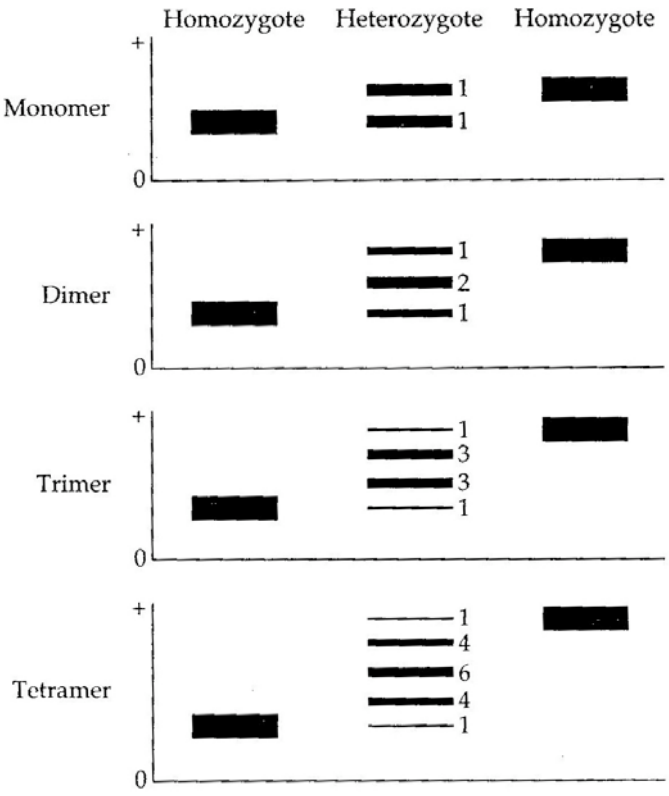


Figure 1-1 : Diagramme de patrons isozymiques attendus chez les homozygotes et les hétérozygotes pour les enzymes protéiques composées d'une ou de plusieurs sous unités. Les ratios de l'intensité enzymatique chez les hétérozygotes sont indiqués. (d'après Hillis *et al*. 1996).

Bien que de nombreux marqueurs moléculaires soient devenus accessibles ultérieurement après la découverte des endonucléases ou enzymes de restriction et le

séquençage direct du matériel génétique, les alloenzymes restent encore d'actualité et sont parfois utilisés soit seuls, soit en association avec d'autres marqueurs (nucléotidiques ou morphologiques) dans de nombreuses études chez les insectes. De tels exemples se retrouvent en taxonomie (Pinto *et al*. 1992) ; en génétique des populations et en écologie (Loxdale *et al*. 1985 ; Strong & Mahon 1991 ; Wainhouse & Jukes 1997 ; Jeffrey *et al*. 1999 ; Thomas *et al*. 2003 ; Schwarz *et al*. 2003 ; Al-Barrak *et al*. 2004), avec des problématiques variées, liées à l'étude des effets fondateurs chez les drosophiles hawaïennes (DeSalle & Hunt 1987), à l'analyse de spéciation et de la formation de races chez *Rhagoletis* (Feder *et al*. 1989), à la résistance aux insecticides des insectes ravageurs (conférée par les estérases) (Maa & Terriere 1983), à la différenciation des chenilles d'espèces différentes de noctuelles ravageuses (cas de *Helicoverpa* spp. ; Fisk & Daly 1989) et à la cartographie chromosomique (Loukas *et al*. 1979). L'étude des alloenzymes a permis d'estimer la proportion moyenne des gènes polymorphes chez les insectes. Cette méthode a révélé que la fraction des gènes polymorphes est plus importante, de l'ordre de 45%, chez les drosophiles que chez d'autres groupes insectes. Cette analyse a montré qu'entre les espèces de *Drosophila* la proportion de gènes polymorphes varie largement entre espèces (30 à 56 %) (Hartl 1997). Caterino *et al*. (2000) ont fait une synthèse sur des séquences utilisées dans les différents ordres d'insectes à des fins de systématique moléculaire (Tableau 1-1).

Cependant, les marqueurs isozymiques présentent des limites liées à leur pouvoir de résolution. La variabilité génétique observée n'est que l'expression phénotypique des enzymes étudiées. En effet, le degré de polymorphisme tend à être sous-estimé du fait que l'électrophorèse ne permet pas de révéler toutes les substitutions d'acides aminés ou de détecter les allèles nuls. La substitution des acides aminés ne conduit pas nécessairement à un changement de point isoélectrique. Si la charge de l'acide aminé substitué ne change pas, la substitution ne sera pas non plus détectée par électrophorèse isoenzymatique : la charge globale de la chaîne polypeptidique mutée restera alors identique à celle de la chaîne initiale. Il en résulte que le polymorphisme allélique est sous-estimé. La migration des allèles multiples à un locus donné est considérée comme un trait qualitatif à état multiple qui ne permet pas d'établir leurs relations phylogénétiques par la difficulté de mettre en évidence les allèles rares : les enzymes révélées sont généralement soumises à la sélection naturelle. Dans ce cas, seuls les allèles les plus fréquents sont facilement identifiés à un locus, les autres, plus rares, seront difficilement révélés du fait de leur fréquence très faible. Richardson *et al*. (1986) ont estimé qu'un sixième de la variabilité moléculaire est révélée à l'aide des protéines enzymatiques.

Tableau 1-1 : Protéines codées par les locus nucléaires qui ont été utilisées dans les études de systématique moléculaire chez insectes. Les taxons sont listés en ordre de nombre décroissant des séquences existantes : Di = Diptera ; Le = Lepidoptera ; Tr = Trichoptera ; Me = Mecoptera, Si = Siphonaptera ; Hy = Hymenoptera ; He = Hemiptera ; Co = Coleoptera ; Or = Orthoptera ; Th = Thysanura ; Cm = collembola ; Ar = Archaeognatha ; Bl = Blattodea (d'après Caterino *et al*. 2000 ; modifié).

Locus	Séquences	Taxons
α-amylase	38	Di, Co
Acetylcholine esterase (*achE*)	21	Di
Actin	5	Hy, Le
Alchol desdrogenase (*adh*)	157	Di
Arylphorin (hexamerin family)	12	Le
Cecropin family	36	Di
Chorion genes : *s18, s15 and s19*	5	Di
Cpn1-1 (anonymous nuclear DNA)	8	Or
Dopa decarboxylase (*ddc*)	57	Le
Elongation factor-1-alpha (*EF-1α*)	196	Le, Hy, He, Cm, Ar, Bl
EF-1α (copy 2)	1	Hy
Esterase (*est*)	31	Di
Glycerol-3-phosphate dehydrogenase (gpdh or g3pdh)	39	Di
Glycerol-6-phosphate dehydrogenase (g6pdh)	13	Di, Le, Th
Guanylate cyclase	13	Di
Globin family	33	Di
Histone H1	8	Di
Histone H4	7	Hy
Hunchback (*hb*)	39	Di
Kruppel	3	Di
Luciferase	2	Co
Lysozyme intron	16	Le
Myosin alkali light chain (*M1c1*) intron	30	Di
Nullo	7	Di
Opsin	28	Hy, Di
Period (*per*)	111	Di, Le
Phosphoglucose isomerase (*Pgi*)	2	Or
Phosphoenolpyruvate carboxykinase (*PEPCK*)	12	Le, Tr, Di, Me, Si
Prune	10	Di
Resistance to dieldrin (*Rdl*)	17	Co
Retrotransposon reverse transcriptase	30	Hy
Ribosomal protein 49 (*rp49*)	9	Di
Serendiptiy-α	7	Di
Cu, Zn-superoxide dismutase (*sod*)	30	Di
Sodium channel para locus 1	27	Le, He, Bl, Co
Snail	3	Di
Timeless	1	Di
Triosephosphate isomerase (*Tpi*)	39	Di
Vestigial (*vg*)	5	Di
White	15	Di
Wingless (*wg*)	57	Le, Di
Xanthine dehydrogenase (*xdh*)	37	Di
Yolk protein 1 (*yp1*)	55	Di
Yolk protein 2 (*yp2*)	40	Di
Zeste	24	Di

1.3. Les marqueurs moléculaires nucléotidiques

Les marqueurs moléculaires nucléotidiques peuvent être des fragments de taille ou des séquences nucléotidiques. Dans les deux cas, il existe chez tous les animaux deux types principaux de marqueurs selon leur origine cellulaire : les marqueurs mitochondriaux et les marqueurs nucléaires. Chacun de ces compartiments génomiques comporte des gènes avec leurs fragments codants (exons) et non codants (introns) ; les pseudogènes et les régions intergéniques dont l'importance varie en fonction de leur localisation cellulaire : le non codant domine dans le génome nucléaire alors qu'il est très réduit dans le génome cytoplasmique.

1.3.1. Les fragments d'ADN

La source fondamentale de la variabilité génétique se situe au niveau de l'ADN (Hartl 1997). La mise au point des techniques d'hybridation moléculaire (*DNA-DNA hybridization*), d'amplification et de digestion moléculaire des fragments d'ADN et de séquençage (dans les années 80) a ouvert la voie à l'exploitation des marqueurs nucléotidiques. On a alors assisté à une explosion des approches méthodologiques et analytiques qui permettent de mettre en évidence le polymorphisme nucléotidique qui existe au sein des différents génomes (nucléaire, mitochondrial et chloroplastique) et au sein des populations constituant une espèce.

Hybridation d'ADN

Les premières techniques nucléotidiques sont basées sur l'hybridation des molécules d'ADN (*DNA-DNA hybridization*). Celle-ci était utilisée pour établir les relations évolutives entre les organismes vivants. Son avantage est lié au fait que les forces qui relient les deux brins de la molécule d'ADN dépendent de leur séquence nucléotidique. L'adénine et la thymine reliées par deux liaisons hydrogènes nécessitent moins d'énergie pour se dénaturer que la guanine et la cytosine reliées par trois liaisons hydrogènes. Les températures de dénaturation dépendent donc du pourcentage en GC. Après un chauffage à 100°C de la molécule d'ADN, ces liaisons hydrogènes sont rompues et les deux brins complémentaires sont séparés. Un refroidissement doux de cette solution conduit à la renaturation des brins séparés. Les molécules d'ADN provenant de deux différentes espèces peuvent être combinées, dénaturées et réassociées ou hybridées. Les molécules doubles brins d'ADN qui se forment après hybridation des brins complémentaires des deux espèces en présence contiennent la disparité de distribution (mismatch distribution) des paires de bases qui est

caractérisée par leur divergence évolutive à partir d'un ancêtre commun. Le polymorphisme obtenu résulte simplement des différences des températures observées (dues au pourcentage en CG) au sein des fragments d'ADN formés après leur hybridation en double brin pour des séquences provenant de deux ou plusieurs espèces (figure 1-2). L'évaluation de ces différences de températures (T_m à 50%) de stabilité entre les molécules homo- et hétéroduplex fournit une estimation quantitative sur la divergence génétique entre les espèces en présence. Les approches d'hybridation *DNA-DNA* ont eu un effet appréciable sur la génétique moléculaire en révélant de nombreux aspects de l'organisation et de la structure génomique des organismes vivants, par exemple, l'importance des séquences répétées d'ADN, leur longueur et leur distribution avec les copies de séquences (Avise 1994).

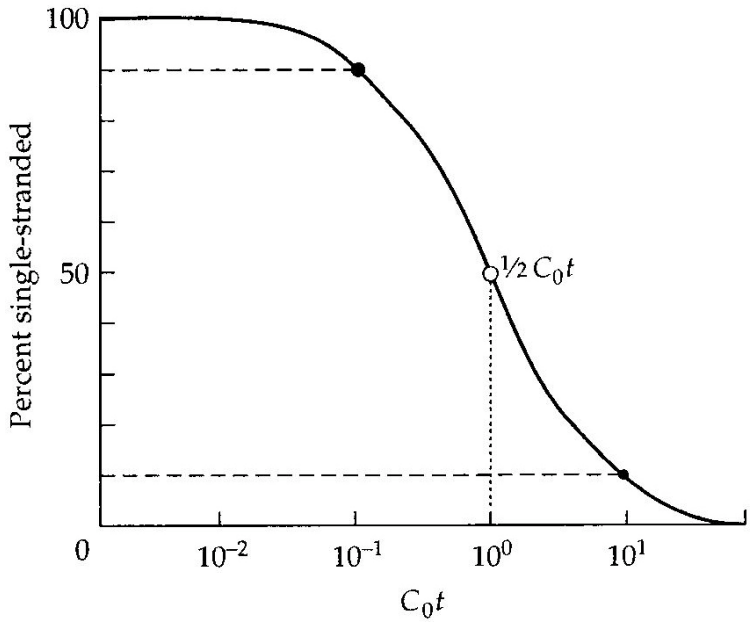


Figure 1-2 : Courbe de réassociation idéale (courbe C_ot) pour une unique classe de l'ADN (c'est-à-dire copie unique ou fréquence unique de la classe répétitive de l'ADN). La courbe représente la perte de brins uniques de l'ADN (déterminée par la formule $1/1 + kC_ot$) et la formation des duplex sur les intervalles log de C_ot exprimée en [moles de nucléotides / litre] x seconde. Le démi-C_ot est la valeur C_ot (dans ce cas = 1) pour laquelle 50% de l'ADN sont réassociés. Dans une réaction idéale, 80% de l'ADN se réassocient sur 2 intervalles de log de sorte que la valeur de C_ot pour laquelle 90% de l'ADN sont réassociés est égale à 10 fois le démi-C_ot (d'après Hillis *et al*. 1996).

RFLP (Restricted Fragment Lengh polymorphism), AFLP (Amplified Fragment Lengh Polymorphism), RAPD (Ramdom Amplified Polymorphic DNA)

La découverte des enzymes de restriction ou endonucléases (Meselson & Yuan 1968 ; Arber & Linn 1969) a donné une impulsion remarquable à la biologie moléculaire. Elles sont produites par des bactéries. Il s'agit des enzymes de type II qui coupent la molécule

double brin d'ADN en des sites spécifiques. Ces endonucléases reconnaissent une très courte séquence d'ADN de 4 à 6 nucléotides, généralement, double brin et le plus souvent palindromique (identique dans les deux sens de lecture). Les fragments nucléotidiques obtenus après la coupure par une enzyme de restriction sont généralement visualisés sur gel d'agarose au moyen de bromure d'éthidium. Plusieurs centaines d'enzymes de restriction ont été identifiées et caractérisées (Roberts 1984). Le tableau 1-2 présente les caractéristiques de quelques-unes. Ces endonucléases de restriction ont ouvert la voie à la mise au point des marqueurs moléculaires tels que les RFLP (Polymorphisme de longueur des fragments de restriction) et les AFLP (Polymorphisme de longueur des fragments amplifiés) qui ont été utilisés dans les études récentes des populations d'insectes.

Pour les RFLP, AFLP et RAPD, la très faible quantité d'ADN nécessaire à la réaction de PCR rend possible l'analyse individuelle d'insectes de très petite taille voire d'un fragment d'insecte ou un stade immature (Harry *et al*. 1998).

Le principe des RFLP est basé sur le polymorphisme de taille de fragments d'ADN lié aux mutations moléculaires (substitutions nucléotidiques, insertion ou délétion de nucléotides, réarrangement de séquences nucléotidiques) produites aux sites de restriction enzymatique entre les différents individus d'une espèce. Ce marqueur moléculaire présente l'avantage majeur de permettre d'accéder facilement à certaines substitutions nucléotidiques qui, souvent, représentent une variation moléculaire élevée. Le nombre d'allèles va changer en fonction des mutations nucléotidiques survenues sur les fragments géniques ou intergéniques. L'utilisation de ce marqueur moléculaire a permis d'établir des cartographies du génome (carte chromosomique chez les papillons ravageurs (*Heliothis*, Zraket *et al*.1990)), d'étudier la génétique des populations des mouches de fruit (Haymer *et al*. 1992) de la noctuelle légionnaire d'automne, *Spodoptera frugiperda* (Lu *et al.* 1992), du complexe d'espèces *Thaumetopoea pityocampa-wilkinsoni* (Lépidoptères) (Salvato *et al*. 2002) et de faire des analyses parentales chez des insectes élevés à des fins commerciales tels que les abeilles domestiques (Hall 1990). Les analyses ont surtout porté sur les petits génomes (mitochondries, chloroplastes, plasmides) pour caractériser et évaluer la diversité biologique des espèces (Caterino *et al*. 2000).

Les AFLP constituent une variante des RFLP. Ils ont été mis au point par Vos *et al*. (1995). Le principe de base repose sur l'amplification sélective de fragments d'ADN générés par digestion à l'aide d'une combinaison de deux enzymes de restriction, par exemple, *Sau*3A et *Eco*RI. Cette amplification sélective est rendue possible par la ligation d'adaptateurs (séquences oligonucléotidiques d'ADN connues et ayant l'extrémité complémentaire à la

séquence générée par la digestion de restriction) aux fragments digérés. La variation génétique des AFLP peut être exprimée de plusieurs façons. Il s'agit d'un marqueur moléculaire dominant car il est impossible de différencier les individus homozygotes des hétérozygotes par la présence ou absence des bandes (Carlson *et al*. 1991). Les bandes visualisées sur le gel sont de poids moléculaires différents et sont considérées comme étant des locus différents. Ces marqueurs restent très peu utilisés en entomologie bien qu'ils aient récemment permis d'établir la carte de liaison dans l'étude de papillons ravageurs (Heckel *et al*. 1998). Ils sont utilisés dans certaines études actuelles pour déterminer la structure génétique des populations et de patron biogéographique des insectes (Carisio *et al*. 2004 ; Timmermans *et al*. 2005, Prowell *et al*. 2004) ou pour établir la corrélation entre la structure génétique d'un insecte phytophage avec la distribution de ses plantes hôtes (Sword *et al*. 2005).

Enfin, l'identification et la caractérisation d'amorces universelles nucléaires et mitochondriales (Simon *et al*. 1994) ont permis d'inférer les informations portées par l'ADN en se servant de la méthode RAPD (amplification aléatoire du polymorphisme de l'ADN). Les amorces courtes ou peu spécifiques se lient aléatoirement à des séquences homologues se trouvant dans le génome. Dans la méthode RAPD, contrairement à une PCR "classique", une seule amorce non spécifique (et non deux spécifiques) est utilisée. Cette amorce présente une séquence nucléotidique aléatoire de taille standardisée, généralement, à 10 nucléotides. En raison de sa petite taille, l'amorce a une grande probabilité de s'hybrider à des sites complémentaires proches les uns des autres et en orientations inversées sur l'ADN matriciel. Pour un génome donné, il peut y avoir un certain nombre de sites d'hybridation pour cette amorce. A l'issue de la RAPD, un profil multi locus est obtenu. Les produits d'amplification, visualisés sur gel d'agarose en présence de bromure d'éthidium, varient en longueur et par la nature de leur séquence. Au sein d'une population, les mutations influencent les sites de fixation des amorces. Cela permet de mettre en évidence le haut niveau de polymorphisme des RAPD existant au sein de cette population. C'est une technique qui ne nécessite pas une connaissance préalable du génome étudié et qui repose sur l'emploi de basses températures d'hybridation (de 35 à 40°C).

L'atout majeur des marqueurs RAPD est la rapidité avec laquelle il est possible de les révéler en grand nombre. L'ADN ciblé par la RAPD est essentiellement l'ADN nucléaire et notamment des régions répétitives (répétées). L'application de cette méthode a conduit à caractériser la variabilité génétique des populations ou des espèces d'insectes (Baumann *et al*. 2003 ; Al-Barrak *et al*. 2004). Ces marqueurs sont utilisés pour la diagnose d'espèces affines

ou cryptiques d'insectes d'importance agronomique ou médicale, ou pour l'étude des systèmes de reproduction d'espèces, de populations ou de clones d'insectes spécialistes ou présentant des cycles de reproduction complexes (Harry *et al*. 1998). Ils permettent d'identifier la provenance de ravageurs immigrants ; ils sont très utiles dans les études de recherche de paternité ou de parenté. L'étude de la diversité génétique des populations ibériques de *Bemisia tabaci* à l'aide de la technique RAPD a montré que le flux génique est faible entre les différents groupes coexistant au sein d'une même population géographique par rapport à celui observé entre populations géographiques d'un même groupe (Moya *et al*. 2001). Ce marqueur a été utilisé pour étudier la structure génétique chez d'autres espèces d'insectes phytophages. On peut citer le cas des populations de charançons (*Anthonomus grandis*) ravageurs du cotonnier aux USA (Kim & Sappington 2004), de populations de *Diprion pini* (Baumann *et al*. 2003) et de la noctuelle des tomates (*Helicoverpa armigera*) (Zhou *et al*. 2000). Des études similaires conduites avec ce même marqueur moléculaire a permis d'établir la carte génétique de *Tribolium confusum* (Yezerski *et al*. 2003), à définir les profils caractéristiques des espèces avec des fragments diagnostiques dans le cas du complexe d'espèces *Culicoides imicola* (Diptera, Ceratopogonidae) (Sebastiani *et al*. 2001).

Toutefois, elle présente l'inconvénient d'être peu reproductible du fait de variations stochastiques de l'efficacité des cycles d'amplifications liées à la petite taille et à la spécificité imparfaite des amorces (Harry *et al*. 1998).

De façon générale, comme pour les allozymes, il est impossible d'établir les relations phylogénétiques qui existent entre les différents allèles visualisés à l'aide de ces marqueurs moléculaires (RFLP, AFLP, RAPD).

Tableau 1-2 : Propriétés de plusieurs endonucléases couramment utilisées (*b, c, d* représentent le nombre de fragments obtenus après la digestion enzymatique spécifique du génome chloroplastique pour le tabac et mitochondrial respectivement pour l'homme et l'abeille) (d'après Hillis *et al*. 1996).

		Number of sites		
Enzyme	Site	Tobacco cpDNA[b]	Human mtDNA[c]	Bee mtDNA[d]
*Alu*I	AG↓CT	341	64	19
*Apa*I	GGGCC↓C	13	5	0
*Apa*LI	G↓TGCAC	8	1	0
*Ase*I	AT↓TAAT	114	10	147
*Ava*I	C↓YCGUG	70	3	1
*Bam*HI	G↓GATCC	40	1	0
*Ban*I	G↓GYUCC	32	8	0
*Ban*II	GUGCY↓C	75	15	0
*Bcl*I	T↓GATCA	54	4	6
*Bgl*I	GCN$_5$↓GGC	9	2	0
*Bgl*II	A↓GATCT	60	0	2
*Bst*BI	TT↓CGAA	94	7	0
*Bst*EII	G↓GTNACC	12	2	0
*Bst*NI	CC↓(A/T)GG	128	4	5
*Bst*UI	CG↓CG	89	6	0
*Bst*XI	CCAN$_6$↓TGG	26	4	0
*Bsu*36I	CCT↓NAGG	15	1	2
*Cla*I	AT↓CGAT	59	1	4
*Dde*I	C↓TNAG	309	72	4
*Dra*I	TTT↓AAA	64	4	103
*Eco*0109	UG↓GNCCY	67	14	1
*Eco*NI	CCT↓N$_8$AGG	18	5	1
*Eco*RI	G↓AATTC	97	3	5
*Eco*RV	GAT↓ATC	36	3	1
*Hae*II	UGCGC↓Y	25	7	0
*Hae*III	GG↓CC	196	50	0
*Hin*cII	GTY↓UAC	57	12	1
*Hin*dIII	A↓AGCTT	33	3	2
*Hin*fI	G↓ANTC	718	36	12
*Hin*PI	G↓CGC	89	17	1
*Hph*I	GGTGAN$_8$↓	128	55	8
*Kpn*I	GGTAC↓C	15	3	0
*Mbo*I	↓GATC	623	22	35
*Mlu*I	A↓CGCGT	9	0	0
*Msp*I	C↓CGG	214	23	1
*Nci*I	CC↓(C/G)GG	113	8	0
*Nco*I	C↓CATGG	NA	4	0
*Nde*I	CA↓TATG	NA	3	4
*Nhe*I	G↓CTAGC	7	1	0
*Nsi*I	ATGCA↓T	43	3	2
*Pst*I	CTGCA↓G	14	2	1

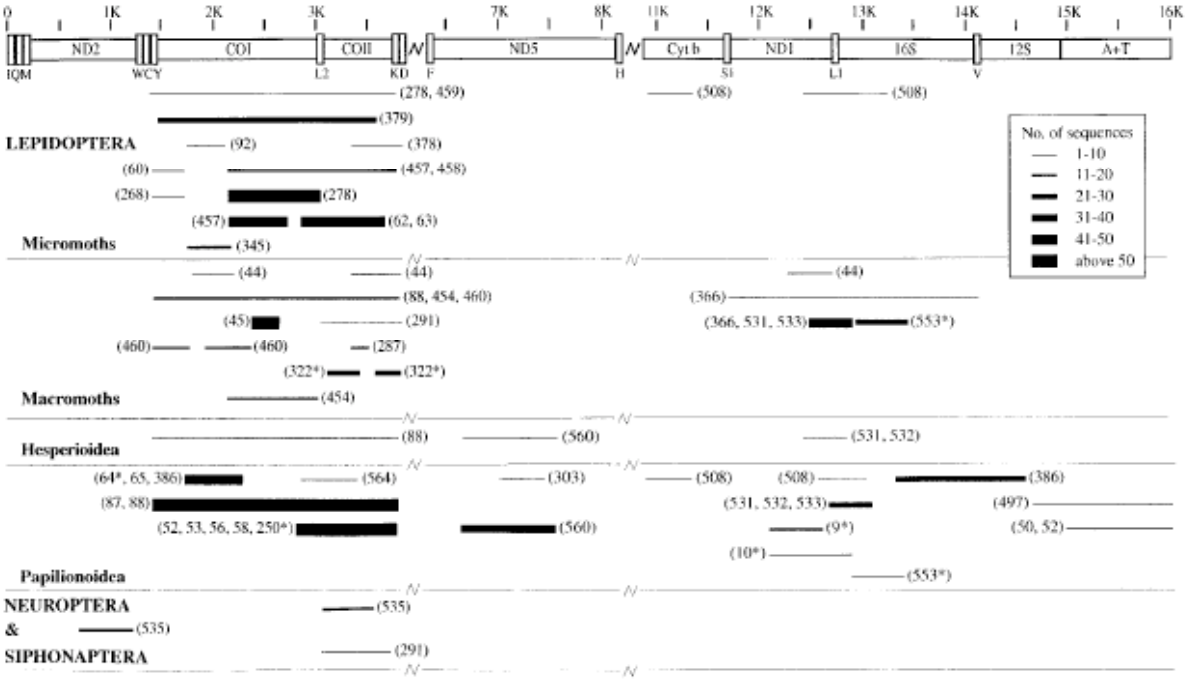
1.3.2. Les séquences nucléotidiques

Le séquençage de l'ADN a permis la découverte et l'utilisation de nombreuses régions génomiques ayant des caractéristiques particulières et conduisant aux inférences sur les mécanismes de l'évolution des espèces. Les marqueurs nucléotidiques nucléaires (ITS1, ITS2, 18S, 28S, microsatellites,…) et mitochondriaux (COI, COII, COIII, cytochrome *b*…) sont utilisés dans le règne animal pour des études phylogénétiques, démographiques, biogéographiques, phylogéographiques. Ces analyses visent avant tout à reconstituer l'histoire évolutive des organismes étudiés et à établir les liens de parenté entre eux. Il s'agit d'une nouvelle révolution de la biologie moléculaire dont le champ d'application recouvre plusieurs disciplines de biologie fondamentale et appliquée.

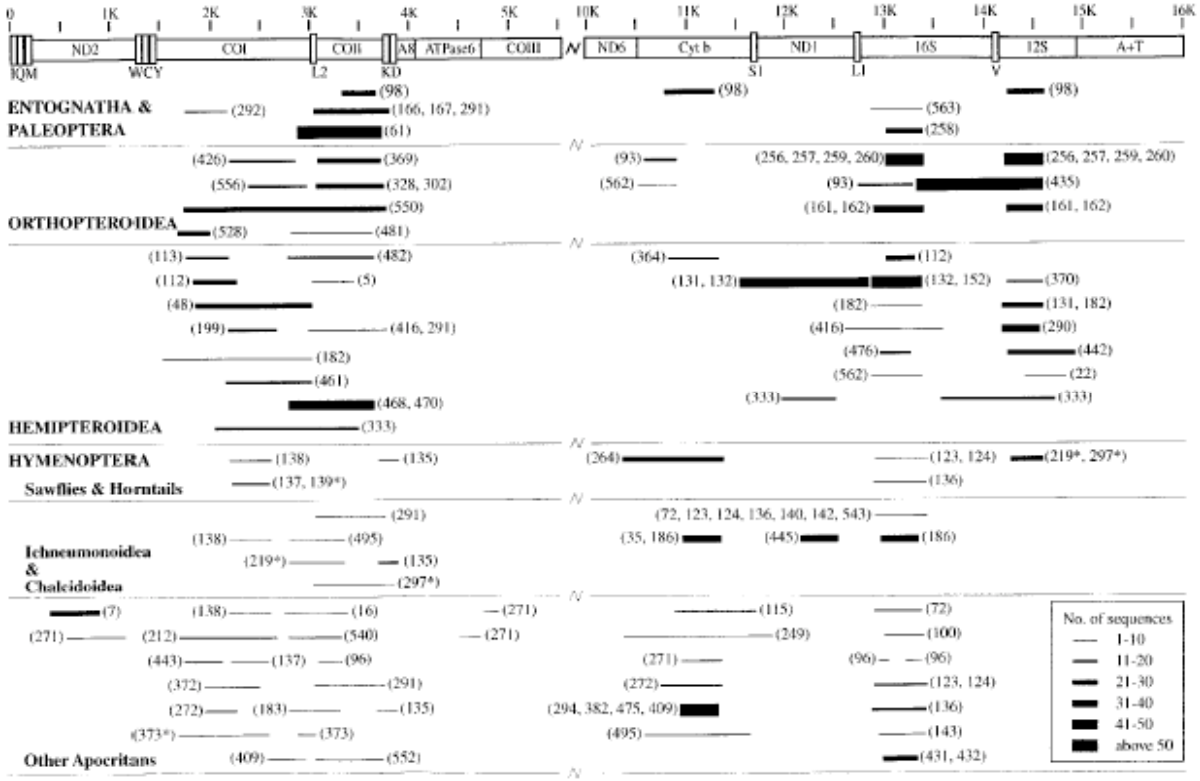
Il faut cependant rappeler que la plupart des marqueurs moléculaires sont apparus à la suite de la mise au point de la technique d'amplification (PCR) de l'ADN.

L'ADN mitochondrial

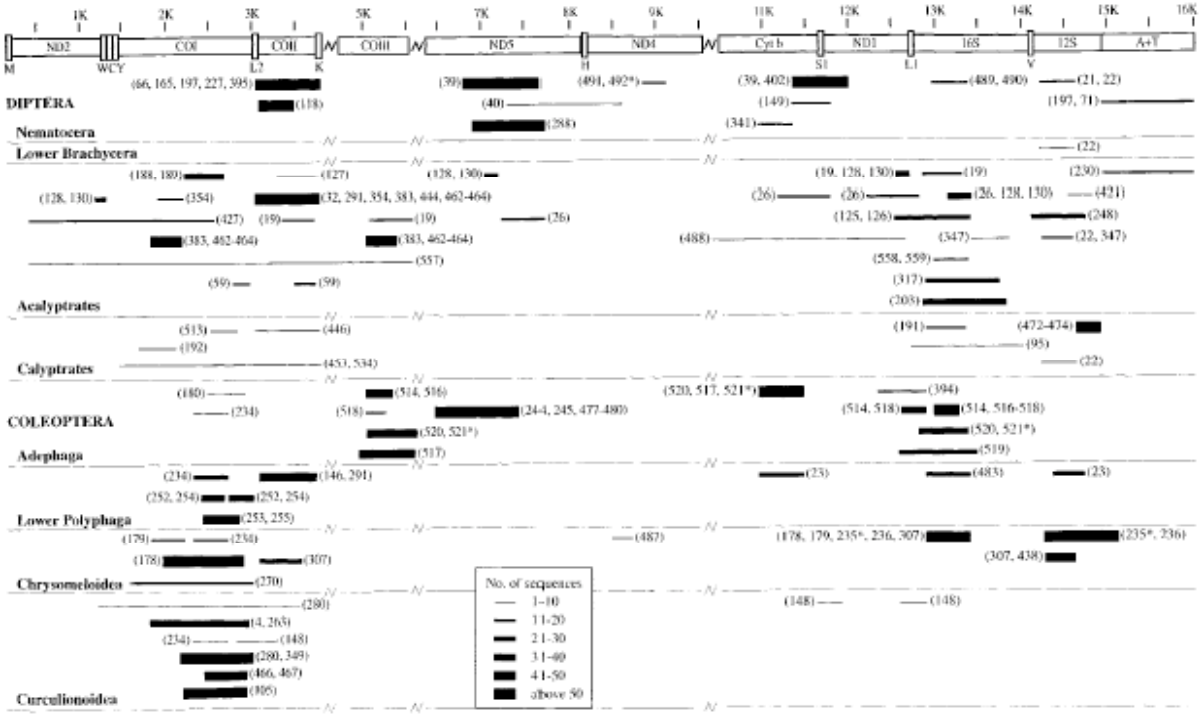
Les marqueurs mitochondriaux sont de loin les marqueurs nucléotidiques les plus utilisés chez les insectes comme chez les animaux en général (Hillis *et al*. 1996). Les figures 1-3a, 1-3b, 1-3c présentent l'utilisation relative des différentes régions nucléotidiques mitochondriales dans les différents ordres d'insectes (d'après Caterino *et al*. 2000).



1-3a



1-3b



1-3c

Figure 1-3 : Régions mitochondriales séquencées pour les études de systématique moléculaire chez les hexapodes. (1-3a) Cette figure traite de toutes les études publiées avant 15 mars 1999. Celles basées uniquement sur des données précédemment publiées dans les bases de données (comme GenBank) ne sont pas prise en compte. Une étude est considérée comme

relevant de la systématique si les données ont servi à construire un arbre phylogénétique. Pour une région donnée, la longueur de chaque barre indique la taille de la région séquencée alors que l'épaisseur de cette barre donne une idée sur le nombre de séquences publiées pour ce taxon. (1-3b) Dans les régions à longueur variable (la région de contrôle, ADN ribosomique) la longueur des barres figurées peut ne pas représenter exactement la longueur du fragment séquencé. (1-3c) En complément des études citées ici, des séquences mitochondriales complètes sont disponibles pour les six espèces d'insectes suivants : *Locusta migratoria* (164), *Drosophila yakuba* (103), *D. melanogaster* (289), *Anopheles gambiae* (31), *A. quadrimaculatus* (327) et *Apis mellifera* (114). Les régions illustrées par chaque figure ne sont pas identiques. *(a)* Régions mitochondriales séquencées chez les Lépidoptères, les Siphonaptères et les Neuroptères. *(b)* Régions mitochondriales séquencées chez les Holométaboles et les Hémimétaboles. *(c)* Régions mitochondriales séquencées chez les Diptères et les Coléoptères (d'après Caterino *et al*. 2000).

On constate que les différentes parties du génome mitochondrial sont utilisées pour les études de systématique moléculaire chez les insectes avec des fréquences variables. Les plus utilisés sont les gènes codant les sous-unités I et II de la cytochrome oxydase (COI, COII) et ceux codant des ARN ribosomiques (16S et 12S). D'autres n'ont fait l'objet que de quelques études, comme par exemple, les gènes codant la sous-unité III de la cytochrome oxydase (COIII), la NADH déshydrogénase et le Cytochrome *b*, la région de contrôle de l'ADN mitochondrial. D'autres gènes encore n'ont fait l'objet que d'analyses isolées : ND1, ND2, ND4. Enfin, le reste des gènes mitochondriaux a été le plus souvent ignoré.

L'ADN nucléaire

Les régions codantes du génome nucléaire comportent des gènes qui sont exprimés et dont les produits donnent des ARN ribosomiques ou des chaînes polypeptidiques. Beaucoup de gènes nucléaires sont en copies multiples rendant difficile leur utilisation dans les études moléculaires. La phylogénie d'un gène représente de façon certaine l'histoire généalogique particulière de ce dernier. Celle-ci peut différer entre les différents gènes de l'espèce étant donné qu'ils ne sont pas toujours génétiquement liés entre eux et que les effets d'autres forces évolutives, comme la dérive génétique et la sélection naturelle, seront différents d'un gène à un autre. Pour cette raison, plusieurs gènes non liés doivent être utilisés pour la reconstruction phylogénétique. De plus, la constatation que l'arbre de gènes mitochondriaux peut représenter seulement un volet, parfois, partiellement biaisé (par les contraintes évolutives ou par les méthodes des reconstructions phylogénétiques) de l'histoire phylogénétique des organismes a conduit les systématiciens molécularistes à s'orienter vers des marqueurs nucléaires pour compléter les données mitochondriales. Quelques gènes nucléaires ont été utilisés pour la

phylogénie des arthropodes (Friedlander *et al*. 1992 ; Friedlander *et al*. 1994 ; Friedlander *et al*. 1996 ; Brower & DeSalle 1994 ; Lin & Danforth 2004 ; Mallatt *et al*. 2004 ; Saux *et al*. 2004 ; Jost & Shaw 2006). La figure 1-4 donne un exemple de gènes nucléaires couramment utilisés en systématique moléculaire.

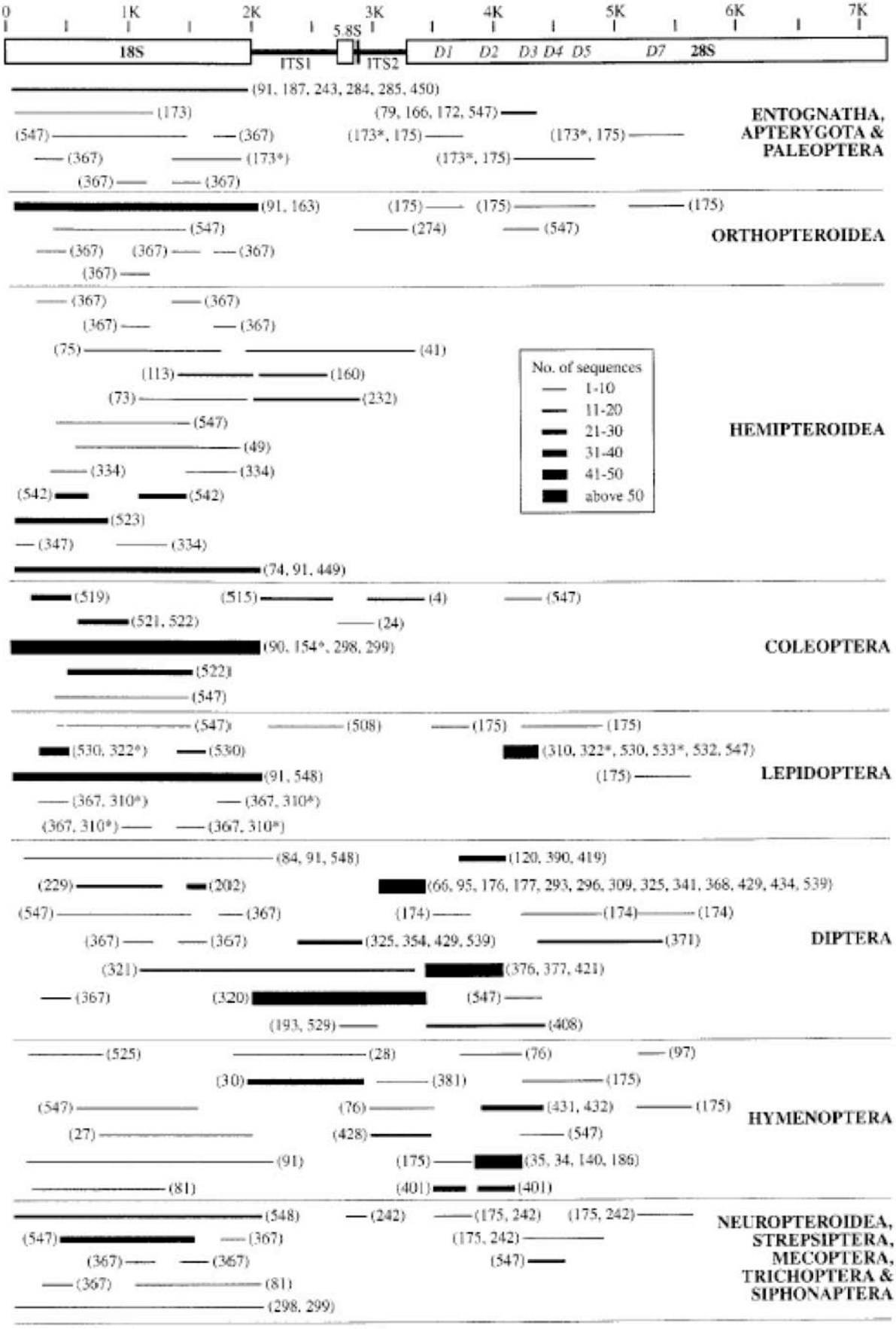


Figure 1-4 : Régions séquencées de l'ADN ribosomique pour les études de systématique moléculaire chez les hexapodes avant mars 1999 (d'après Caterino *et al*. 2000)

Comme dans le cas de l'ADN mitochondrial, seuls quelques gènes nucléaires sont couramment utilisés en systématique moléculaire les autres étant ignorés ou rarement utilisés. Les gènes nucléaires les plus fréquemment utilisés chez les arthropodes sont : l'ARN ribosomique 28S, l'ARN ribosomique 18S, le facteur d'élongation (*EF1-*α), le gène wingless (*Wg*). Dans certaines études isolées, quelques autres gènes nucléaires ont été étudiés : le gène de phosphoenolpyruvate carboxykinase (*PepCK*), le gène de dopa décarboxylase (*DDC*), le gène de tektin (*tektin*), le gène de l'histone H3 (*H3*). Les séquences de ces différents gènes nucléaires sont les plus utilisées en phylogénie (Wiegmann *et al*. 2000 ; Rokas *et al*. 2002 ; Mallatt *et al*. 2004). Dans ces études, plusieurs gènes nucléaires et mitochondriaux sont combinés et, parfois, associés aux caractères morphologiques afin d'examiner la contribution relative et la fiabilité de chaque gène dans la compréhension de l'histoire évolutive de l'espèce ou du groupe taxonomique concerné. Une telle analyse peut conduire à faire le choix de gènes différents pour résoudre les nœuds profonds et des divergences plus récentes. Ces études ont permis d'établir la phylogénie et l'histoire évolutive des espèces (Damgaard *et al*. 2001 ; Downie & Gullan 2004 ; Sota & Ishikawa 2004 ; Mallarino *et al*. 2005 ; Brower *et al*. 2006 ; Mitchell *et al*. 2006), de préciser la position systématique, la biogéographie et parfois la phylogéographie des espèces proches avec inférence de leur histoire passée (Arensburger *et al*. 2004 ; Balke *et al*. 2005 ; Zhang *et al*. 2005 ; Wahlberg *et al*. 2005 ; Horn *et al*. 2006).

On regroupe dans la catégorie des régions non codantes les introns, les pseudogènes et les espaces intergéniques, les séquences répétées et les ADN satellites, en particulier, les microsatellites sur lesquels nous insisterons chez les insectes. Les séquences introniques des gènes nucléaires sont parfois utilisées soit seules pour retracer l'histoire de l'invasion d'une espèce (cas de la mouche de fruits *Ceratitis capitata*, Villablanca *et al*. 1998) ou en association avec d'autres marqueurs pour préciser la systématique de certains genres (cas du genre *Ithomia*, Mallarino *et al*. 2005). Les pseudogènes sont très peu utilisés chez les insectes. Les espaceurs ITS1 et ITS2 sont des régions intergéniques couramment utilisés en génétique des populations (Szalanski *et al*. 1999 ; Gallego & Galian 2001), mais aussi, en phylogénie comme le cas des espèces du genre *Diabrotica* (Clark *et al*. 2001).

Les microsatellites sont reconnus comme des marqueurs moléculaires hypervariables avec un taux de substitution allant de 10^{-6} à 10^{-2} par kb (kilobase) (Harry 2001). Ils sont de ce fait très utilisés au niveau intraspécifique pour l'étude de la structure génétique et l'inférence des évènements historiques qui se sont produits au sein des populations d'organismes vivants. Les microsatellites sont des répétitions d'un nombre variable de di, tri, ou tétranucléotides, très représentées dans le génome des eucaryotes (Beckmann & Weber 1992), avec une

distribution plus ou moins régulière (Hamada *et al*. 1982). Ils sont de taille variable, mais, celle-ci n'excède généralement pas 200 paires de bases. Dans certains cas, l'augmentation de la taille des microsatellites peut être directement source de graves pathologies ou associée aux cancers dans l'espèce humaine (Caskey *et al*. 1992 ; Imbert *et al*. 1993 ; Thibodeau *et al*. 1993). L'origine des microsatellites reste peu connue même si actuellement deux hypothèses sont fréquemment avancées pour expliquer leur apparition : (1) la duplication des sites cibles de certains types d'éléments transposables avant leur excision ou (2) la méthylation des résidus adényliques en guanine. L'évolution et la différenciation ultérieures de ces sites cibles se produit au cours de la réplication de l'ADN. Leur duplication se réalise grâce au glissement de l'ADN polymérase sur des courtes séquences nucléotidiques répétées qui sont initialement apparues du simple fait du hasard lors du processus de mutation. Il peut aussi y avoir délétion de quelques répétitions de ces microsatellites grâce à la formation des boucles de délétion D'une manière générale, le pouvoir de résolution des microsatellites en analyse génétique et historique fait qu'ils sont de plus en plus utilisés pour répondre aux questions de microévolution.

Il existe plusieurs avantages à l'utilisation des microsatellites : (1) le degré de polymorphisme est très élevé car le nombre de répétitions peut être très variable du fait de leur grande vitesse d'évolution moléculaire ; (2) ces marqueurs moléculaires sont repartis plus ou moins régulièrement sur l'ensemble du génome nucléaire, ce qui assure une bonne couverture de celui-ci ; (3) les différents allèles des microsatellites sont facilement identifiables à partir d'une simple amplification par PCR. Ils ont permis le développement d'outils puissants pour la modélisation statistique de leur évolution et leur utilisation en génétique des populations.

Dans la classe des insectes, les microsatellites, isolés chez des espèces d'Hyménoptères (guêpes, abeilles et bourdons, fourmis) et d'Isoptères (termites) sociaux, montrent une forte fréquence dans le génome et un niveau de polymorphisme comparable à celui observé chez les vertébrés (Choudhary *et al*. 1993 ; Evans 1993 ; Hughes & Queller 1993 ; Estoup *et al*. 1994), les acridiens (Ustinova *et al*. 2006), les mouches de fruits (Baliraine *et al*. 2004). Cependant, on observe au sein d'un groupe donné d'organismes que la fréquence et la distribution des microsatellites varient largement comme l'ont montré Nève & Meglécz (2000) dans le cas des Lépidoptères et des Hyménoptères. D'une manière générale chez les Lépidoptères, ordre d'insectes auquel appartient *B. fusca* (le présent matériel d'étude), Zhang (2004) conclut que les microsatellites sont caractérisés par deux propriétés importantes : une faible fréquence avec une forte redondance des séquences flanquantes (les

régions qui sont situées de chaque côté des microsatellites), ce qui conduit souvent à l'amplification des séquences microsatellites paralogues. La différentiation ultérieure de ces copies se produit à la suite des mutations moléculaires (duplication et délétion) issues des nombreux mécanismes cellulaires liés à la division de la cellule. Les séquences nucléotidiques microsatellites des Lépidoptères pourraient n'en être qu'au début de leur évolution moléculaire ou, alternativement, il existerait dans ce groupe d'organismes des mécanismes cellulaires (association avec les éléments mobiles ou éléments transposables) qui empêchent l'évolution et la différenciation ultérieures de ces marqueurs nucléaires (Zhang 2004). Plusieurs études ont souligné les difficultés d'identification et d'isolement des microsatellites chez les Lépidoptères (Palo *et al*. 1995 ; Meglecz & Solignac 1998). Néanmoins, les techniques moléculaires de plus en plus adéquates ont permis récemment d'isoler des locus microsatellites pour caractériser la diversité biologique et la structure génétique des populations de cet ordre d'insectes (Williams *et al*. 2003 ; Harper *et al*. 2003 ; Meglecz *et al*. 2004, Faure & Silvain 2005).

1.4. Intérêts

D'une façon générale, les marqueurs moléculaires permettent d'inférer la phylogénie, la biogéographie, la phylogéographie et l'écologie moléculaires des groupes taxonomiques. Du fait de leur facilité croissante d'utilisation, les marqueurs moléculaires sont de plus en plus utilisés en entomologie moderne. Les exemples sont nombreux et illustrent parfaitement les différents champs de la biologie évolutive utilisant ces marqueurs moléculaires. De nouvelles méthodes apparaissent pour bien établir le lien entre l'entomologie classique et les outils moléculaires. Bien que la comparaison des données issues de l'analyse de différents gènes présente un avantage certain, Brower & DeSalle (1994) soulignent qu'il ne faut pas retenir un gène particulier pour une analyse donnée simplement sur la seule base de la fréquence de son utilisation ou de sa popularité dans la littérature. L'utilité de chaque région génique de l'ADN mitochondrial a été discutée par Simon *et al*. (1994). Des marqueurs moléculaires ayant des caractéristiques précises et répondant à des problématiques biologiques précises ont été utilisés ces dernières décennies afin d'évaluer la diversité génétique (le cytochrome *b*, les microsatellites), de retracer l'histoire des populations (gènes codant pour le COI, le COII), d'étudier l'histoire évolutive et démographique des espèces (les alloenzymes et la région de contrôle de l'ADN mitochondrial) ou d'établir la phylogénie des populations, des espèces ou des groupes d'espèces (COI, COII, 28S, 18S) (Simon *et al*. 1994). Les premières études

basées sur les séquences nucléotidiques utilisaient surtout les marqueurs mitochondriaux auxquels se sont ajoutés ces dernières années les gènes nucléaires. Les études récentes en génétique des populations ou en biologie évolutive associent les deux types de marqueurs dans le but de compléter ou de combiner les informations que pourrait apporter individuellement l'un ou l'autre génome. C'est le cas de la phylogénie des Neuroptères où l'utilisation combinée des marqueurs moléculaires mitochondriaux et nucléaires ont clarifié les relations phylogénétiques entre les différents sous-ordres (Haring & Aspöck 2004). Il en de même pour les chironomes australiens pour lesquels l'utilisation des marqueurs nucléotidiques a permis de différencier les espèces du genre *Cladopelma* (Diptera, Chironomidae) (Carew *et al*. 2005).

Section II : Application et utilisation des outils moléculaires en biologie évolutive des populations et des espèces d'insectes

Les marqueurs moléculaires ont vu leur domaine d'application s'élargir ces dernières années car ils permettent d'aborder une diversité de questions de biologie fondamentale et appliquée. La caractérisation des populations, des races d'hôtes, des espèces ou des complexes d'espèces est utile tant à la biologie de la conservation, qu'à la connaissance du fonctionnement des populations d'espèces nuisibles aux cultures ou à celles de leurs auxiliaires. Elle vise à élaborer des stratégies de lutte biologique ou intégrée contre les insectes ravageurs des plantes cultivées (Coléoptères Bruchidae ravageurs des légumineuses alimentaires (Sembène & Delobel 1998), les Lépidoptères nuisibles aux cultures (Timm *et al*. 2006), les Orthoptères ravageurs de cultures (Yang *et al*. 2005)), à développer une politique de conservation des différentes races d'une espèce d'insecte d'intérêt agronomique (Coccinelles prédatrices des Aphides (Chen *et al*. 2000), Hyménoptères parasitoïdes (Baker *et al*. 2003)) ou à envisager des moyens de protection des espèces en voie de disparition ou menacées suite à la modification des habitats ou des écosystèmes qui peut engendrer des goulots d'étranglement ou l'extinction des espèces d'insectes.

2.1. Etude des capacités de dispersion

L'histoire démographique des populations d'une espèce donnée ou d'un groupe d'espèces peut avoir un impact profond sur la distribution de la variation génétique au sein de

populations actuelles. La capacité de dispersion des insectes est connue grâce au différents indices de différenciation. Par exemple, l'obtention d'une valeur faible de Φ_{ST} (indice de différenciation génétique) est souvent considérée comme la preuve d'un flux génique continu entre les différentes populations étudiées alors qu'une forte valeur de cet indice est interprétée comme la preuve de l'absence d'échange génétique. Il faut signaler qu'une structuration génétique observée aujourd'hui peut être le résultat d'un isolement ancien plutôt que d'une fragmentation récente, liée par exemple aux actions humaines (Cunningham & Moritz 1998). Dans le cas d'un isolement très récent, la différenciation moléculaire entre populations peut être non perceptible du fait du maintien des caractères ancestraux (Avise *et al*. 1987). Le maintien d'un polymorphisme partagé entre deux ou plusieurs populations d'une espèce donnée peut donc indiquer un polymorphisme ancestral ou au contraire une homogénéisation génétique des populations. On peut estimer le nombre de migrants échangés entre les différentes populations d'une espèce. Ce paramètre, lié directement à l'indice Φ_{ST}, explique en partie la structure génétique présentée par les populations. L'insecte phytophage *Gonioctena pallida* (Coleoptera, Chrysomelidae) présente une forte différenciation génétique interpopulationnelle du fait de sa capacité de dispersion limitée associée à la fragmentation de l'aire de distribution de sa plante hôte (Mardulyn 2001). En effet, il est montré chez cette espèce que la dispersion est assurée uniquement par les larves qui réalisent seulement des déplacements sur quelques mètres de distances, les adultes bien que dotés d'une capacité de vol le font très rarement (Mason & Lawson 1982). On peut aussi citer le cas des populations d'autres chrysomèles comme *Chrysomela aeneicollis* (Rank 1992) ; *Plagiodera versicolora* (McCauley *et al*. 1998) où de fortes valeurs de Φ_{ST} ont été observées entre des localités proches. Les populations des espèces appartenant aux genres *Oreina* et *Gonioctena* présentent une forte différenciation génétique bien qu'elles ne soient séparées que de quelques kilomètres (Knoll *et al*. 1996). La répartition géographique de la ressource alimentaire peut aussi limiter la dispersion des espèces.

Par contre pour d'autres espèces, la différenciation est presque inexistante du fait de la présence continue de flux géniques. La structure des populations de *Phyllodecta vulgatissima* (Coleoptera, Chrysomelidae) et de *P. vitellinae* (Coleoptera, Chrysomelidae), ravageurs des plantations de l'osier (saule) au Royaume-Uni révélée à l'aide des microsatellites, montre un taux d'échanges génétiques très fort (nombre élevé de migrants) et peu de preuves en faveur d'une subdivision des populations (Φ_{ST} faible) en relation avec la géographie, la période d'échantillonnage ou de la plante hôte (Batley *et al*. 2004). L'étude des

populations australiennes de *Plutella xylostella* (Lepidoptera, Plutellidae), ravageur de la culture de *Brassica*, montre également l'absence de structuration génétique entre celles-ci, ce qui démontre vraisemblablement l'existence d'un flux génique continu entre les populations appartenant aux différentes zones géographiques d'Australie et de Nouvelle-Zélande (Endersby *et al*. 2006).

2.2. Histoire des populations et des espèces

L'analyse combinée des microsatellites et de l'ADN mitochondrial a permis de déterminer la structure génétique des populations et de déduire l'histoire de colonisation de la mouche de l'olive, *Bactrocera oleae* (Diptera, Tephritidae), considérée comme un ravageur majeur de l'olive à travers les régions de culture d'Oleacae (Nardi *et al*. 2005). Une étude similaire conduite avec les microsatellites dans les populations du papillon *Polyommatus bellargus* (Lepidoptera, Lycaenidae) a permis de conclure que l'isolement par la distance et les changements démographiques n'influencent pas la diversité génétique de cette espèce mais que la cause cruciale est le faible effectif des populations (Harper *et al*. 2003).

Les marqueurs moléculaires permettent de suivre l'histoire évolutive des plantes hôtes et des insectes phytophages qui leur sont associés en mettant en évidence les changements, opérés par ces derniers, entre des plantes botaniquement proches (même famille botanique) ou éloignées (différentes familles botaniques) au cours du temps (Yotoko *et al*. 2005 ; Ohshima & Yoshizawa 2006). Des cas identiques sont connus chez *Rhagoletis pomonella* et *R. zephyria* (Jeffrey *et al*. 1999) ; *Apis mellifera* (Hymenoptera : Apidae) (Franck *et al*. 2001). Les études ont mis en évidence dans certains cas l'histoire d'une divergence sympatrique via le changement de plantes hôtes (*Rhagoletis pomonella* et de *R. zephyria*) et ont permis dans d'autres l'identification des lignées évolutives et la caractérisation des zones de différenciation des populations (*Apis mellifera*).

2.3. Evolution de la phytophagie : races d'hôtes, résistance aux insecticides

En ce qui concerne l'étude des mécanismes à l'origine de la biodiversité chez les insectes phytophages, les reconstructions phylogénétiques à l'aide des marqueurs moléculaires permettent de mettre en évidence les relations coévolutives qui se sont produites au sein des différents taxons étudiés (Funk *et al.* 1995 ; Becerra 1997 ; Mardulyn *et al.* 1997 ; Silvain & Delobel 1998). Farrell (1998) a pu récemment affirmer que l'accroissement de la

diversité spécifique chez les Coléoptères phytophages (Chrysomeloidea) était dû à une série de radiations adaptatives sur les Angiospermes. Kergoat *et al*. (2004) ont utilisé les marqueurs moléculaires mitochondriaux avec différentes approches de reconstruction phylogénétique pour établir la phylogénie des espèces de bruches européennes et leur relation avec leurs plantes hôtes, et appréhender les mécanismes à l'origine de la diversification de ce groupe d'insectes.

D'un point de vue plus fondamental, l'étude des races d'hôtes permet de décrire les patrons de diversification et d'en comprendre les mécanismes. Chez les espèces d'insectes phytophages, de nombreuses observations montrent que la tendance à la spécialisation dans l'exploitation des plantes hôtes a été l'un des principaux facteurs ayant favorisé la diversification de ce groupe animal (Bernays & Chapman 1994). Ce qui conduit à la formation des races d'hôtes dont les exemples sont très nombreux dans le groupe des insectes. On peut citer le cas de *Caryedon serratus* (Coleoptera, Bruchidae) (Sembène 2000), bruche inféodée à l'arachide. D'autres études, utilisant différents marqueurs moléculaires, ont permis de mettre en évidence des races d'hôtes écologiquement distinctes chez plusieurs espèces d'insectes phytophages (Ryckewaert & Alauzet 2001 ; Levy *et al*. 2002 ; Thomas *et al*. 2003 ; Nagoshi & Meagher 2003 ; Pashley Prowell *et al*. 2004).

Les analyses moléculaires en biologie comportementale et en génétique des populations des insectes ravageurs ont permis d'identifier et de caractériser certains gènes de résistance à certaines substances chimiques constituant les insecticides et de découvrir les mécanismes évolutifs qui sous-tendent les relations tritrophiques : un cas très spectaculaire est celui du puceron ravageur du cotonnier, *Aphis gossypii*, chez qui deux mutations dans le gène codant l'acétylcholinestérase (*AchE*) confèrent la résistance aux insecticides de type carbamate et organophosphate (organophosphorés) (Andrews *et al*. 2004). Cette résistance aux insecticides permet aux individus de cette espèce de puceron d'être des ravageurs potentiels du cotonnier en rendant la lutte chimique inefficace et en entraînant des nuisances économiques importantes. Toda *et al*. (2004) ont montré que ces mutations nucléotidiques entraînent le remplacement de deux acides aminés. Par ailleurs, une analyse préliminaire des séquences nucléotidiques de ce gène chez les individus résistants aux insecticides de certaines espèces d'Hémiptères phytophages (*Myzus persicae*, *A. gossypii*, *B. tabaci*, *Trialeurodes vaporarium*) montre qu'elles sont parfaitement identiques, suggérant des mécanismes communs de résistance aux produits chimiques (Javed *et al*. 2003). Dans les programmes de recherche à finalité agronomique, plusieurs gènes de résistance aux insecticides ont été identifiés chez les insectes ravageurs de cultures (Enayati *et al*. 2005).

2.4. Espèces cryptiques

La combinaison séquentielle de différentes approches méthodologiques et analytiques a permis de clarifier la position et, parfois, de déterminer l'identité de certains taxons cryptiques. Ce cas est couramment rencontré dans plusieurs taxons d'insectes sympatriques ayant des larves et des adultes morphologiquement semblables qui exploitent les mêmes ressources. On peut citer le genre *Tetramesa* (Hymenoptera, Eurytomidae) où les analyses moléculaires ont permis de préciser le niveau taxonomique et d'établir les liens phylogénétiques entre certaines espèces (Al-Barrak *et al*. 2004). Une analyse basée sur les RFLP du cytochrome *b* chez les Aphides, une approche moléculaire simple, a identifié quatre espèces cryptiques : *Aphis gossypii*, *Aphis crassivora*, *Myzus persicae* et *Macrosiphum euphorbiae* (Raboudi *et al*. 2004). C'est aussi le cas des espèces du criquet de genre *Laupala* dont les AFLP ont précisé la limite des espèces cryptiques tout en caractérisant la diversité génétique au sein des espèces hawaïennes (Parsons & Shaw 2001). L'identification des espèces cryptiques à l'aide des marqueurs moléculaires a pu se faire dans les différents ordres d'insectes (Berkov 2002). En Amérique du Nord, l'histoire et la diversification au sein des espèces des papillons du genre *Lycaeides* ont été retracées grâce aux analyses de conformation du polymorphisme de simple brin (Nice *et al*. 2005).

Les outils moléculaires peuvent permettre de résoudre des questions plus ponctuelles de systématique. Ils ont conduit à la description de nouvelles espèces. Alvarez *et al*. (2004) en réalisant l'étude combinée des caractères morphologiques et moléculaires sont parvenus à conclure que les bruches *Acanthoscelides obtectus* et *Acanthoscelides obvelatus* sont deux bonnes espèces ayant divergé il y a 22 millions d'années. Ces auteurs ont par ailleurs souligné que la forte ressemblance morphologique entre ces deux espèces de bruches pourrait résulter de la stase évolutive (conservation des certains caractères) qui découlerait d'une sélection stabilisante.

Parallèlement, des variations morphologiques importantes des individus entre les populations réparties sur de grande échelle géographique sont souvent à l'origine de la confusion dans la détermination des espèces. Ce qui consiste à assimiler certaines populations à des espèces cryptiques. Ce cas est bien connu avec les populations d'Argentine d'*Anastrepha fraterculus* (Diptera, Tephritidae). Les analyses moléculaires au sein de ses populations ont montré que celles-ci, considérées comme des espèces cryptiques, appartenaient à une espèce biologique déjà décrite : il s'agit de l'espèce nominale *A. fraterculus* (Alberti *et al*. 2002). Les études moléculaires peuvent aussi confirmer les

observations morphologiques comme dans le cas des espèces du genre *Liriomyza* (Kox *et al*. 2005). Chez les espèces du complexe *Drosophila serrata*, les microsatellites ont permis d'identifier les différentes espèces qui forment ce sous-groupe (Schiffer *et al*. 2004). Le taxon *Calligrapha multipunctata* var. *suturella* a été élevé au rang d'espèce, *Calligrapha suturella* (Coleoptera, Chrysomelidae), grâce aux analyses moléculaires de génétique des populations (Gómez-Zurita *et al*. 2004).

2.5. Rôle des espèces modèles (I) : définition de patrons biogéographiques

L'identification de clades géographiquement distincts et bien structurés résultant vraisemblablement d'un événement de fragmentation passée est mise en évidence dans les études phylogéographiques. Gómez-Zurita (2002), en étudiant les populations du complexe d'espèce *Timarcha goettingensis* (Coleoptera, Chrysomelidae) à l'aide du gène mitochondrial COII, a montré qu'il existe au sein de ce taxon deux clades issus d'un événement de fragmentation passée.

La phylogéographie des populations thaïlandaises de la mouche *Simulium tani* (Diptera, Simuliidae) au moyen des séquences de l'ADN mitochondrial a mis en évidence une forte différenciation génétique entre trois régions géographiques majeures (Pramual *et al*. 2005). Cette étude a en outre identifié la région Nord comme le centre d'origine des populations Est et Centre/Sud thailandaises. Les analyses de mismatch distribution indiquent que ces populations ont connu des phénomènes d'expansion démographique au Pléistocène moyen et supérieur suggérant que leur structure génétique actuelle et leur diversité moléculaire seraient, en partie, liées à l'adaptation des espèces aux fluctuations climatiques de cette période géologique. La phylogéographie comparée de deux espèces parthénogénétiques *Aneugmenus padi* et *Eurhadinoceraea ventralis* (Hymenoptera, Trenthredinidae) a révélé des différences de structure génétique entre ces espèces qui sont vraisemblablement corrélées avec les différences dans la distribution géographique de leurs plantes hôtes (Müller *et al*. 2004).

2.6. Rôle des espèces modèles (II) : compréhension des mécanismes de changement d'hôte et origine des insectes ravageurs

Les insectes phytophages constituent de bons modèles d'étude pour examiner les facteurs et les conséquences évolutives des différenciations génétiques, par exemple, ceux en relation avec la distribution géographique de leurs plantes hôtes. Ils nous renseignent sur leur propre histoire évolutive ainsi que celle des différentes plantes hôtes qu'ils ont connues. L'origine des insectes ravageurs est à rechercher dans l'évolution de la phytophagie (exploitation des plantes sauvages apparentées pour aboutir ultérieurement à l'utilisation de leurs formes cultivées), phénomène biologique couramment rencontré chez les espèces d'insectes. Elle influence la structure génétique des différentes populations constituant l'espèce. En effet, une partie ou la totalité des populations d'un insecte phytophage commence à un moment de l'histoire évolutive de l'espèce à s'attaquer aux plantes cultivées présentes dans le même environnement que ses plantes hôtes sauvages suite, vraisemblablement, à une forte anthropisation (agriculture intensive, destruction des formations végétales naturelles). Des exemples illustrant de tel changement d'hôtes sont bien connus. Au sein d'une espèce, ce phénomène est initié par certaines populations qui, par sélection naturelle et dérive génétique, peuvent se différencier génétiquement des populations restées sur les plantes hôtes sauvages originelles. Cette divergence génétique, plus ou moins rapide selon les espèces d'insectes consécutivement à l'anthropisation et à la domestication des phytophages, pourrait être à l'origine de l'apparition de races d'hôtes ou nouvelles espèces.

Comprendre l'évolution de la spécialisation alimentaire chez les insectes phytophages va donc nous éclairer quant aux contraintes et aux facteurs biologiques et écologiques agissant sur les processus de changement d'hôte impliqués dans l'apparition des insectes ravageurs. Les cas les plus spectaculaires de ségrégation écologique associée à une structuration génétique sont souvent observés chez les insectes phytophages. Les études conduites avec les marqueurs moléculaires ont apporté plusieurs résultats en faveur de l'hypothèse de spéciation sympatrique. Cette hypothèse est fréquemment retenue chez les insectes phytophages où l'on peut observer une structuration génétique ou une ségrégation écologique suite à un changement de plante hôte ou à l'ajout d'autre nouvelle espèce au spectre de plantes hôtes d'un insecte donné (Berlocher 1998). Les insectes phytophages constituent de ce fait des modèles pour tester l'hypothèse d'une spéciation sympatrique via les plantes hôtes. Celle-ci a été défendue historiquement par Bush (1975) qui pense que les accouplements préférentiels entre partenaires provenant de la même plante hôte, les

préférences alimentaires et/ou les sites d'oviposition, les adéquations physiologiques au nouvel hôte sont les facteurs qui rendent possible cette spéciation sympatrique. La phénologie différentielle des plantes hôtes exploitées favorise des accouplements préférentiels (choix privilégié du partenaire issu d'un même habitat) au sein des niches écologiques. Ces processus conduisent finalement à une ségrégation écologique avec formation des races d'hôtes. Seuls quelques caractères, impliquant une partie du génome de l'insecte, seraient responsables dans cette adaptation à une nouvelle plante hôte qui est à l'origine de la structuration écologique sympatrique (Thomas *et al*. 2003 ; Schwartz *et al*. 2003), d'où l'intérêt d'une diversification des marqueurs utilisés. Dans le cas de *Rhogoletis cerasi*, les études de la ségrégation écologique ont montré que la différenciation de races d'hôte concerne principalement l'enzyme mannose 6-phosphate isomérase (*Mpi*), probablement impliquée directement dans la différenciation adaptative des populations (Schwarz *et al.* 2003). La même enzyme fait partie de celles permettant de différencier des races d'hôtes chez l'espèce sœur *R. pomonella*. Les auteurs ont alors conclu à l'existence d'un pool génétique commun aux espèces du genre *Rhagoletis* qui est vraisemblablement à l'origine de la formation de races d'hôte au sein de ce groupe. Dans les cas des Coléoptères ravageurs de l'écorce du bois, les marqueurs moléculaires ont permis de quantifier l'effet de la spécialisation sur la différenciation génétique des espèces sœurs (Kelley *et al*. 2000).

2.7. Quelques travaux de biologie moléculaires sur les espèces d'insectes ravageurs de cultures

Les tableaux 1-3, 1-4, font la synthèse des études d'analyse moléculaires chez les insectes phytophages.

Tableau 1-3 : quelques études de la diversité et de la structure génétique des populations d'insectes phytophages

Etudes moléculaires	Marqueurs	Auteurs et année
Microsatellites reveal a lack of structure in Australian populations of the diamondback of the diamondback moth, *Plutella xylostella* (L.).	Microsatellites	Endersby *et al*. 2006
Genetic structure of the whitefly *Bemisia tabaci* in the Asia-Pacific region revealed using microsatellite markers.	Microsatellite	De Barro 2005
Population structure of the beetle pests *Phyllodecta vulgatissima* and *P. vitellinae* on UK willow plantations.	Microsatellites	Batley *et al*. 2004
Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites.	Microsatellites	Llewellyn *et al*. 2003
Hierarchical genetic structure of the introduced wasp *Vespula germanica* in Australia	Microsatellites	Goodisman *et al*. 2001
Genetic variation and origin of the red turpentine beetle (*Dendroctonus valens* LeConte) introduced to the people's Republic of China.	COI	Cognato *et al*. 2005
Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae)	NADH dehydrogenase gene	Nardi *et al*. 2005
Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host.	PCR-RFLP-SSCP	Burban *et al*. 1999
The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations.	AFLP, COII	Grapputo *et al*. 2005
Genetic variation in geographical populations of western and Mexican corn rootworm.	ITS1, mtDNA	Szalanski *et al*. 1999

Tableau 1-4 : quelques études de la cartographie, de la phylogénie et de la caractérisation moléculaire chez les insectes phytophages

Etudes moléculaires	Marqueurs	Auteurs et année
The internal transcribed spacers (ITS1 and ITS2) of the rDNA differentiates the bark beetle forest pests *Tomicus destruens* and *T. piniperda*.	ITS1, ITS2.	Gallego & Galian 2001
Sibling species of bean bruchids: a morphological and phylogenetic study of *Acanthoscelides obtectus* Say and *Acanthoscelides obvelatus* Bridwell.	Morphological characters, COI, 28S, 12S, microsatellites.	Alvarez *et al*. 2005
A genetic linkage map for *Tribolium confusum* based on random amplified polymorphic DNAs and recombinant inbred lines.	RAPD	Yezerski *et al*. 2003
Lack of nucleotide variability in a beetle pest with extreme inbreeding	Gene of resistance to dieldrin, ITS2, COI	Andreev *et al*. 1998
Two amino acid substitution in acetylcholinesterase associated with pirimicarb and organophosphorous insecticide resistance in cotton Aphid, *Aphis gossypii* Glover (Homoptera: Aphididae).	Séquences polypeptidiques	Toda *et al*. 2004
Identification of mutations conferring insecticide-insensitive AChE in the cotton-melon aphid, *Aphis gossypii* Glover.	Séquences polypeptidiques	Andrews *et al*. 2004
Characterization of acetycholinesterases, and their genes, from the hemipteran species *Myzus persicae* (Suler), *Aphis gossypii* (Glover), *Bemisia tabaci* (Gennadius) and *Trialeurodes vaporariorum* (Westwood).	Séquences polypeptidiques	Javed *et al*. 2003

Dans la littérature scientifique, les taxons nord-américains sont les plus représentés et les plus étudiés en analyse moléculaire. Ils sont suivis des espèces européennes. Quelques études ont été faites en Asie et Australie alors que très peu d'attention a été portée sur les insectes africains, en particulier, sur ceux qui sont phytophages parmi lesquels les ravageurs de cultures. Cela justifie, en grande partie, cette étude de génétique des populations et de phylogéographie sur *Busseola fusca*, un ravageur africain des cultures de céréales.

Chapitre 2
Le modèle *Busseola fusca*

Section I : Biologie et écologie de *B. fusca*

Busseola fusca (Fuller 1901) appartient taxonomiquement au phylum des Arthropodes ; à la classe des Hexapoda ; à la sous-classe des Pterygotae ; à l'ordre des Lepidoptera ; à la famille des Noctuidae ; à la sous-famille des Amphipyrinae ; au genre *Busseola* et à l'espèce *fusca*.

Il est, sur le continent africain, l'un des ravageurs les plus importants sur le plan économique de certaines graminées cultivées (Polaszek & Khan 1998). Ses chenilles forent principalement les tiges de maïs et du sorgho cultivé. *B. fusca* est une espèce endémique en Afrique subsaharienne où il présente une large distribution géographique (Ajayi 1998; Moyal 1998 ; Ndemah *et al*. 2001 ; Haile & Hofsvang 2001 ; Kfir *et al*. 2002) : il s'étend approximativement entre 12°N et 30°S (Appert 1970). Les figures 2-1, 2-2 représentent respectivement une chenille et un mâle de *B. fusca*.

Figure 2-1 : œufs et chenilles de *Busseola fusca* aux différents stades (UR-IRD 072, Gif-sur-Yvette).

Figure 2-2 : Un mâle de *Busseola fusca* (G. Goergen, Biodiversity Center, IITA Benin).

1.1. Biologie de *B. fusca*

Le cycle biologique de *B. fusca* comprend les différents stades classiques du développement des noctuelles (œufs, chenilles, chrysalides, imagos), mais, sa durée, notamment celle des stades larvaires, dépend des conditions écologiques.

Comme chez toutes les noctuelles, le cycle biologique de *B. fusca* débute par une phase de rapprochement des sexes qui se traduit par l'attraction des mâles par les femelles via l'émission de phéromones sexuelles (substances chimiques volatiles et attractives). Le rapprochement des sexes conduit à l'accouplement et à la fécondation des ovocytes. Celle-ci est suivie par la ponte. Les œufs sphériques jaunes sont pondus par les femelles en amas (30 à 100) entre la gaine foliaire et la tige de la graminée. La période d'incubation des œufs est influencée par la température et l'humidité relative (RH). Usua (1968) observe qu'une élévation de la température associée à une humidité relative de l'ordre de 80%, ou plus, entraîne l'éclosion des œufs entre les 5ème et 6ème jours après la ponte alors que pour des taux d'humidité relative de 70% et 60%, l'éclosion se produit respectivement entre 6-7 jours et 7-8 jours. Ce même auteur rapporte qu'une température de 22°C entraîne l'arrêt du développement des chenilles alors que celles de 32°C et plus sont considérées comme létales. Il conclut que les températures comprises entre 26°C et 30°C sont les plus favorables et les plus appropriées pour le développement des œufs de *B. fusca*. Après émergence, les jeunes chenilles, au comportement grégaire, décapent les feuilles tendres du verticille des céréales et peuvent détruire le point de croissance de la plantule entraînant le symptôme du "cœur mort" (Bosque-Pérez & Schulthess 1998). Les chenilles restent phyllophages pendant quelques jours puis pénètrent à l'intérieur des tiges en réalisant au préalable des perforations de celles-ci. Une fois à l'intérieur des tiges de céréales, les chenilles se nourrissent de la moelle de la tige puis elles expulsent vers l'extérieur de celle-ci de grande quantité de fines poussières correspondant aux fèces et aux déchets alimentaires. Le temps de développement des chenilles à l'intérieur des tiges varie entre 26 et 33 jours (Bosque-Pérez & Schulthess 1998). Ce développement conduit les chenilles jusqu'aux stades 6 et 8 entre le 30ème et le 45ème jour après l'éclosion (Holloway 1998). Les dégâts entraînent un affaiblissement général de la plante qui, en cas de fortes précipitations, est alors soumise à un taux de verse ou de casse important. Dans certains cas, ces dégâts peuvent toucher directement les jeunes épis de maïs par destruction de leur axe porteur des grains. Avant de se métamorphoser en chrysalide, la chenille fore un orifice dans les tissus extérieurs de la tige du végétal par lequel sortira le papillon.

Une étape supplémentaire peut intervenir au cours des derniers stades larvaires. Lorsque les conditions sont défavorables (saison sèche ou période plus froide), les chenilles de *B. fusca* entrent en diapause qui est un processus biologique qui correspond à un arrêt du développement. Cette diapause est obligatoire chez *B. fusca* dans les mauvaises conditions (Kfir 1991) et peut durer jusqu'à six mois (Holloway 1998 ; Kfir *et al*. 2002) suivant les conditions climatiques. Avec la reprise des pluies à la saison suivante, ces chenilles se métamorphosent en chrysalides à l'intérieur des tiges. Elles émergent 10 à 12 jours plus tard en donnant des papillons adultes (Bosque-Pérez & Schulthess 1998). En général, *B. fusca* présente deux générations par an, mais dans certaines régions où la diapause est réduite à quelques semaines ou absente, il peut y avoir chaque année jusqu'à quatre générations de cette espèce (Holloway 1998). La figure 2-3 résume les différentes étapes du développement de *B. fusca*.

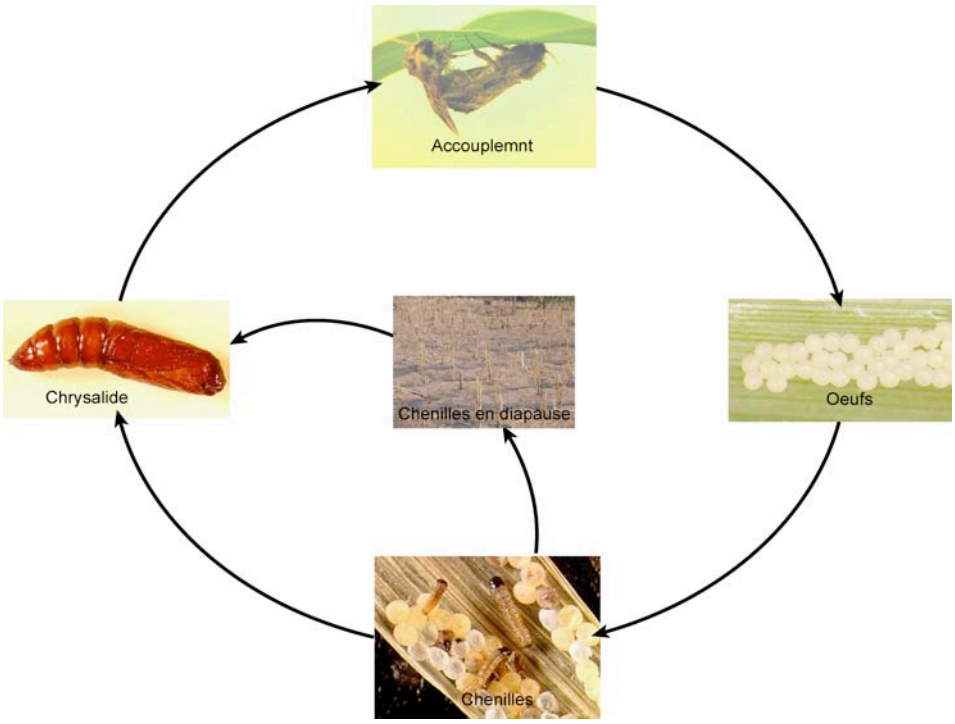


Figure 2-3 : Cycle biologique de *B. fusca*.

1.2. Ecologie de *B. fusca*

Plusieurs espèces d'insectes foreurs peuvent exploiter la tige d'une même espèce de céréale. Leur coexistence implique qu'il y ait des phénomènes de compétition et de différenciation des niches écologiques entre les espèces susceptibles d'exploiter la même

ressource. La plupart des espèces vont se spécialiser dans l'exploitation d'une niche écologique bien précise. A l'échelle régionale, on constate en Afrique de l'Est, du Centre et du Sud que cette spécialisation de niche se fait en fonction du climat et de l'altitude. Dans ces régions, *B. fusca* est un ravageur majeur des cultures de maïs et de sorgho à des altitudes moyennes et élevées (> 600 m) (Nye 1960 ; Sithole 1989 ; Ndemah *et al*. 2001 ; Kfir *et al*. 2002). En Afrique de l'Ouest, cette ségrégation de niche écologique se fait en fonction du climat et de la culture céréalière. *B. fusca* semble s'être adapté à des biotopes de plus basses altitudes avec des dégâts plus importants en zones de savanes sur les cultures de sorgho qu'en zones de production de maïs (Harris 1962 ; Harris & Nwanze 1992 ; Ajayi 1998). Les figures 2-4, 2-5 et 2-6 montrent quelques biotopes de *B. fusca* en Afrique de l'Ouest.

Figure 2-4 : Champ de maïs à Attogon infesté majoritairement par les chenilles de *B. fusca*. On y trouve aussi les chenilles de *S. calamistis* et quelques fois celles de *E. saccharina* (Rép. du Bénin) (Octobre 2004).

Figure 2-5 : Tiges sèches partiellement brûlées de sorgho cultivé hébergeant exclusivement les chenilles diapausantes de *B. fusca* dans un champ situé sur la Route Boukounbé (Rép. du Bénin) (Mars 2004).

Figure 2-6 : Recherche dans un champ de Nikki (Rép. du Bénin) des chenilles et chrysalides de *B. fusca* dans les repousses et tiges sèches de sorgho cultivé. On y rencontre aussi celles de *S. calamistis* (Mars 2004).

Au sein des foreurs de tiges, les niches écologiques se recoupent. *B. fusca* évolue dans une communauté diversifiée de foreurs, en particulier, avec des ravageurs des mêmes cultures comme *Chilo partellus*, *Sesamia calamistis* à travers toute l'Afrique Subsaharienne (Polaszek 1998). En fonction des régions écologiques, d'autres foreurs de tiges peuvent être mentionnés. Il s'agit de *C. orichalcociliellus*, *C. diffusilineus*, *S. poephaga, S. cretica, S. penniseti, S. nonagrioides, S. botanephaga, Coniesta ignefusalis* ou *Spodoptera littoralis* (Ajayi 1998). Dans un même champ, la coexistence des ravageurs est rendue possible par leur spécialisation sur les différents organes de la plante cultivée : on citera le cas de la présence simultanée de *B. fusca*, de *S. calamistis, E. saccharina* qui sont les foreurs de tiges et du Pyralidae *Mussidia nigrivenella* qui est un ravageur d'épis (Cardwell *et al*. 1997).

Il existe une dynamique dans l'écologie des foreurs. Les distributions spatiales évoluent du fait des changements environnementaux liés au réchauffement climatique et aux invasions biologiques. Les études écologiques conduites en Afrique de l'Est et Australe (Kfir 1997 ; Songa 1999) soulignent clairement qu'on assiste aujourd'hui dans ces régions à un déplacement compétitif de *B. fusca* suite à l'introduction accidentelle et à l'invasion du Lépidoptère Pyralidae *C. partellus*. Ce déplacement compétitif est plus visible dans le cas du

sorgho cultivé où la proportion de *C. partellus* sur l'ensemble des populations de foreurs de tiges de céréales est passée de 3% en 1986 à 91% en 1992 (Kfir *et al*. 2002).

Les nuisances économiques liées aux activités des foreurs de tiges de céréales varient largement parmi les zones écologiques, les régions et les saisons de cultures (Bosque-Pérez & Schulthess 1998). Les foreurs de tiges de céréales produisent généralement des dégâts similaires sur les plantes de maïs et de sorgho cultivé (Seshu Reddy 1998). En Afrique Australe, les pertes de la production de maïs et de sorgho cultivé dues aux dégâts causés par les foreurs, principalement *B. fusca* et *C. partellus*, sont estimées de 10% à la perte totale (Kfir 1988). La perte de la production céréalière engendrée par les foreurs varie largement entre les champs : on observe à la fois des champs non attaqués de céréales et des champs totalement détruits (van Rensburg & Bate 1987). Les pertes de la production de sorgho et de maïs en Afrique de l'Est liées aux foreurs de tiges sont du même ordre de grandeur que celles observées en Afrique Australe avec les mêmes variances (Seshu Reddy 1998). Cet auteur a conclu que l'infestation des champs de céréales par les insectes ravageurs de cultures constitue l'un des facteurs majeurs limitant la production de maïs et de sorgho cultivé en Afrique de l'Est. Bien que l'attaque de sorgho cultivé par les foreurs de tiges soit un phénomène répandu en Afrique de l'Ouest et se traduise parfois par la destruction totale de la culture céréalière, Ajayi (1998) reconnaît qu'actuellement des pertes très significatives de produits céréaliers sont rarement mentionnées. Mercer-Quarshie (1970) a indiqué que les foreurs de tiges provoquent une réduction de 4% de la production céréalière au Ghana. En 1990, dans la savane nord guinéenne du Nigeria, Ajayi (1991), en réalisant une comparaison des produits céréaliers provenant de 22 champs pulvérisés ou non avec des insecticides, constate que la perte moyenne des grains céréaliers est de 21% pour trois variétés locales. Par ailleurs, il souligne qu'il est difficile de dissocier les effets liés à *B. fusca* et *S. calamistis*.

Il se dégage de tout ce qui précède qu'il existe une relation étroite entre l'écologie des foreurs de tiges et les nuisances économiques : l'intensité et la nature des dégâts causés par les ravageurs phytophages dépendent des espèces présentes, du stage de la croissance végétale de la plante, du nombre de chenilles exploitant la plante et de la réaction de cette dernière face à l'attaque des foreurs de tiges (Bosque-Pérez & Schulthess 1998).

Les principales plantes hôtes cultivées de *B. fusca* sont le maïs et le sorgho (Kfir *et al*. 2002) qui occupent respectivement les 1er et 2ème rangs parmi les céréales produites en Afrique (Murty & Renard 2001). Ces céréales ont des histoires et des biologies extrêmement diverses (espèces introduites ou autochtones, variétés culturales résistantes ou non aux attaques d'insectes ravageurs, cycle de développement court ou long) dont la connaissance est

nécessaire pour l'élaboration des méthodes de lutte biologique contre leurs nombreux ravageurs, plus particulièrement, contre leurs foreurs de tiges (Polaszek & Khan 1998). A ces deux plantes hôtes cultivées s'ajoute, selon certains auteurs et suivant les régions phytogéographiques, un nombre important de plantes hôtes sauvages appartenant à trois familles de Monocotylédones : Poaceae, Cyperaceae et Typhaceae. A titre d'exemples, Haile & Hofsvang (2001) ont identifié en Erythrée (Afrique de l'Est) seize plantes hôtes différentes de *B. fusca* allant des plantes cultivées aux sauvages. Il en est de même pour l'Afrique Centrale où les plantes hôtes sauvages semblent jouer un rôle important dans l'écologie de *B. fusca*. En revanche, les études conduites par Le Rü *et al*. (in press) au Kenya et dans plusieurs autres pays Est africains ne semblent pas montrer que les plantes hôtes sauvages présentes dans les milieux naturels jouent un rôle significatif dans l'écologie de *B. fusca* : on ne trouve en effet que très rarement *B. fusca* sur des espèces sauvages. Il en est de même pour l'Afrique de l'Ouest : seul *Sorghum arundinaceum*, une espèce botaniquement proche de sorgho cultivé, est signalé au Ghana comme la plante hôte sauvage de *B. fusca* (Gounou *et al*. 1994). Quelques figures (2-7, 2-8) montrent quelques dégâts causés par l'activité biologique de *B. fusca* sur les tiges de mais dans une localité béninoise (Attogon).

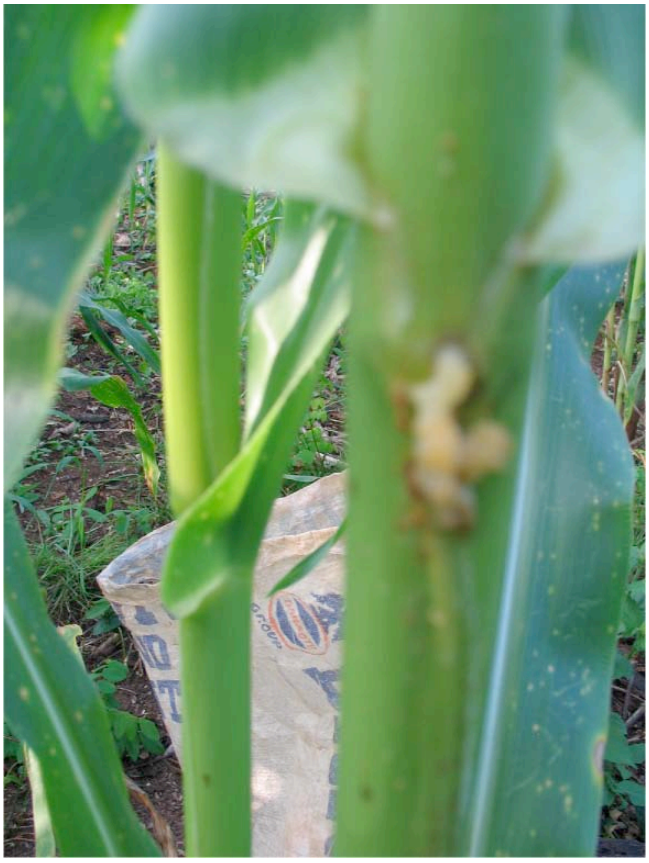


Figure 2-7 : Dégât causé sur le maïs par la chenille de *B. fusca* dans un champ à Attogon (Rép. du Bénin) (Octobre 2004).

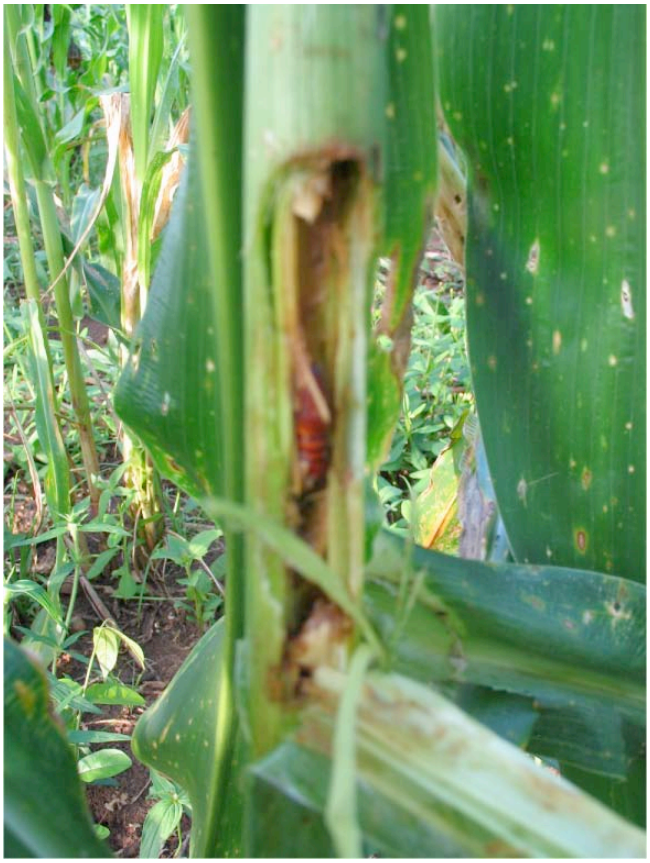


Figure 2-8 : Chrysalide de *B. fusca* à l'intérieur d'une tige de maïs dans un champ à Attogon (Rép. du Bénin) (Octobre 2004).

Parmi les principaux antagonistes de *B. fusca*, figurent les parasitoïdes appartenant principalement à deux ordres d'insectes (Diptères et Hyménoptères) (Kfir *et al*. 2002). D'autres antagonistes dont l'importance varie suivant les régions phytogéographiques régulent les populations de *B. fusca* vivant sur le sorgho cultivé : des prédateurs (cas de la fourmi *Dorylus helvolus* (Kfir 1988)), des nématodes parasites (*Hexamermis* sp.) (Poinar & Polaszek 1998) et des micro-organismes entomopathogènes qui regroupent des champignons (*Aspergillus flavus*, *A. sydowii*, *Metarrhizum* sp.) ; des protozoaires et des virus. Par rapport aux études écologiques réalisées en Afrique du Sud qui ont identifié la fourmi *D. helvolus* comme prédateur de *B. fusca* (Kfir 1988), celles conduites par Mohyuddin & Greathead (1970) en Ouganda (Afrique de l'Est) ont révélé quatre autres espèces de fourmis (*Tetramorium guineense* ; *Phediole megacephala* ; *Carciocondyla badonei* et *C. emeryi*) qui detruisent 90 % des œufs et des chenilles de premier stade de *C. ppartellus* et de *B. fusca*.

De façon générale, tous les stades pré-imaginaux de *B. fusca* sont potentiellement affectés par les effets du parasitisme ou de la prédation des antagonistes : œufs, chenilles, chrysalides et imagos. On constate ainsi que les parasitoïdes mentionnés dans la littérature s'attaquent principalement aux stades larvaires de *B. fusca* en Afrique de l'Est et aux œufs en Afrique de l'Ouest.

Section II : Problématique générale de la thèse

A contrario des taxons nord américains, notamment de vertébrés, on dispose de peu de connaissance sur la génétique des populations et la phylogéographie des insectes africains malgré leur diversité biologique et, pour certains d'entre eux, leur importance économique et agricole (Kutcha & Meyer 2001). En dehors des études écologiques, peu d'études de génétique des populations et de phylogéographie portant sur des espèces d'insectes nuisibles aux cultures en général, et en particulier, sur des espèces foreuses de tiges de céréales ont été réalisées en Afrique. Dans le cas des foreurs de tiges, aucune étude phylogéographique n'avait été réalisée à ce jour. Pourtant, la connaissance de la génétique des populations et de la phylogéographie d'un insecte africain ravageur de cultures à large distribution géographique peut nous éclairer sur sa propre histoire évolutive et sur celle des autres espèces d'insectes africains foreurs des plantes cultivées ou celles associées au peuplement graminéen. L'importance économique des insectes phytophages africains et l'hétérogénéité écologique de certains d'entre eux font des modèles de choix pour des études de biogéographie et d'histoire démographique. Cette hétérogénéité écologique des insectes ravageurs est bien corrélée avec des différences dans les niveaux de nuisances économiques, les communautés d'antagonistes et la géographie. Bien que l'hétérogénéité qualitative ou quantitative des interactions plantes - insectes ravageurs - antagonistes ait été fréquemment observée entre différentes populations géographiques ou écologiques d'un foreur de tiges ou d'autres espèces d'insectes, le phénomène était jusqu'à présent peu pris en compte dans les études à finalité appliquée. Les insectes ravageurs de cultures ont été généralement considérés comme des ensembles écologiquement et génétiquement homogènes. Cette non-prise en compte de la diversité biologique explique, en grande partie, certains échecs ou les succès relatifs d'opérations de lutte biologique ou intégrée. La prise en compte du patron phylogéographique d'un ravageur de cultures présente un intérêt agronomique particulier car la connaissance de la distribution originelle d'une telle espèce peut donner des indications sur les conditions climatiques ayant permis l'évolution ultérieure de celle-ci et l'établissement des relations coévolutives hôte-parasites. Ce qui, par conséquent, peut conditionner la recherche plus rationnelle des agents les plus efficaces de la lutte biologique. Cette recherche doit notamment tenir compte des compatibilités climatiques entre la zone d'origine et la zone d'introduction de l'agent de lutte biologique.

Les études récentes montrent que l'introduction de plantes exotiques peut entraîner des changements d'hôtes chez les espèces autochtones, oligophages ou monophages, qui

deviennent des ravageurs des plantes introduites. On citera le cas de *R. pomonella* qui en Amérique du Nord est passé de l'Aubépine au Pommier (Bush 1994). Il en est de même de *B. fusca* et de *S. calamistis*, tous les deux des espèces de noctuelles et endémiques en Afrique subsaharienne, qui sont devenus des ravageurs majeurs du sorgho cultivé suite vraisemblablement à sa domestication puis du maïs après son introduction sur le continent (Madeira Santos & Ferraz Torrao 1998 ; Chastanet 1998).

Le changement de plantes hôtes est un des facteurs favorisant la spécialisation écologique chez les insectes phytophages. Lorsqu'une population d'une espèce d'insecte phytophage, associée à une plante hôte donnée, se trouvera confrontée à une nouvelle ressource alimentaire certains génotypes, conférant sur cette nouvelle ressource un avantage adaptatif à leurs porteurs, seront sélectionnés. La sélection différentielle de génotypes adaptés à chacune des plantes hôtes conduira à une différenciation génétique de plusieurs populations. L'étude menée par Via *et al*. (2000) sur l'Aphide du pois, *Acyrthosiphon pisum,* qui exploite deux espèces de légumineuses montre que la sélection directe contre les migrants et les hybrides dans chaque milieu parental pourrait favoriser l'évolution ultérieure des races d'hôte avec des choix de plus en plus efficaces d'habitats. Ces auteurs ont, en outre, mis en évidence que la différenciation phénologique des plantes hôtes renforce la ségrégation écologique en favorisant l'isolement reproducteur entre les populations d'une même espèce. De tels phénomènes renforcent les accouplements préférentiels des individus issus d'un même habitat. Les analyses moléculaires de structure génétique des populations ont mis en évidence que certaines espèces d'insectes ont vu leurs populations inféodées aux plantes introduites diverger génétiquement dans des intervalles de temps compris entre plusieurs siècles (*R. pomonella*, *C. serratus*), plusieurs dizaines d'années (*J. haematoloma*) et moins de 20 ans (*Euphydrias editha*). Des exemples similaires avec la formation de races d'hôtes sont connus chez d'autres insectes. C'est le cas la pyrale européenne de maïs, *Ostrinia nubilalis* (Thomas *et al*. 2003) ; de *Spodoptera frugiperda* (Pashley Prowell *et al*. 2004). Il est donc possible, si l'on considère l'ancienneté de la domestication (6.000 ans) du sorgho et de l'introduction (16ème siècle) du maïs en Afrique, que l'agriculture ait pu, au cours des temps historiques, être à l'origine de tels phénomènes de divergence génétique. Si ces derniers sont accompagnés d'un isolement reproducteur, ils pourraient conduire à l'apparition de nouvelles races d'hôte chez les espèces africaines inféodées à ces céréales.

Comme nous l'avons indiqué précédemment, *B. fusca*, signalé auparavant comme polyphage (Polazsek & Khan 1998 ; Haile & Hofsvang 2001), exploite préférentiellement et majoritairement aujourd'hui, suivant les zones écologiques, le maïs ou le sorgho cultivé (Kfir

et al. 2002, Le Rü *et al*. in press). Cela suggère que l'anthropisation des milieux naturels découlant de l'extension et de l'intensification de l'agriculture céréalière a créé des conditions écologiques favorables à une adaptation de *B. fusca* aux plantes cultivées. Cette anthropisation généralisée des milieux naturels n'a pas conduit à un abandon systématique des plantes sauvages par *B. fusca*. Dans certaines régions phytogéographiques, ces plantes semblent encore jouer un rôle significatif dans l'écologie de ce ravageur. Le passage de *B. fusca* sur les plantes cultivées reste par conséquent un phénomène plus ou moins complet selon les régions et en fonction des conditions écologiques et de la dynamique spatiale des populations de cet insecte (Le Rü *et al*. in press). Certains auteurs, au Cameroun notamment, ont signalé une forte abondance *B. fusca* sur *Pennisetum purpureum* (Ndemah *et al*. 2000). Si la préférence comportementale de *B. fusca* pour les plantes cultivées se généralise à travers toute son aire de distribution, ce phénomène représente certainement l'un des facteurs qui pourront conduire à rendre définitif le changement d'hôte. Etant donné qu'il existe encore des populations de *B. fusca* vivant à la fois dans les milieux naturels et anthropisés (cas de l'Afrique Centrale et de l'Erythrée par exemple) (Ndemah *et al*.2000 ; Le Rü *et al*. in press), on peut tenter de mettre en évidence une spécialisation écologique et une différenciation génétique entre les populations de *B. fusca* exploitant les différentes plantes hôtes (entre sauvages et cultivées ou entre sorgho cultivé et maïs). Une étude exhaustive et extensive en Afrique de l'Ouest (Bénin, Togo, Ghana) et dans certains pays d'Afrique de l'Est (Kenya, Tanzanie, Ouganda) montre cependant que *B. fusca,* très rare dans les habitats sauvages, est fortement associé à l'agriculture céréalière (Le Rü *et al*. in press, Sezonlin *et al*. non publié). Pour la présente étude, il n'a donc pas été possible de tester la ségrégation écologique entre les plantes hôtes sauvages et cultivées : nous n'avons pu collecter que très peu d'échantillons de *B. fusca* sur ces plantes sauvages.

Le passage des plantes sauvages aux plantes cultivées est un phénomène récurrent chez les insectes phytophages. La fréquence, la multiplicité et l'indépendance de ces évènements de changement de plantes hôtes restent encore mal élucidées. Les analyses phylogéographiques moléculaires peuvent permettre de comprendre les éventuels effets de la domestication du sorgho et de l'introduction du maïs sur la structure génétique de *B. fusca* tout en mettant en évidence les conditions écologiques et historiques des changements d'hôtes. Cela est rendu possible grâce la datation des phénomènes génétiques et des évènements historiques qui se sont produits au cours des périodes géologiques. Cette datation repose principalement sur l'estimation des distances génétiques corrigées entre populations. Ndemah *et al*. (2001), se basant sur des études écologiques des foreurs africains de tiges,

soulignent que des barrières géographiques telles que les montagnes ou les forêts pourraient faciliter le développement des races géographiques qui se différencieront par leur écologie et leur spécificité pour les plantes hôtes. Si tel est le cas, il serait alors possible d'identifier de telles barrières en analysant la structure génétique des populations à l'aide des marqueurs moléculaires. Ainsi, plusieurs questions fondamentales se dégagent clairement :

Au plan de la génétique des populations, les populations de *B. fusca* sont-elle structurées en fonction de la géographie ou des facteurs écologiques ?

S'il y a une structuration géographique, quels sont les éléments du paysage (structures géomorphologiques : montagnes, rifts, cours d'eau,…) qui en sont à l'origine et quelle en est l'ancienneté ?

Alternativement, s'il y a structuration en fonction des facteurs écologiques, quels facteurs (abiotiques comme climat, altitude ; biotiques) sont en cause ? Parmi les facteurs biotiques, quel rôle ont joué les plantes hôtes, d'une part, les plantes hôtes sauvages – si tant est que l'on puisse les identifier – et d'autre part, les plantes cultivées avec d'un côté le sorgho (plante autochtone) et le maïs (plante introduite et surtout généralisée très récemment) ?

L'agriculture, au travers de ces deux cultures, a-t-elle influé la structure génétique des populations de *B. fusca* ? Quels sont les caractères propres à l'histoire (les types de domestication, les voies et les dates d'introduction) et aux techniques culturales (la diversification ou au contraire l'homogénéisation des variétés cultivées) de ces cultures qui ont modelé ou homogénéisé la structure des populations de ce foreur de tiges ?

Le passage de *B. fusca* des plantes sauvages aux plantes cultivées est un phénomène encore mal connu. Est-il produit une ou plusieurs fois dans des populations indépendantes de *B. fusca* dans plusieurs régions africaines ? Ce changement d'hôte constitue-t-il un événement irréversible et généralisé à travers toutes les régions d'Afrique Subsaharienne ? Peut-on alors identifier les facteurs intrinsèque (traits de vie de l'espèce) ou extrinsèques (facteurs écologiques – concept d'opportunité de niche notamment) ayant favorisé le changement d'hôtes chez *B. fusca* ? L'espèce est-elle vraiment polyphage ou simplement un oligophage ou encore un monophage exploitant accidentellement ou occasionnellement suivant des zones phytogéographiques les plantes hôtes d'une même famille botanique ? Au cas où plusieurs populations de *B. fusca* seraient caractérisées et identifiées à l'aide des outils moléculaires, est-il possible de connaître leur histoire régionale et déterminer leur centre d'origine ?

Des hypothèses ont été avancées pour expliquer les mécanismes biologiques qui font qu'un insecte, originellement inféodé à une ou plusieurs plantes sauvages, devient, à un moment de son histoire évolutive, ravageur de cultures dans une ou plusieurs régions

biogéographiques : on peut citer l'introduction de nouvelles plantes cultivées, l'intensification de l'agriculture et la forte anthropisation des milieux naturels.

Les études de génétique des populations et de phylogéographique d'un ravageur apparaissent alors comme l'un des moyens pouvant apporter des réponses aux questions concernant l'histoire évolutive de ses populations tout en mettant en évidence la nature des facteurs de différenciation de celles-ci. Dans le cas de *B. fusca*, nous avons choisi un marqueur moléculaire, le cytochrome *b*, pour tenter d'apporter des réponses aux questions soulevées ci-dessus.

Chapitre 3
Matériel et méthodes

Section I : Les données biologiques

1.1. Echantillonnage des individus de *B. fusca*

L'échantillonnage des individus de *B. fusca* a été réalisé entre 2001 et 2005 par les différents participants au programme. Les insectes ont été principalement collectés dans les tiges de maïs et de sorgho cultivé au cours des saisons de culture de ces céréales qui correspondent aux saisons des pluies. La collecte des chenilles en diapause a été aussi réalisée au cours des différentes saisons sèches dans les chaumes de sorgho cultivé et de maïs. Les sites de récolte de *B. fusca* ont été choisis dans les zones identifiées comme celles de la culture de ces deux céréales dans des régions écologiques caractéristiques. La figure 3-1 indique la répartition géographique des cinq races principales (*bicolor*, *caudatum*, *dura*, *guinea*, *kafir*) de sorghos cultivés (*S. bicolor*) en Afrique. La race *bicolor* est la plus cultivée. Elle est suivie des races *caudatum* et *guinea* alors que les races *dura* et *kafir* sont très localement cultivées (Deu & Hamon 1994). Le maïs est cultivé dans toute l'Afrique mais la production varie largement suivant les pays et les régions (Chastanet 1998).

En Afrique de l'Ouest, 19 localités du Bénin, Togo, Ghana, Burkina-Faso et Mali, appartenant aux différentes zones phytogéographiques décrites par White (1983), ont été échantillonnées (fig. 3-2a). En Afrique Centrale, représentée par le Cameroun, 17 localités des régions forestières et savanicoles ont été échantillonnées (fig. 3-2b) alors qu'en Afrique de l'Est, 55 localités du Kenya, Ouganda, Rwanda, Tanzanie, Ethiopie et l'Erythrée, appartenant aux différentes zones agroécologiques ou phytogéographiques où ce ravageur est présent, ont été échantillonnées (fig. 3-2c). Enfin en Afrique Australe, la collecte des individus de *B. fusca* est réalisée dans 21 localités comprenant plusieurs zones phytogéographiques de plusieurs pays de cette région (Malawi, Mozambique, Zambie, Zimbabwe, Afrique du Sud) (fig. 3-2d). Toutes ces collectes à travers des différentes grandes régions africaines ont permis d'analyser le génome mitochondrial d'un grand nombre d'individus (590) de *B. fusca*. Ces individus ont été échantillonnés le plus souvent au stade larvaire ou nymphal.

Afin de déterminer les limites géographiques des populations qui pourraient éventuellement être distinguées, et de révéler le maximum de leur diversité moléculaire, les localités d'échantillonnage dans certains pays (Bénin, Ghana, Cameroun, Kenya) sont, en général, séparées les unes des autres d'une distance moyenne d'environ 50 km alors que pour d'autres, elles sont beaucoup plus éloignées (figure 3-3). Nous avons retenu le Bénin (Afrique de l'Ouest), le Cameroun (Afrique Centrale) et le Kenya (Afrique de l'Est) pour étudier la

structure génétique fine des populations et identifier les facteurs de différenciation des populations d'insectes africains.

Dans la perspective des analyses phylogénétiques, une séquence haplotypique de *Busseola phaia* (Lepidoptera, Noctuidae) collecté au Kenya en Afrique de l'Est, a été utilisée comme groupe extérieur. Bien que la systématique des Lépidoptères foreurs de céréales reste encore confuse dans la littérature (Holloway 1998), les études récentes (Moyal *et al*. non publié) ont montré que *B. phaia* échantillonné dans les différentes régions d'Afrique de l'Est est l'espèce sœur de *B. fusca*.

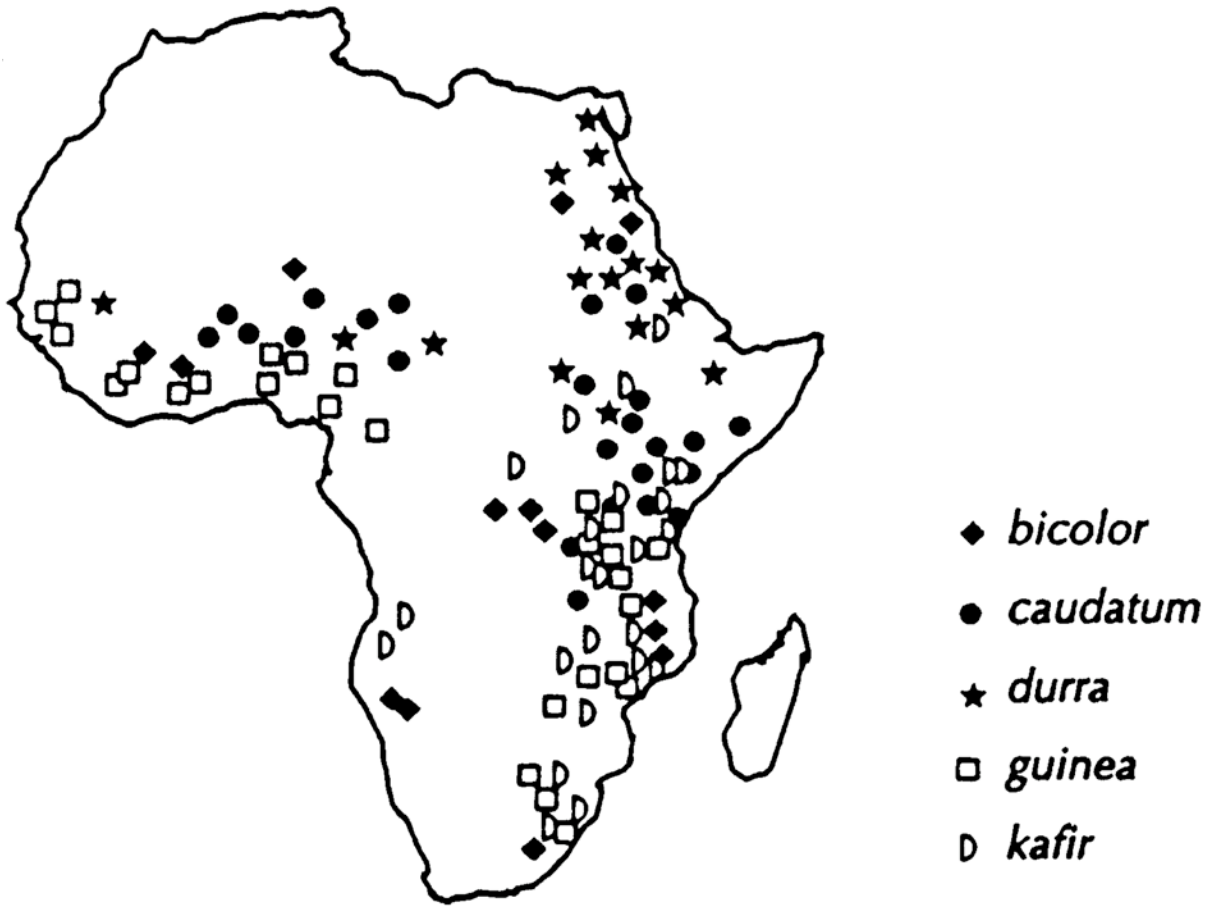


Figure 3-1 : distribution géographique des 5 races principales de sorghos cultivés en Afrique (d'après Deu & Hamon 1994).

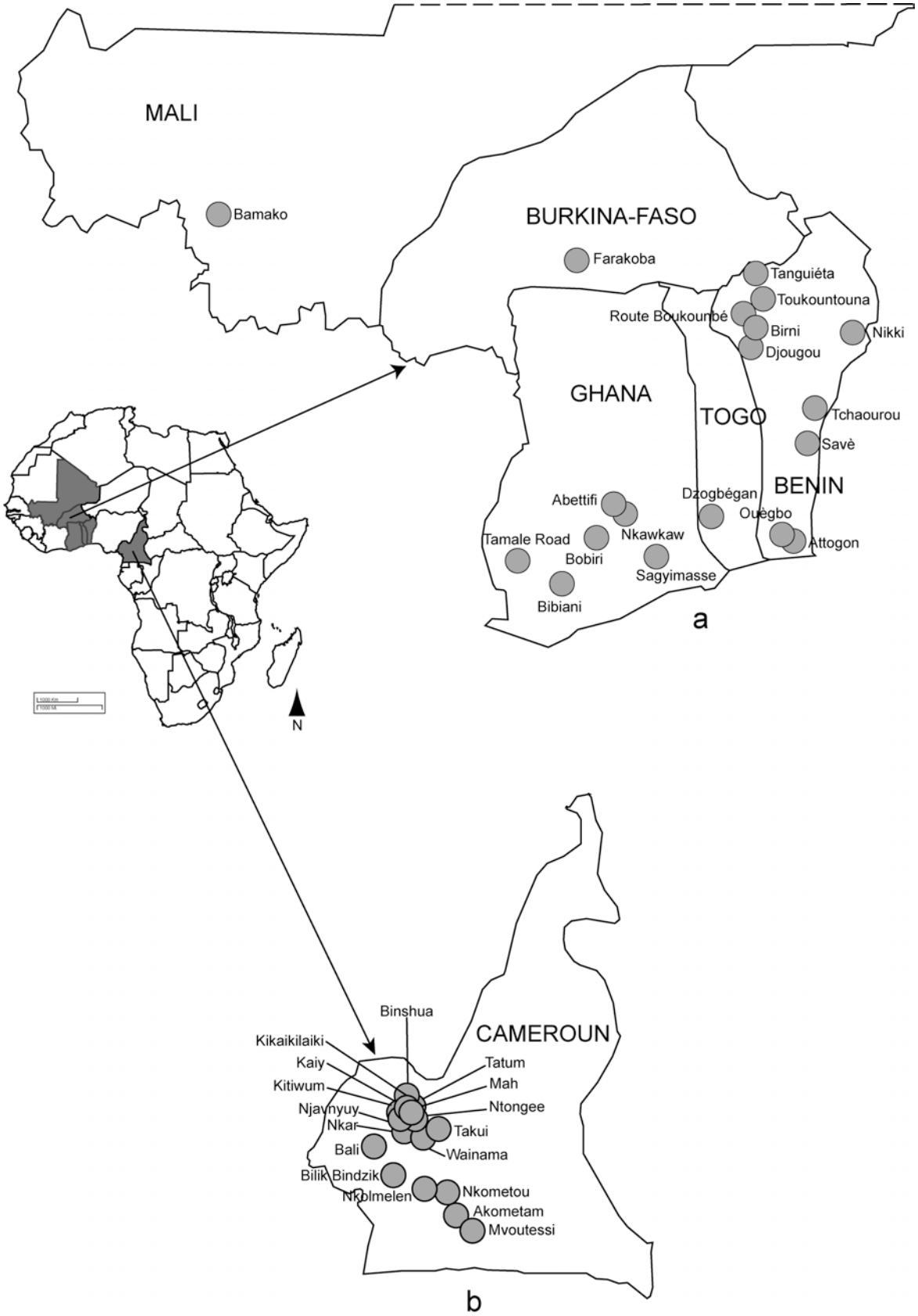


Figure 3-2(a,b) : distribution géographique des localités où *B. fusca* a été échantillonné (a) : Afrique de l'Ouest (Bénin, Togo, Ghana, Mali, Burkina-Faso) ; (b) Afrique Centrale (Cameroun).

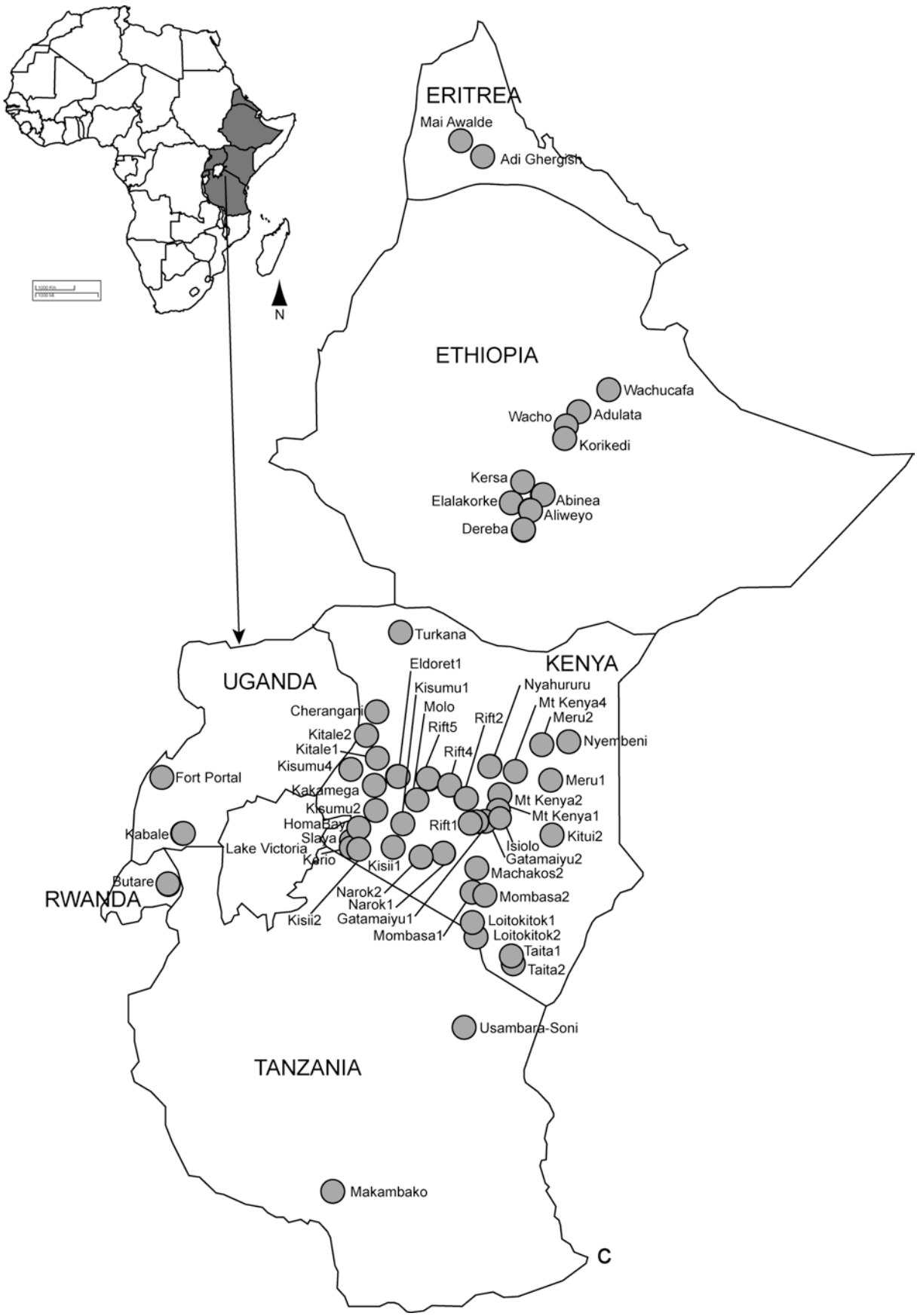


Figure 3-2c : distribution géographique des localités où *B. fusca* a été échantillonné en Afrique de l'Est (Erythrée, Ethiopie, Kenya, Ouganda, Ruanda, Tanzanie).

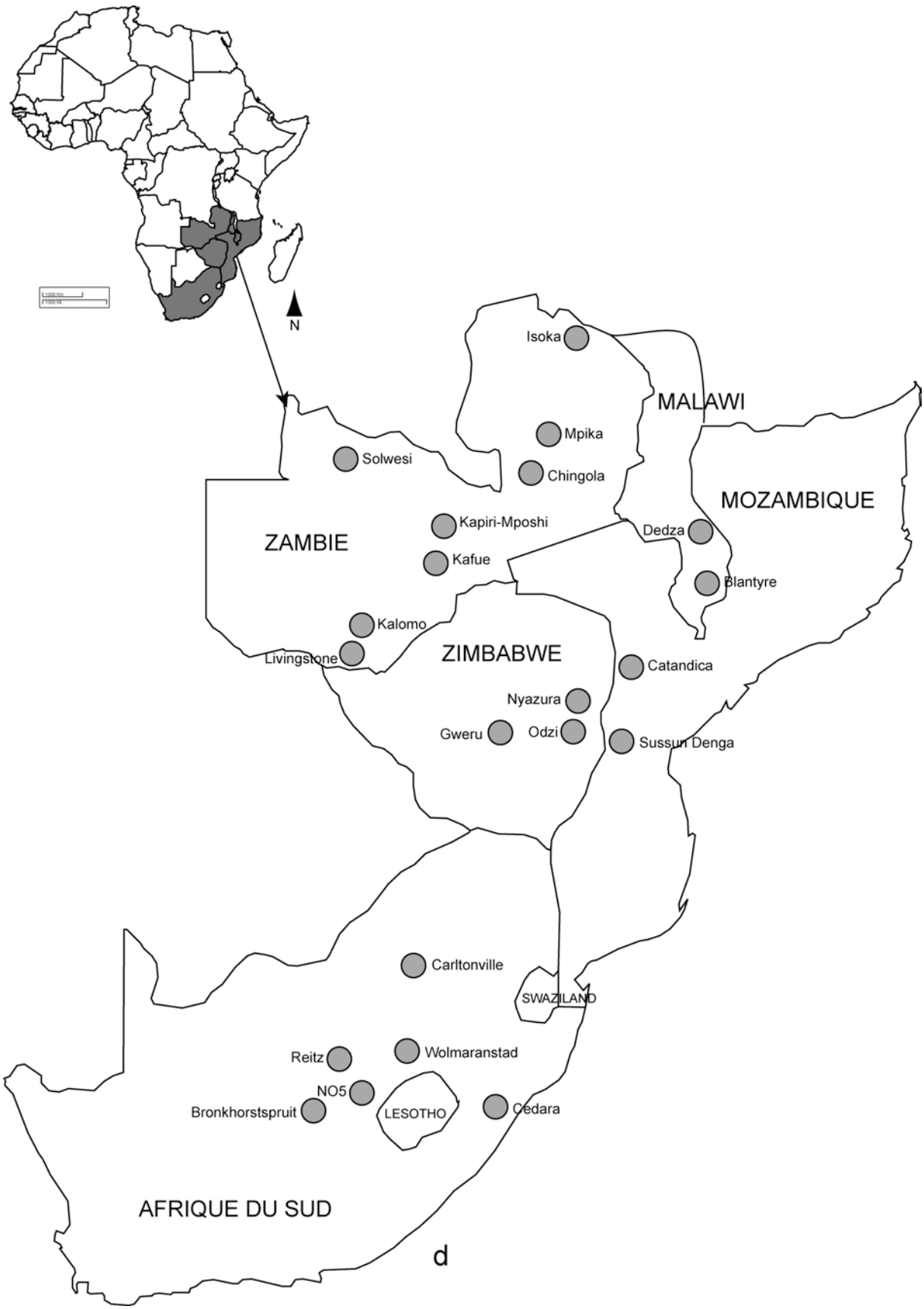


Figure 3-2d : distribution géographique des localités où *B. fusca* a été échantillonné en Afrique Australe (Malawi, Mozambique, Zambie, Zimbabwe, République d'Afrique du Sud).

Figure 3-3 : les pays avec échantillonnage exhaustif de *B. fusca* sont marqués en noir alors que ceux avec un échantillonnage relatif sont marqués en gris dégradé.

1.2. Elevage et conservation des papillons

Les chenilles et chrysalides récoltées étaient ramenées aux différents laboratoires des organismes de recherche associés au programme "foreurs" (International Institut of Tropical Agriculture (IITA), Station de Cotonou pour le Centre et l'Ouest de l'Afrique et International Centre of Insect Physiology and Ecology (ICIPE) au Kenya pour les régions Est et Australe de l'Afrique). Elles ont été élevées jusqu'au stade adulte sur des milieux semi-naturels faits directement de tiges fraîches de maïs et de sorgho cultivé et parfois de sorgho sauvage ou simplement des milieux synthétiques élaborés à partir de broyats de tiges fraîches de ces graminées. Les papillons étaient tués dès l'émergence et conservés dans l'éthanol absolu. Ils étaient finalement ramenés au laboratoire Evolution, Génomes et Spéciations (UPR 9034 - CNRS, Gif-sur-Yvette) en France pour les analyses moléculaires.

Section II : Choix et mise au point des marqueurs et obtention des données moléculaires

2.1. Choix des marqueurs moléculaires

Dans la perspective d'apporter des réponses aux différentes questions énumérées lors de l'exposé de la problématique générale de la thèse, un marqueur mitochondrial (cytochrome *b*) a été retenu. L'objectif est de rechercher des évènements de la paléogéographie africaine ou du développement de l'agriculture céréalière pouvant expliquer l'histoire démographique inférée à partir du marqueur moléculaire.

Avise (2000) conseille particulièrement les marqueurs mitochondriaux pour des études phylogéographiques en s'appuyant sur les nombreuses particularités de ce génome : transmission uniparentale et maternelle, évolution moléculaire rapide, absence de recombinaison.

La vitesse d'évolution de l'ADN mitochondrial chez les insectes a été largement discutée. Elle est sensiblement du même ordre que celle des gènes nucléaires en copie unique, même si, certaines études tendent à montrer qu'il existe un fort taux de substitutions synonymes au sein du génome mitochondrial par rapport au génome nucléaire pour *Drosophila* (Moriyama & Powell 1997). Le taux de substitutions nucléotidiques chez les Lépidoptères est en général de l'ordre 2 % par million d'années (Brower 1994). Le taux de substitution nucléotidique synonyme (qui ne change pas l'acide aminé) est plus élevé que le taux non synonyme (qui change l'acide aminé) (Irwin *et al*. 1991). Le taux de substitution non synonyme est très hétérogène. La vitesse d'évolution n'est alors pas uniforme pour les différentes bases nucléotidiques constituant une séquence moléculaire. Les marqueurs mitochondriaux présentent l'avantage d'exister sous forme haplotypique au sein d'un même individu ce qui facilite leur séquençage direct. Cependant, certaines études ont mis en évidence l'existence de deux ou de plusieurs haplotypes mitochondriaux au sein d'un même individu (Olivo *et al*. 1983 ; Solignac *et al*. 1984 ; Harrison *et al*. 1987 ; Bermingham *et al*. 1986). Ce phénomène connu sous le nom de l'hétéroplasmie provient vraisemblablement plus de mutations mitochondriales que d'une transmission à la fois paternelle et maternelle des molécules mitochondriales (Avise 2000). Il reste de ce fait un événement très rare (Lunt & Hyman 1997) et transitoire (Meusel & Moritz 1993 ; Anderson *et al*. 1995 ; Shitara *et al*. 1998).

Le cytochrome *b* constitue un marqueur mitochondrial de premier choix pour une analyse phylogéographique d'un Lépidoptère ravageur de céréales. Il est plus variable que le COI. Malgré l'hétérogénéité de son taux de substitutions nucléotidiques entre les différents ordres d'insectes, le cytochrome *b* présente une variation intragénérique très élevée dans ce groupe d'organismes (Simmons & Weller 2001).

2.2. Extraction d'ADN génomique

Pour chaque individu de *B. fusca* analysé, la tête, les pattes, les ailes ainsi que l'abdomen sont prélevés et conservés pour d'éventuelles études morphologiques visant à confirmer le statut spécifique ou à identifier une race morphologique ou une espèce nouvelle. L'ADN total est extrait à partir du thorax des papillons en utilisant le Dneasy® Tissue kit (Qiagen GmbH, Germany) approprié. Il s'agit d'un kit qui propose un protocole spécifique pour l'extraction de l'ADN génomique des insectes. L'ADN génomique extrait est récupéré dans un volume total de 100 µl. Les ADN sont conservés au congélateur à –20°C pour les analyses moléculaires.

2.3. Amplification et séquençage d'un fragment du gène de cytochrome *b*

2.3.1. Amplification à la PCR et purification des produits

Un fragment 965 paires de bases (pb) du gène codant le cytochrome *b* a été amplifié avec les amorces CP1 (modifié à partir de Harry *et al.* 1998) : 5'GATGATGAAATTTTGGATC3' et TRs (Simon *et al.* 1994) : 5'TATTTCTTTATTATGTTTTCAAAAC3'. Le programme de PCR utilisé est suivant : une phase de dénaturation initiale de 5 min à 94°C puis 40 cycles de dénaturation (1 min à 94°C), hybridation (1 min 30 sec à 46°C) et élongation (1 min 30 sec à 72°C), suivi d'une phase d'élongation finale de 10 min à 72°C. Chaque réaction de PCR contient : 3 mM $MgCl_2$; 0,4 mM de chaque amorce ; 0,24 mM dNTPs ; environs 100 ng d'ADN plus 2 unités de Taq Promega polymerase pour un volume total de 50µl. Etant donnée que la Taq promega n'est pas des plus fidèles, certains échantillons de *B. fusca* ont été simplement repris jusqu'à l'amplification par PCR de la région cible de leur ADN. Les produits amplifiés sont purifiés à l'aide du kit de purification QIAquick® (Qiagen).

2.3.2. Réaction de séquence et précipitation

Les produits de PCR purifiés sont repris pour la réaction de séquence avec les mêmes amorces ayant servi à l'amplification. Afin de confirmer les séquences nucléotidiques, la réaction est faite dans le deux sens. Pour chaque échantillon, dans un tube eppendorf de 0,5 ml, on ajoute dans l'ordre 2µl de Mix de séquence ; 3µl de tampon de dilution, 1µl de l'amorce ; 1 à 2µl d'ADN purifié le tout complété avec de l'eau distillée pour un volume total de 20µl. Les tubes ainsi préparés sont placés dans l'appareil de PCR et mis en route avec le programme de séquençage décrit. Les produits obtenus à la suite de la réaction de séquence sont précipités et déposés pour le séquençage.

2.3.3. Séquençage et alignement des séquences nucléotidiques

Les produits des réactions de séquence sont analysés avec un séquenceur automatique *ABI prism* 377 puis 3100. Les séquences obtenues sont alignées manuellement avec le logiciel MacClade 4.06 (Maddison & Maddison 2001). Les tableaux (3-1a, 3-1b, 3-1c, 3-1d) présentent les différentes localités échantillonnées, leurs coordonnées géographiques, ainsi que le nombre d'individus analysés pour le gène cytochrome *b*.

Tableau 3-1a : Les différentes localités Ouest africaines échantillonnées, leurs coordonnées géographiques, le nombre d'individus analysés pour le gène du cytochrome *b*.

Pays	Localités échantillonnées	Latitude	Longitude	Individus analysés avec cyt. *b*
Bénin	Attogon	06 42 13 N	02 09 59 E	18
	Ouègbo	06 50 34 N	02 10 13 E	5
	Tchaourou	09 05 34 N	02 33 29 E	10
	Djougou	09 41 41 N	01 37 17 E	4
	Birni	09 56 47 N	01 32 53 E	4
	Route Boukounbé	10 18 29 N	01 21 04 E	22
	Toukountouna	10 34 03 N	01 19 06 E	2
	Tanguiéta	10 40 02 N	01 18 03 E	2
	Savè	08 15 00 N	02 39 03 E	9
	Nikki	09 56 10 N	03 10 37 E	3
Togo	Dzogbégan	07 14 24 N	00 41 38 E	13
Ghana	Sagymasse	06 14 09 N	00 31 20 W	15
	Nkawkaw	06 40 21 N	00 45 08 W	7
	Abettifi	06 48 36 N	00 52 40 W	3
	Bobiri	06 18 00 N	01 22 00 W	1
	Bibiani	06 18 00 N	02 14 33 W	8
	Tamale Road	06 27 40 N	02 54 01 W	1
Burkina-Faso	Farakoba	10 48 23 N	04 35 00 W	11
Mali	Bamako	13 20 10 N	08 14 30 W	3

Tableau 3-1b : Les différentes localités Centre africaines échantillonnées, leurs coordonnées géographiques, le nombre d'individus analysés pour le gène du cytochrome *b*.

Pays	Localités échantillonnées	Latitude	Longitude	Individus analysés avec cyt. *b*
Cameroun	Akometam	03 27 40 N	11 34 20 E	3
	Bali	05 51 38 N	09 58 17 E	2
	Bilik Bindzik	04 18 07 N	11 17 51 E	1
	Binshua	06 32 51 N	10 44 14 E	4
	Kaiy	06 17 09 N	10 44 04 E	18
	Kikaikilaiki	06 16 11 N	10 41 42 E	21
	Kitiwum	06 14 23 N	10 42 27 E	16
	Mah	06 17 29 N	10 43 37 E	11
	Mvoutessi	03 16 17 N	11 47 19 E	3
	Njavnyuy	06 11 37 N	10 40 22 E	2
	Nkar	06 06 13 N	10 38 13 E	2
	Nkolmelen	04 07 43 N	11 31 21 E	1
	Nkometou	04 05 34 N	11 35 54 E	17
	Ntongee	06 16 51 N	10 42 21 E	4
	Takui	06 16 06 N	10 45 07 E	8
	Tatum	06 19 51 N	10 45 41 E	6
	Wainama	06 04 48 N	10 36 25 E	1

Tableau 3-1c : Les différentes localités Est africaines échantillonnées, leurs coordonnées géographiques, le nombre d'individus analysés pour le gène du cytochrome *b*.

Pays	Localités échantillonnées	Latitude	Longitude	Individus analysés avec cyt. *b*
Kenya	Cherangani	01 20 48 N	35 15 14 E	2
	Eldoret 1	00 34 53 N	35 13 33 E	10
	Gatamaiyu 1	01 00 03 S	36 43 01 E	11
	Gatamaiyu 2	01 05 05 S	36 51 38 E	2
	HomaBay	00 27 51 S	34 33 06 E	10
	Isiolo	00 19 34 N	37 33 24 E	2
	Kakamega	00 13 31 N	34 53 48 E	22
	Kerio	00 30 49 N	35 34 34 E	1
	Kisii 1	00 46 13 S	34 58 47 E	5
	Kisii 2	00 54 47 S	34 31 44 E	2
	Kisumu 1	00 21 40 S	35 15 12 E	3
	Kisumu 2	00 10 21 S	34 54 48 E	6
	Kisumu 4	00 35 46 N	34 27 09 E	12
	Kitale 1	00 47 27 N	34 49 30 E	9
	Kitale 2	01 11 44 N	34 49 06 E	13
	Kitui 2	01 19 29 S	38 03 41 E	11
	Loitokitok 1	02 43 06 S	37 31 10 E	2
	Loitokitok 2	02 55 46 S	37 30 59 E	11
	Machakos 2	01 29 21 S	37 16 36 E	12
	Meru 1	00 23 02 S	37 36 38 E	4
	Meru 2	00 08 20 N	37 28 17 E	6
	Molo	00 10 09 S	35 37 38 E	5
	Mombasa 1	01 50 09 S	37 15 20 E	5
	Mombasa 2	02 05 27 S	37 29 23 E	1
	Mt Kenya 1	00 55 47 S	37 09 20 E	7
	Mt Kenya 2	00 43 02 S	37 16 02 E	8
	Mt Kenya 4	00 02 50 S	37 03 22 E	2
	Narok 1	01 05 32 S	36 07 01 E	2
	Narok 2	01 00 38 S	35 38 28 E	6
	Nyahururu	00 04 44 S	36 34 16 E	9
	Nyembeni	00 14 39 N	37 54 25 E	6
	Rift 1	00 57 26 S	36 33 28 E	3
	Rift 2	00 39 53 S	36 23 09 E	4
	Rift 4	00 19 24 S	36 09 02 E	3
	Rift 5	00 06 33 N	35 54 01 E	5
	Slaya	00 08 05 N	34 19 00 E	1
	Taita 1	03 23 37 S	38 20 20 E	3
	Taita 2	03 26 17 S	38 21 57 E	3
	Turkana	03 32 17 N	35 51 38 E	1
Erythrée	Mai Awalde	15 98 24 N	38 27 47 E	11
	Adi Ghergish	15 68 21 N	38 64 27 E	1
Ethiopie	Abinea	07 57 19 N	39 17 06 E	1
	Adulata	08 28 42 N	39 17 40 E	1
	Aliweyo	07 24 28 N	38 40 23 E	5
	Dereba	07 10 06 N	38 32 53 E	3
	Elalakorke	07 16 54 N	38 39 07 E	3
	Kersa	07 19 42 N	38 39 37 E	2
	Korikedi	08 06 37 N	38 43 12 E	1
	Wacho	08 07 48 N	38 48 16 E	2
	Wachucafa	08 35 14 N	39 21 06 E	1
Ouganda	Fort Portal	00 37 48 N	30 14 56 E	6
	Kabale	01 08 23 S	30 08 41 E	7
Rwanda	Butare	02 39 37 S	29 44 02 E	4
Tanzanie	Usambara-Soni	04 51 24 S	38 20 40 E	3
	Makambako	08 50 44 S	34 19 42 E	4

Tableau 3-1d : Les différentes localités Sud africaines échantillonnées, leurs coordonnées géographiques, le nombre d'individus analysés pour le gène du cytochrome *b*.

Pays	Localités échantillonnées	Latitude	Longitude	Individus analysés avec cyt. *b*
Malawi	Blantyre	15 55 06 S	34 56 57 E	7
	Dedza	14 13 12 S	34 08 22 E	2
Mozambique	Catandica	17 40 02 S	33 15 20 E	4
	Sussun Denga	19 29 49 S	33 16 19 E	3
Zambie	Chingola	12 23 46 S	27 22 02 E	2
	Isoka	09 46 58 S	32 44 39 E	2
	Kafue	15 40 36 S	28 14 05 E	2
	Kalomo	17 01 43 S	26 28 23 E	1
	Kapiri Mposhi	13 50 12 S	28 38 58 E	2
	Livingstone	17 38 03 S	25 56 05 E	2
	Mpika	11 32 17 S	31 35 45 E	2
	Solwesi	13 05 57 S	25 54 02 E	1
Zimbabwe	Gweru	19 20 47 S	30 11 15 E	3
	Nyazura	18 43 46 S	32 09 53 E	2
	Odzi	18 51 58 S	32 15 42 E	1
Afrique du Sud	Bronkhorstspruit	25 44 42 S	28 44 40 E	1
	Carltonville	26 20 00 S	27 22 00 E	4
	Cedara	29 37 00 S	30 20 00 E	1
	NO5	26 44 01 S	27 04 22 E	3
	Reitz	27 47 10 S	28 28 43 E	1
	Wolmaranstad	26 44 35 S	28 23 08 E	1

Section III : Les méthodes d'études en génétique des populations et en phylogéographie / Analyse des données moléculaires fournies par le séquençage du cytochrome *b*

Les méthodes utilisées regroupent (1) les méthodes classiques qui comprennent les méthodes d'études de la phylogénie des haplotypes et les méthodes hiérarchiques qui traitent de la structure génétique, de la diversité moléculaire et de la démographie des populations ; (2) les méthodes synthétiques qui comprennent l'analyse phylogéographique des clades emboîtés (NCPA).

3.1. Reconstruction de la phylogénie des haplotypes

Les relations phylogénétiques entre haplotypes ont été évaluées à l'aide des méthodes de parcimonie et de neighbour-joining.

Les analyses de parcimonie (MP) ont été réalisées en procédant à une recherche heuristique avec les options "best trees only", "stepwise addition", "random addition sequence" (1000 réplicats). La robustesse des topologies de MP a été estimée par les procédures de bootstrap en utilisant 1000 répétitions (full heuristic search) de 10 réplications aléatoires de "stepwise addition" à chaque fois pour toutes les analyses. Le résultat est présenté sous forme de l'arbre consensus majoritaire (50% majority-rule consensus tree) ou de consensus strict.

Les distances utilisées pour les analyses en Neighbor-Joinging ont été calculées par le maximum de vraisemblance sous PAUP en utilisant les paramètres des modèles de substitutions nucléotidiques sélectionnés par MODELTEST. La version 3.06 du MODELTEST a été utilisée pour sélectionner les modèles de substitutions nucléotidiques qui décrivent au mieux le jeu de données de *B. fusca*. Les valeurs de bootstrap de cette analyse sont obtenues à la suite de 1000 réplications.

3.2. Structure et diversité génétique des populations de *B. fusca*

L'analyse moléculaire de la variance (AMOVA) et de la diversité moléculaire a été réalisée, telles que décrites par Excoffier *et al*. (1992), avec le programme Arlequin (version 2.00). La structure génétique des populations de *B. fusca* a été explorée en calculant les différents indices de la variance moléculaire (Φ_{ST}, Φ_{CS} et Φ_{CT}). Les hypothèses de structuration géographique testées, à l'échelle régionale, sont issues des résultats des études de la biogéographie portant sur des espèces d'herbivores africains associés au peuplement graminéen (Arctander *et al*. 1999 ; Flagstad *et al*. 2001 ; Alpers *et al*. 2004). L'Afrique Centrale, de l'Est et Australe sont groupées et séparées de l'Afrique de l'Ouest. Pour d'autres analyses, la structure génétique a été testée en tenant compte des résultats de l'analyse phylogénétique.

Dans certaines régions à structures phytogéographiques bien définies, et pour lesquelles nous disposions d'un échantillonnage important des individus de *B. fusca*, des regroupements des sites d'échantillonnage ont été faits sur la base des zones phytogéographiques ou écologiques, et lorsque cela était possible sur la base des barrières géographiques naturelles afin d'étudier la structure génétique fine des populations de ce foreur. Etant donnée l'existence d'une hétérogénéité des formations végétales en Afrique, plusieurs cartes phytogéographiques ont été utilisées afin de mieux décrire la structure génétique de *B. fusca*. En Afrique de l'Ouest, le regroupement des sites d'échantillonnage a

été définie à l'aide de la carte phytogéographique de White (1983). Dans cette région, trois zones concernent notre aire d'étude : forêt ombrophile planitiaire guinéo-congolaise (types relativement secs), mosaïque de forêt ombrophile planitiaire guinéo-congolaise et de formation herbeuse secondaire, forêt claire soudanienne avec abondance d'*Isoberlinia*. Au Cameroun, où les populations ont été regroupées sur la base des groupes phytogéographiques identifiés par Mayaux *et al.* (1999) en Afrique Centrale, deux zones agro-écologiques couvrent notre aire d'étude : forêt dense sempervirente (type relativement humide); forêt claire et savane boisée. Au Kenya où l'échantillonnage a pu être réalisé selon un maillage relativement serré, des regroupements ont été définis en fonction de la Rift Valley, une dépression naturelle sèche, et des zones agro-écologiques d'Hassan (1998). Cet auteur identifie cinq zones agro-écologiques au Kenya : zone sèche d'altitude moyenne, zone de transition, zone humide d'altitude moyenne, zone humide de transition, zone de hautes montagnes tropicales. Tous les regroupements possibles de ces cinq zones agro-écologiques ont été testés afin d'évaluer les différentes valeurs de Φ_{CS} et Φ_{CT}. Des indices de diversité haplotypique et nucléotidique ont été calculés à plusieurs niveaux de regroupements afin d'évaluer le degré de polymorphisme des populations et d'inférer l'histoire démographique des clades (Grant & Bowen 1998) indépendamment de la longueur du fragment et de la taille de l'échantillon (Nei & Li 1979 ; Nei 1987).

3.3. Démographie des populations de *B. fusca*

L'inférence de l'histoire démographique des clades identifiés à la suite des analyses phylogénétiques est essentielle pour déterminer les facteurs ayant modelé l'évolution des populations. Dans le cas de *B. fusca*, la disparité de distribution (mismatch distribution) des différences observées entre les paires de séquences nucléotidiques (distribution of pairwise differences) a été analysée (Slatkin & Hudson 1991; Schneider & Excoffier 1999). La distribution attendue est unimodale pour les populations en expansion démographique alors qu'elle est multimodale pour les populations à l'équilibre (Slatkin & Hudson 1991). Nous avons calculé l'indice D de Tajima. Cet indice donne des informations sur la démographie des populations avec une valeur négative dans le cas d'une expansion démographique et une valeur positive dans les cas de populations subdivisées ou en équilibre démographique (Tajima 1989a, b). Tous ces paramètres démographiques ont été estimés avec le logiciel Arlequin 2.00 (Schneider *et al.* 2000).

3.4. Les méthodes synthétiques : la NCPA

Utilité et pouvoir de résolution : l'analyse phylogéographique des clades emboîtés (NCPA ou Nested Clade Phylogeographic Analysis) s'est avérée très utile en phylogéographie comme en témoignent les résultats obtenus par son application (Bernatchez 2001, Alpers *et al.* 2004). Dans plusieurs études portant sur des espèces animales ou végétales, la NCPA a été utilisée pour inférer l'histoire évolutive ayant produit les patrons géographiques observés actuellement entre les populations ou les espèces (Arctander *et al*. 1999 ; Pitra *et al*. 2002). Néanmoins, Knowles & Maddison (2002) font remarquer l'absence de test statistique pour valider les différents scénarios évolutifs inférés. Cela représente une des faiblesses de cette méthode. Ces critiques méthodologiques ont conduit Templeton (2004) à réévaluer la fiabilité et la validité de la NCPA en l'appliquant au cas d'espèces dont l'histoire évolutive est *a priori* connue. Les résultats de ce test méthodologique ont conduit cet auteur à conclure que la NCPA déduit des inférences correctes dans la majorité des cas.

Applications aux données mitochondriales (cyt. b*) de* B. fusca : l'analyse phylogéographique des clades emboîtés a été réalisée telle qu'elle est décrite par Templeton (1998, 2004). Les relations généalogiques sont représentées par un réseau parcimonieux d'haplotypes qui définit une série de clades emboîtés. Les probabilités de connections haplotypiques étaient calculées selon la théorie de la coalescence en utilisant successivement le logiciel TCS 1.17 et TCS 1.21 (Clement *et al*. 2000). Les branches ayant des probabilités au-dessus du seuil parcimonieux (0,95) sont retenues pour constituer le réseau. L'hypothèse d'une distribution géographique aléatoire est testée en réalisant des tests de permutation pour chaque clade et ses sous-clades. Les analyses statistiques sur les distances géographiques intra- et interclades ont été calculées en utilisant le logiciel GeoDis 2.1 (Posada *et al*. 2000). Les coordonnées GPS des localités échantillonnées ont été utilisées au cours de ces analyses. Les valeurs significatives des distances géographiques *Dc* (distance géographique moyenne entre individus portant les haplotypes dans un groupe donné et le centre géographique de ce groupe particulier), *Dn* (distance géographique moyenne entre individus portant les haplotypes d'un groupe particulier et le centre géographique du groupe de niveau supérieur qui le contient), *(I-T) Dc* (distance géographique entre clade intérieur et clade apical pour les *Dc*) et *(I-T) Dn* (distance géographique entre clade intérieur et clade apical pour les *Dn*) ($p<0,05$) sont interprétées avec la clé d'inférence proposée par Templeton en 2004. Cette inférence permet d'identifier les évènements récents ou anciens ayant affecté la structure des populations, mais aussi, de déterminer les patrons principaux de l'histoire de ces populations.

Chapitre 4
Résultats et discussion

Section I : Les articles publiés

1.1 Article 1

Phylogeography and population genetics of the maize stalk borer *Bussseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa

M. SEZONLIN, S. DUPAS, B. LE RÜ, P. LE GALL, P. MOYAL, P-A. CALATAYUD, I. GIFFARD, N. FAURE and J-F. SILVAIN

Molecular Ecology (2006) 15, 407-420

Phylogéographie et génétique des populations du foreur de tiges de maïs *Busseola fusca* (Lepidoptera, Noctuidae) en Afrique Subsaharienne

Ce premier article correspond à l'exploitation des données acquises au cours des deux premières années de mon travail de thèse. Il a permis de définir une démarche analytique s'appuyant sur l'analyse des travaux similaires récents et de formuler des hypothèses explicatives pour les résultats observés.

Contexte de l'étude : la génétique des populations et la phylogéographie des insectes africains et, en particulier, des espèces phytophages restent très peu connues. Pourtant, de telles analyses chez les insectes phytophages peuvent permettre de comprendre les facteurs qui modèlent l'adaptation et la structuration des populations d'une espèce donnée.

Objectifs de l'étude : il s'agit (1) d'analyser la structure génétique et la phylogéographie des populations de *B. fusca* présentes sur maïs et sorgho cultivé dans différents biotopes de plusieurs pays d'Afrique de l'Ouest et de l'Est et (2) de déterminer quels sont les facteurs historiques et démographiques qui permettent d'expliquer cette structure génétique.

Méthodes utilisées : l'échantillonnage des individus de *B. fusca* sur le sorgho cultivé et le maïs a été réalisé au Bénin, Togo, Ghana, Mali, Cameroun et Kenya. Un marqueur moléculaire présentant une variabilité intraspécifique *a priori* suffisante a été choisi : un fragment de 1kb du gène codant pour le cytochrome *b* a été séquencé. Les individus séquencés ont été pris au hasard dans chaque population afin que toutes les localités échantillonnées soient représentées. L'analyse des données moléculaires est faite en combinant les différentes méthodes utilisées en phylogéographie : les méthodes traditionnelles (généalogiques et hiérarchiques) et la méthode synthétique (analyse phylogéographique des clades emboîtés) (voir détails dans le chapitre 3).

Principaux résultats et points discutés : les analyses phylogénétiques et celles des clades emboîtés séparent les populations de *B. fusca* en trois clades majeurs : un clade regroupant les haplotypes observés en Afrique de l'Ouest et deux clades - Kenya I et Kenya II - rassemblant

les haplotypes de l'Est et du Centre de l'Afrique. Le clade Kenya I localisé sur une partie de l'Afrique Orientale est partiellement sympatrique avec le clade Kenya II qui présente une large distribution géographique. La similitude des niveaux de divergence haplotypique et nucléotidique entre et au sein de ces clades suggère qu'ils se sont isolés à la même période (environ un million d'années) dans trois aires géographiques différentes en Afrique Subsaharienne. L'analyse de mismatch distribution suggère que ces trois clades ont bénéficié d'une expansion démographique. L'analyse moléculaire de la variance indique des différenciations géographiques à différents niveaux (populationnel, régional). L'analyse phylogéographique des clades emboîtés suggère que la distribution des haplotypes à plusieurs niveaux hiérarchiques au sein des trois clades majeurs peut s'interpréter dans plusieurs cas par des hypothèses de restriction des flux géniques avec isolement par la distance. La présence d'une structuration génétique ancienne impliquant de forts niveaux de divergence entre unités géographiques montre que la domestication du sorgho et l'introduction de maïs en Afrique n'ont eu aucun effet sur le patron géographique observé chez *B. fusca* en ce qui concerne du moins le génome mitochondrial.

Molecular Ecology (2006) 15, 407–420 doi: 10.1111/j.1365-294X.2005.02761.x

Phylogeography and population genetics of the maize stalk borer *Busseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa

M. SEZONLIN,*† S. DUPAS,* B. LE RÜ,‡ P. LE GALL,*† P. MOYAL,* P.-A. CALATAYUD,‡ I. GIFFARD,* N. FAURE* and J.-F. SILVAIN*

*Unité de Recherche IRD 072, CNRS, Laboratoire Populations, Génétique et Evolution, BP1, 91198 Gif-sur-Yvette cedex, France, †Unité de Recherche IRD 072, IITA, Cotonou, Bénin, ‡Unité de Recherche IRD 072, ICIPE, Nairobi, Kenya

Abstract

The population genetics and phylogeography of African phytophagous insects have received little attention. Some, such as the maize stalk borer *Busseola fusca*, display significant geographic differences in ecological preferences that may be congruent with patterns of molecular variation. To test this, we collected 307 individuals of this species from maize and cultivated sorghum at 52 localities in West, Central and East Africa during the growing season. For all collected individuals, we sequenced a fragment of the mitochondrial cytochrome *b*. We tested hypotheses concerning the history and demographic structure of this species. Phylogenetic analyses and nested clade phylogeographic analyses (NCPA) separated the populations into three mitochondrial clades, one from West Africa, and two — Kenya I and Kenya II — from East and Central Africa. The similar nucleotide divergence between clades and nucleotide diversity within clades suggest that they became isolated at about the same time in three different refuges in sub-Saharan Africa and have similar demographic histories. The results of mismatch distribution analyses were consistent with the demographic expansion of these clades. Analysis of molecular variance (AMOVA) indicated a high level of geographic differentiation at different hierarchical levels. NCPA suggested that the observed distribution of haplotypes at several hierarchical levels within the three major clades is best accounted for by restricted gene flow with isolation by distance. The domestication of sorghum and the introduction of maize in Africa had no visible effect on the geographic patterns observed in the *B. fusca* mitochondrial genome.

Keywords: African biogeography, *Busseola fusca*, phylogeography, Pleistocene, population genetics

Received 14 February 2005; revision received 17 June 2005; accepted 1 September 2005

Introduction

Several species of noctuid moths behave as larval stem borers of cultivated plants. *Busseola fusca*, *Sesamia calamistis*, and *Sesamia nonagrioides* are major pests on maize (*Zea mays* L.) and cultivated sorghum [*Sorghum bicolor* (L.) Moench] throughout sub-Saharan Africa (reviewed in Kfir *et al.* 2002). Their larvae damage cereal crops, reducing yields. Some of these species, such as *B. fusca*, are endemic to sub-Saharan Africa. *B. fusca* known as the maize stalk borer displays ecological preferences that differ between regions. This pest is more adapted to middle- and high-altitude conditions in East Africa, where it is abundant in the highlands (above 1500 m) and is rarely reported in the lowlands (Kfir *et al.* 2002). In Cameroon (Central Africa), it is abundant from high- (above 1000 m) to mid-altitude (700 m) regions (Ndemah *et al.* 2001). However, in West Africa, it is generally more abundant at lower altitudes than in other regions. Moyal (1998) reported that this insect was more frequent in Guineo-Congolian rainforest (dry type) (White 1983) areas than in savannah, whereas other studies reported that it is more frequent in dry savannah than in forests (Kfir *et al.* 2002). The pest status of *B. fusca* on cereal crops also varies between regions. It is the most important pest in the highlands of East and Southern Africa (Nye 1960; Sithole

Correspondence: Stéphane Dupas, Fax: +33169070421; E-mail: dupas@pge.cnrs-gif.fr

© 2006 Blackwell Publishing Ltd

1989; Ndemah *et al.* 2001; Kfir *et al.* 2002) whereas in West Africa, it generally has a lower economic impact (Schultess *et al.* 1997) except in sorghum-growing areas where it is the most important stem borer (Dakouo & Ratnadass 1997; Ajayi 1998). In Cameroon, it was shown to be the main pest across years, seasons and ecological zones (Ndemah *et al.* 2001). Locally, for each region, the population density of *B. fusca* varies according to season and year. Climatic factors, and biotic factors such as human activity, availability of wild and cultivated plants and natural enemies, are responsible for the population fluctuations from year to year (Dakouo & Ratnadass 1997; Moyal 1998; Ndemah *et al.* 2000). *B. fusca* has been reported to attack a wide range of host plants from three monocotyledonous families: Cyperaceae, Poaceae and Typhaceae (Khan *et al.* 1997; Polaszek & Khan 1998; Haile & Hofsvang 2001). However, a 2-year exhaustive survey carried out across East Africa (Kenya, Tanzania, Uganda) indicates that *B. fusca* is uncommon in the wild habitat, showing on the contrary that it is specifically adapted to maize and sorghum crops (Le Rü, unpublished). The domestication of sorghum probably began some 5000 years ago in northeast Africa (Doggett 1988; Murty & Renard 2001) whereas maize was introduced more recently, at the end of the 16th century (Chastanet 1998; Madeira Santos & Ferraz Torrao 1998). Since its rapid expansion at the beginning of the 20th century, maize has become one of the most important crops in sub-Saharan Africa (Ristanovic 2001). Several insect species switched from wild plants to cultivated crops, and became pests. The timing of these events, and whether there were multiple or independent host switches is unknown. Molecular phylogeographic analysis can help to understand the possible effects of sorghum domestication and maize introduction on *B. fusca* genetic structure and can provide insights on the ecological and historical conditions of the switch.

Apart from domestication, the genetic structure of *B. fusca* is also likely to be influenced by its evolutionary history on wild host plants before domestication. Very little is known about the population genetics and phylogeography of African insects, despite their diversity and, for some of them, their economic importance (Kuchta & Meyer 2001). For example, only 10 of the 246 articles published in *Molecular Ecology*, in the 'Phylogeography, Speciation and Hybridization' section between January 2001 and June 2004 dealt with African animal species — vertebrates in each case. The ecological and economic importance of African phytophagous insects makes them particularly suitable models for studies of the biogeography and demographic history of African insects. Ndemah *et al.* (2001) suggested that geographic barriers, such as mountains or forests, might facilitate the development of specific races differing in climatic requirements and host-plant specificities. If this is the case, then it should be possible to identify these natural barriers by analysing the genetic structure of molecular markers. An understanding of the ecological and genetic diversity of insect pests would make it possible to develop and to improve monitoring and biological control strategies. Two main questions need to be answered: Do the different ecological preferences of *B. fusca* from different regions correspond to molecular genetic differences? Has the ecological differentiation of populations of *B. fusca* from different regions depended on geography (West vs. East Africa), palaeogeographic history, or biotopes occupied (high- or low-altitude biotopes, wet or dry biotopes, savannah or forests)? The aims of this study were to (i) analyse the genetic structure and phylogeography of the *B. fusca* populations present on maize and cultivated sorghum in various countries and biotopes from West, Central and East Africa, and (ii) determine the historical, demographic and ecological factors shaping the genetic diversity of this species. We studied the mitochondrial cytochrome *b* gene, because it is informative at the intrageneric level in Lepidoptera (Simmons & Weller 2001).

Several approaches must be combined to infer the demographic history of a species from sequence data. Some of these approaches are used to describe the genetic structure, whereas others such as nested clade phylogeographic analyses (NCPA) are used to develop and test hypotheses concerning the evolutionary history responsible for generating the observed structure (Bernatchez 2001). This latter methodological approach is widely used in phylogeography (Althoff & Pellmyr 2002; Duran *et al.* 2004). Knowles & Maddison (2002) pointed out the lack of statistical tests to distinguish between inferences. However, Templeton (2004) re-estimated the reliability and the validity of this method by applying the NCPA methodology to cases in which there were clear a priori expectations. He concluded that NCPA makes the correct inferences in the vast majority of case. We therefore combined descriptive methods, statistical methods, and NCPA to investigate the historical factors of genetic differentiation of *B. fusca* populations.

Materials and methods

Insect DNA sampling

Busseola fusca larvae were collected from maize and sorghum plants during the wet seasons of 2001–2003. The larvae and pupae sampled were reared to adult stage in the laboratory on host plant stems. Moths were preserved in absolute alcohol immediately after emergence. Three hundred seven individuals from sub-Saharan Africa were analysed: 85 individuals from 13 localities in Benin, Togo, Ghana and Mali in West Africa; 19 individuals from three localities in Cameroon in Central Africa; and 203 individuals from 36 localities in five different agroecological zones in Kenya, East Africa (Hassan *et al.* 1998) (Fig. 1a–d). Table 1 provides information about the localities, latitude and longitude and the individuals of *B. fusca* analysed for each locality.

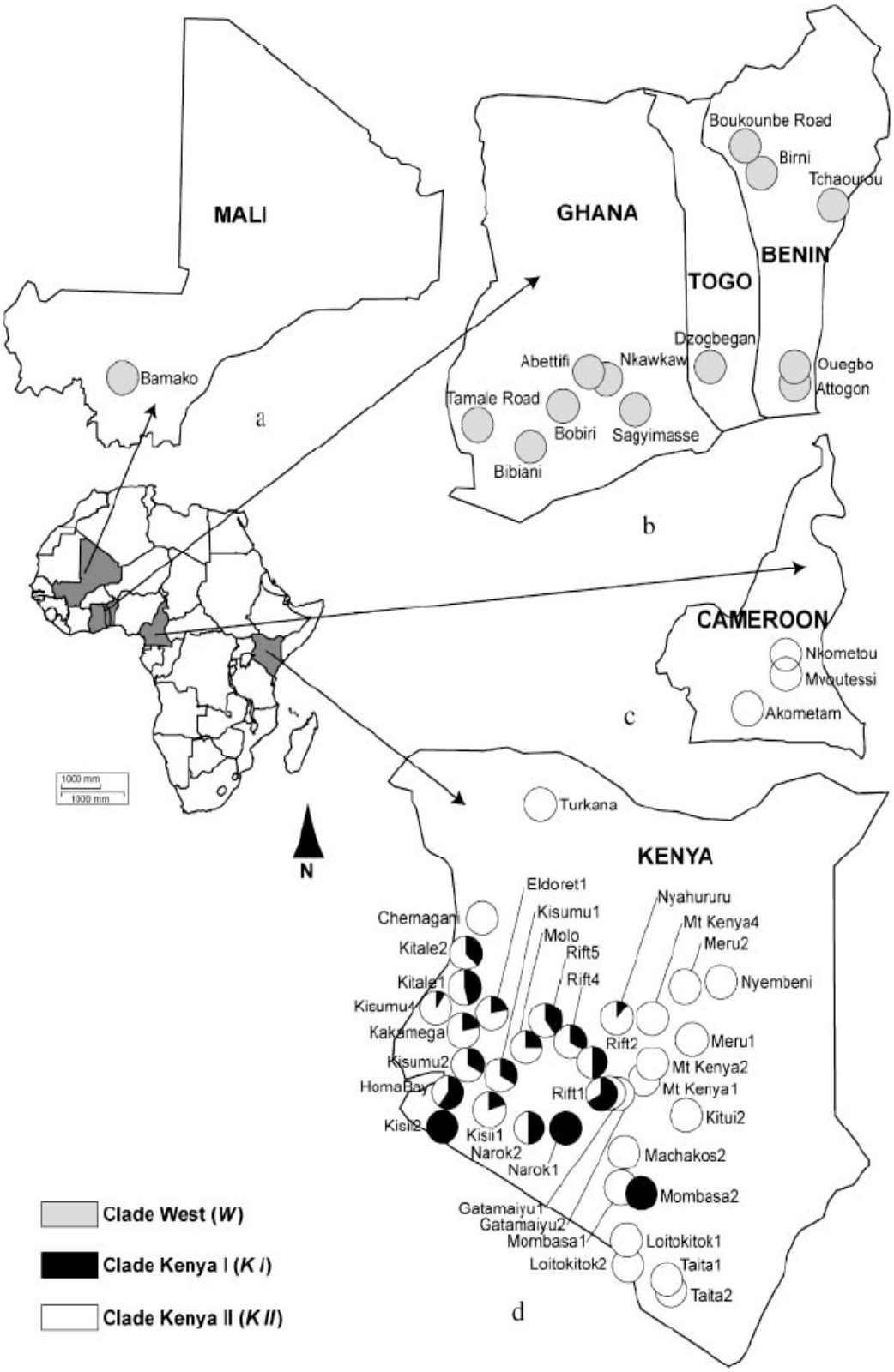


Fig. 1 (a) Geographic distribution of localities where *Busseola fusca* was sampled and the percentage of each clade in Mali (West Africa). (b) Geographic distribution of localities where *B. fusca* was sampled and the percentage of each clade in Benin, Togo and Ghana (West Africa). (c) Geographic distribution of localities where *B. fusca* was sampled and the percentage of each clade in Cameroon (Central Africa). (d) Geographic distribution of localities where *B. fusca* was sampled and the percentage of each clade in Kenya (East Africa).

Table 1 List of sampled localities with geographic coordinates and number of *Busseola fusca* individuals sequenced

Country/ Locality sampled	Latitude	Longitude	Individuals sequenced
Benin			
Attogon	06°42′13″ N	02°09′59″ E	9
Ouegbo	06°50′34″ N	02°10′13″ E	5
Tchaourou	09°05′34″ N	02°33′29″ E	9
Birni	09°56′47″ N	01°32′53″ E	3
Boukounbe Road	10°18′29″ N	01°21′04″ E	12
Togo			
Dzogbegan	07°14′24″ N	00°41′38″ E	13
Ghana			
Sagyimasse	06°14′09″ N	00°31′20″ W	11
Nkawkaw	06°40′21″ N	00°45′08″ W	7
Abettifi	06°48′36″ N	00°52′40″ W	3
Bobiri	06°42′12″ N	01°22′00″ W	1
Bibiani	06°18′00″ N	02°14′33″ W	8
Tamale Road	06°27′40″ N	02°54′01″ W	1
Mali			
Bamako	13°20′10″ N	08°14′30″ W	3
Cameroon			
Nkometou	04°10′50″ N	11°50′01″ E	12
Mvoutessi	03°40′03″ N	11°53′40″ E	4
Akometou	02°55′40″ N	11°10′15″ E	3
Kenya			
Cherangani	01°20′48″ N	35°15′14″ E	2
Eldoret 1	00°34′53″ N	35°13′33″ E	9
Gatamaiyu 1	01°00′03″ S	36°43′01″ E	9
Gatamaiyu 2	01°05′05″ S	36°51′38″ E	1
HomaBay	00°27′51″ S	34°33′06″ E	10
Kakamega	00°13′31″ N	34°53′48″ E	23
Kisii 1	00°46′13″ S	34°58′47″ E	5
Kisii 2	00°54′47″ S	34°31′44″ E	2
Kisumu 1	00°21′40″ S	35°15′12″ E	3
Kisumu 2	00°10′21″ S	34°54′48″ E	6
Kisumu 4	00°35′46″ N	34°27′09″ E	12
Kitale 1	00°47′27″ N	34°49′30″ E	8
Kitale 2	01°11′44″ N	34°49′06″ E	13
Kitui 2	01°19′29″ S	38°03′41″ E	5
Loitokitok 1	02°43′06″ S	37°31′10″ E	2
Loitokitok 2	02°55′46″ S	37°30′59″ E	9
Machakos 2	01°29′21″ S	37°16′36″ E	10
Meru 1	00°23′02″ S	37°36′38″ E	3
Meru 2	00°08′20″ N	37°28′17″ E	6
Molo	00°10′09″ S	35°37′38″ E	4
Mombasa 1	01°50′09″ S	37°15′20″ E	5
Mombasa 2	02°05′27″ S	37°29′23″ E	1
Mt Kenya 1	00°55′47″ S	37°09′20″ E	6
Mt Kenya 2	00°43′02″ S	37°16′02″ E	5
Mt Kenya 4	00°02′50″ S	37°03′22″ E	2
Narok 1	01°05′32″ S	36°07′01″ E	2
Narok 2	01°00′38″ S	35°38′28″ E	6
Nyahururu	00°04′44″ S	36°34′16″ E	9
Nyembeni	00°14′39″ N	37°54′25″ E	5
Rift 1	00°57′26″ S	36°33′28″ E	3
Rift 2	00°39′53″ S	36°23′09″ E	2
Rift 4	00°19′24″ S	36°09′02″ E	3
Rift 5	00°06′33″ N	35°54′01″ E	5
Taita 1	03°23′37″ S	38°20′20″ E	3
Taita 2	03°26′17″ S	38°21′57″ E	3
Turkana	03°32′17″ N	35°51′38″ E	1

Busseola phaia (Lepidoptera, Noctuidae) has been chosen as the outgroup based on recent unpublished work. Although the systematics of African lepidopteran stem borers is still rather confused (Holloway 1998), our studies (Moyal *et al.*, unpublished) show that *B. phaia*, collected from various regions of East Africa, is the sister species of *B. fusca*.

Molecular analyses

Total DNA was extracted from insect thoraxes, using the DNeasy Tissue Kit (QIAGEN GmbH). A 1000-bp fragment of the gene encoding cytochrome *b* was amplified with the primers CP1 (modified from Harry *et al.* 1998): 5′-GATGATGAAATTTTGGATC-3′ and TRs (Simon *et al.* 1994): 5′-TATTTCTTTATTATGTTTTCAAAAC-3′. Polymerase chain reaction (PCR) cycling conditions were as follows: initial denaturation for 5 min at 94 °C; 40 cycles of 94 °C for 1 min, 46 °C for 1 min 30 s, 72 °C for 1 min 30 s and a final extension period of 10 min at 72 °C. The reaction mixture contained 3 mM $MgCl_2$, 0.4 μM primers, 0.24 μM dNTPs, and 100 ng of DNA plus 2 U of Promega *Taq* polymerase per 50 μL of reaction mixture. Amplified products were purified with the QIAquick purification kit (QIAGEN) and sequenced on an automated sequencer using the amplification primers in both directions. The consensus sequences obtained were aligned manually, using MACCLADE 4.05 (Maddison & Maddison 2002).

Haplotype phylogeny

Phylogenetic relationships were estimated by means of maximum parsimony (MP) and neighbour joining (NJ) using maximum-likelihood distances. MP analyses were performed using heuristic search strategy starting with stepwise addition trees replicated 10 times, using a random input order of sequences to get the initial tree for each replicate. Robustness of MP topologies was assessed by bootstrap with 1000 replicates (full heuristic search) of 10 random stepwise addition replicates each, for all analyses. MODELTEST version 3.07 (Posada & Crandall 1998) was used to select the substitution model that best describes the data. An NJ analysis of the ML distances obtained using the parameter estimates derived from MODELTEST was performed. Bootstrap values for this analysis were obtained from 1000 replications.

In order to estimate divergence time between clades, corrected pairwise genetic distances between haplotypes were calculated using the parameters inferred from the nucleotide substitution model selected by MODELTEST. All phylogenetic analyses were performed with PAUP* version 4.0b10 (Swofford 2002).

Population genetic structure and diversity indices

An analysis of molecular variance (AMOVA) was performed as described by Excoffier *et al.* (1992). The population structure of *B. fusca* mitochondrial DNA was explored by calculating the different fixation indices (Φ_{CT} and Φ_{SC}). At the regional scale, Central and East Africa were grouped and separated from West Africa according to biogeographic relationships highlighted by studies on the African herbivore species that are associated with gramineous plant communities (Arctander *et al.* 1999; Flagstad *et al.* 2001; Alpers *et al.* 2004). Within each region, localities were grouped based on hypothetical biogeographic barriers (the effect of East Africa Rift Valley in Kenya) or according to ecological categories. In West Africa, for which ecological categories were defined based on White (1983) phytogeographic map, three zones concerned our study area: Sudanian woodland with abundant *Isoberlinia*, mosaic of Guineo-Congolian lowland rainforest and secondary grassland, and Guineo-Congolian rainforest (drier types). In East Africa, ecological categories were based on Hassan *et al.* (1998) agro-ecological zones. Five zones were used for our study, Dry-mid-altitude, Dry-transitional, Moist-mid-altitude, Moist-transitional, Highland-tropical. All possible regrouping of these five zones were tested for their effect on Φ_{SC} and Φ_{CT} values.

Haplotypic (h) and nucleotide (π) diversities have been used to estimate the level of polymorphism and to infer the demographic history of each clade (Grant & Bowen 1998) independently of fragment length and sample size (Nei & Li 1979; Nei 1987).

AMOVA and diversity index calculations were performed with the ARLEQUIN 2.000 software (Schneider *et al.* 2000).

Demographic history of populations

We investigated the demographic history of the clades identified in the phylogenetic analyses. The distribution of pairwise differences between individual sequences was analysed by means of mismatch distribution analysis (Slatkin & Hudson 1991; Schneider & Excoffier 1999). A unimodal distribution would be expected for populations in expansion and a multimodal distribution for populations at equilibrium (Slatkin & Hudson 1991). The raggedness index of the observed distribution (r) representing the modality of the distribution, and the sum of square deviation from the mismatch expected from a model of sudden population expansion (SSD) were calculated. Since the nucleotide substitution model selected by hierarchical likelihood-ratio tests (hLRTs) and Akaike information criterion (AIC) (HKY + I + G) was not available in the ARLEQUIN 2.000 software, the r and SSD indices were calculated by using pairwise difference. The significance of these statistics was tested as implemented in ARLEQUIN. We calculated Tajima's D index. This index can provide information about demographic history, with demographic expansion leading to negative values, and subdivided populations at equilibrium leading to positive values (Tajima

1989a, b). All demographic parameters were calculated with the ARLEQUIN 2.000 software (Schneider *et al.* 2000).

Nested clade phylogeographic analysis

NCPA was performed as described by Templeton (1998, 2004). The genealogic relationships were represented through a haplotype parsimony network to define a series of nested clades. The hypothesis of random geographic distributions was tested through permutation tests for each clade and subclade components and the causes of the significant geographic patterns observed were inferred by means of an inference key. The probabilities of haplotype connections were calculated using TCS 1.17 software (Clement *et al.* 2000) and the network with probabilities above the parsimony limit (0.95) was selected. Statistical analyses of geographic distances were carried out with GEODIS 2.1 (Posada *et al.* 2000). GPS coordinates of all the sampling localities were used for the analyses. The geographic distances between centres of distributions of clades were tested for significance in permutation tests, within clade (D_c), with nested clade centre (D_n) or between interior and tip [(I-T)D_c and (I-T)D_n] at each level. Significant values were interpreted in terms of population history, using the latest inference key from Templeton (2004).

Results

Phylogenetic reconstruction

A fragment of 965 bp encoding cytochrome *b* was sequenced from 307 individuals of *Busseola fusca* from West, Central and East Africa. We observed 70 different haplotypes (GenBank Accession nos AY769536–AY769605). One hundred eleven nucleotide sites were variable (11.50%) and 50 were informative in parsimony analysis (5.18%). According to both hLRTs and AIC, the HKY + I + G model of evolution (Hasegawa *et al.* 1985; Yang 1993; Gu *et al.* 1995) was selected by MODELTEST version 3.07. The parameters estimated from this model were: transition/transversion (ti/tv) ratio = 13.21; Pinvar = 0.6981; gamma shape parameter = 0.9657; nucleotide base frequencies: A = 0.34; C = 0.14; G = 0.10; T = 0.42. As commonly observed in insect mitochondria, the nucleotide sequenced had a high A + T content (A + T = 76%).

Parsimony analyses generated six equiparsimonious trees (length = 162, CI = 0.698, RI = 0.931). Each of these trees was divided into same three clades: a clade grouping sequences from the West African region only (*W*), a Kenya I clade (*KI*) and a Kenya II clade (*KII*), which also contained sequences from Cameroon. Discrepancies between these six equiparsimonious topologies concerned only the apical nodes. It was therefore possible to construct a strict consensus of the most parsimonious trees (Fig. 2). NJ tree obtained using ML distances was similar to topology derived from MP analyses.

The three conspicuous clades of individuals and haplotypes are supported by bootstrap values exceeding 50% both for MP and NJ analyses. The smallest clade *KI* comprised 8 haplotypes and 43 individuals, all of which came from Kenya. The clade *KII* comprised 39 haplotypes and 179 individuals and had the largest distribution from East to Central Africa. Finally, the clade *W* comprised 23 haplotypes and 85 individuals and was found only in West Africa. No haplotype was shared between West African populations and East–Central African populations. Both in MP and NJ analyses, *W* and *KI* were supported by high bootstrap values whereas *KII* was supported by lower bootstrap values. The phylogenetic relationships between these three major clades remain unclear. The sister group status of clades *W* and *KI* was observed only in NJ analysis and supported by low bootstrap values.

The corrected mean pairwise genetic distances between haplotypes within and between clades are shown in Table 3.

Genetic structure and diversity indices of the B. fusca *populations*

Significant genetic structure was observed at various hierarchical levels (Table 2). Most of the molecular variation was accounted for by the differentiation between large population units corresponding to biogeographic units, West Africa on one hand and Eastern Africa + Cameroon on the other hand ($\Phi_{CT} = 0.708$, $P < 10^{-3}$). Significant genetic differentiation was observed among the populations within these two biogeographic units ($\Phi_{CS} = 0.360$, $P < 10^{-3}$). Within Kenya, most of the variation was accounted for by the Rift Valley ($\Phi_{CT} = 0.397$, $P < 10^{-3}$), suggesting that this geographic feature constitutes a strong natural barrier to gene flow. Other arrangements based on Hassan's agroclimatic zones in Kenya were tested. The differentiation among these agroclimatic zones was significant ($\Phi_{CT} = 0.106$, $P = 0.026 \pm 0.005$) but this value remained low compared to effect of the East Africa Rift Valley. The highest Φ_{CT} value among combinations of agroclimatic zones (0.197, $P = 0.005 \pm 0.002$) was obtained by grouping dry vs. others agro-ecological zones. In West Africa, no genetic structure was observed between White's (1983) phytogeographic zones. For all hierarchical structures tested in West and East Africa, the genetic differentiation between localities (Φ_{SC}) was significant within group (Table 2).

The three conspicuous clades are separated by similar mean distances and displayed similar nucleotide and haplotype diversity indexes (Table 3).

Demographic history of B. fusca

In the absence of a suitable intrinsic calibration we can apply the approximate 2% per million years divergence rate for insect mtDNA of Brower (1994). While this provides

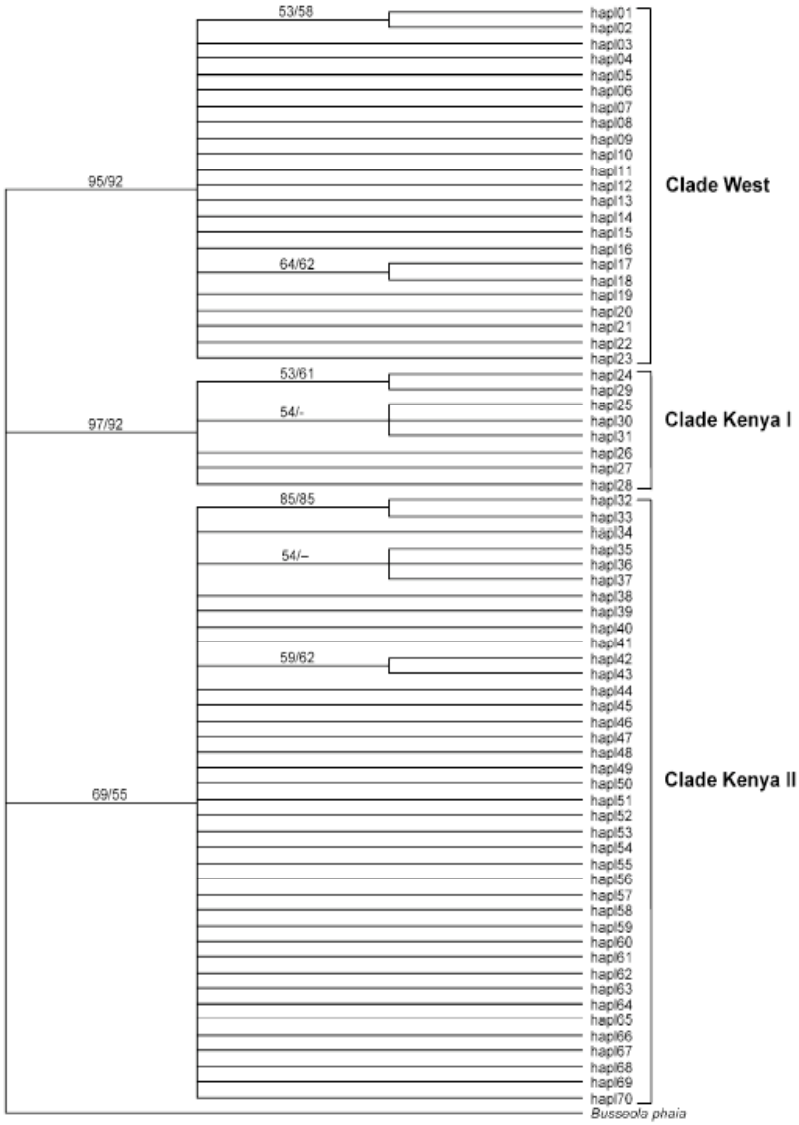


Fig. 2 Strict consensus of the most parsimonious trees. Bootstrap support of > 50% in both MP (first number) and NJ (second number) searches in 1000 replicates are given for the relevant nodes. *Busseola phaia* was used as the outgroup taxon.

Table 2 Values of Φ_{SC} and Φ_{CT} for different hierarchical levels and associated probabilities

Hierarchical level	Φ_{SC}	Φ_{CT}
➢ Major biogeographic zones (East + Central Africa) vs. (West Africa)	0.360 ($P < 10^{-3}$)	0.708 ($P < 10^{-3}$)
➢ East Africa:		
– Rift Valley Barrier	0.104 ($P < 10^{-3}$)	0.397 ($P < 10^{-3}$)
– Hassan's Agroclimatic zones (1, 2, 3, 4, 5)*	0.250 ($P < 10^{-3}$)	0.106 ($P = 0.026 \pm 0.005$)
– Hassan's Agroclimatic zones grouped Dry vs. Others (1 + 2),(3 + 4 + 5)*	0.264 ($P < 10^{-3}$)	0.197 ($P = 0.005 \pm 0.002$)
➢ West Africa:		
– White's phytogeographic regions†	0.443 ($P < 10^{-3}$)	0.000 ($P = 0.332 \pm 0.013$)

Legend: values of Φ_{SC} (localities within groups) and Φ_{CT} (groups) for different hierarchical levels, and associated probabilities.
*Hassan's agroclimatic zones in Kenya; 1: Dry-mid-altitude; 2: Dry-transitional; 3: Moist-mid-altitude; 4: Moist-transitional; 5: Highland-tropical.
†White's phytogeographic regions in West Africa (numbers are from White's map). #27: Sudanian woodland with abundant *Isoberlinia*; #11a: Mosaic of Guineo-Congolian lowland rainforest and secondary grassland; #2: Guineo-Congolian rainforest (drier types).

Table 3 Corrected mean pairwise genetic distance between haplotypes (*p* dist.), haplotype diversity (*h*) and nucleotide diversity (π)

Clades	*p* dist. (%)	Haplotype diversity	Nucleotide diversity (%)
W	0.3775 ± 0.1253	0.9062 ± 0.0166	0.2714 ± 0.1624
KI	0.3228 ± 0.1195	0.8106 ± 0.0426	0.2111 ± 0.1343
KII	0.4888 ± 0.2243	0.8276 ± 0.0260	0.3838 ± 0.2156
W – *KI*	2.7957 ± 0.2223		
W – *KII*	3.1235 ± 0.2387		
KI – *KII*	2.5660 ± 0.2466		

estimates for the placing of divergence dates amongst our clades, it must be recognized that the error associated with these is unknown. The percentages of nucleotide divergence observed between the three clades (2.5–3.1%) suggest they have been separated for about a million years (Myr). Within each of the three clades, the mean distance between haplotypes suggests that their mean divergence time is around 200 000 years. The three clades display strong haplotypic diversity (between 0.81 and 0.91) associated with low levels of nucleotide diversity (between 0.21 and 0.381). This accumulation of haplotypes suggests that the clades experienced bottlenecks at their origins, followed by major population demographic expansion (Grant & Bowen 1998; Avise 2000). Both the variance (SSD) and raggedness index (*r*) tests suggested that the curves (Fig. 3a–c) do not significantly differ from the distribution under a model of population expansion ($P_{SSD} = 0.15$ and $P_r = 0.41$ for *W*; $P_{SSD} = 0.69$ and $P_r = 0.75$ for *KI*; $P_{SSD} = 0.81$ and $P_r = 0.88$ for *KII*). Similarly, the negative values obtained for Tajima's *D* index for each clade (–1.30536; –0.9992; –0.84178 for clades *W*, *KI* and *KII*, respectively) are all consistent with this hypothesis of population expansion since the origin of the clades.

Nested clade phylogeographic analysis

The NCPA network calculation defined the same three clades revealed by MP and NJ analyses. With the parsimony threshold of 0.95, these three clades were not linked (Fig. 4a–c). The West African clade contained the 23 haplotypes observed exclusively in this region and five hypothetical intermediate haplotypes, hierarchically grouped into 10 'one-step clades', 4 'two-step clades', and 2 'three-step clades'. The Kenya I clade contained the eight observed haplotypes and three hypothetical haplotypes: 4 'one-step clades' and 2 'two-step clades'. Finally, the Kenya II clade, with 39 observed haplotypes and only four hypothetical is organized into 11 'one-step' clades and 4 'two-step clades'.

Nested contingency analysis on the haplotype network revealed significant geographic associations in the three major networks at all clade levels. These significant values were interpreted using Templeton's (2004) interpretation key (Table 4). Seven of the eight clades displaying geographic associations were interpreted by restricted gene flow with isolation by distance, and one by contiguous range expansion.

Discussion

All three phylogenetic analyses separate *Busseola fusca* into three major clades corresponding to three geographic units: one originating from the West African region (*W*), one restricted to East Africa (*KI*), and one found from Central to East Africa (*KII*). Geographic overlap was observed only between the two East African clades. Applying a molecular clock of 2% substitution per Myr for insects, these different clades would have expanded from populations that experienced bottlenecks around 1 million years ago (Ma). This is long before sorghum domestication and maize introduction to Africa. These bottlenecks occurred in ancient periods

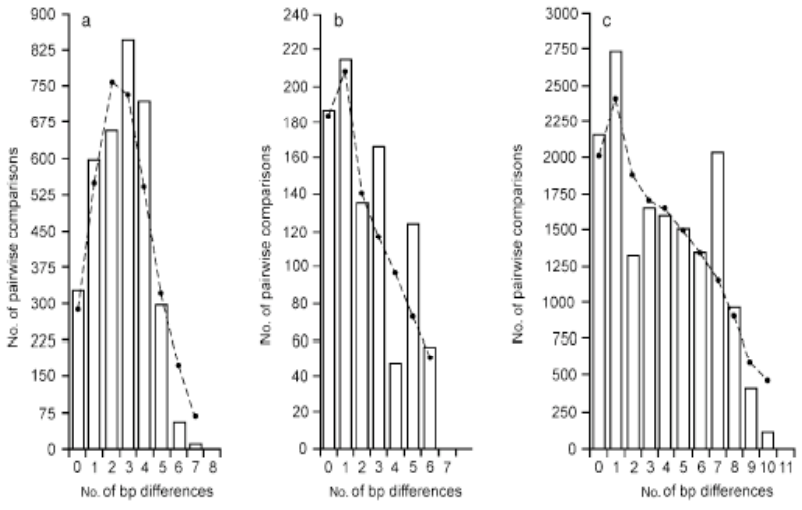


Fig. 3 Mismatch distribution analysis showing histogram of observed mismatch frequencies and best-fit curve of the sudden expansion model. (a) West Africa population; (b) Kenya I population; (c) Kenya II population.

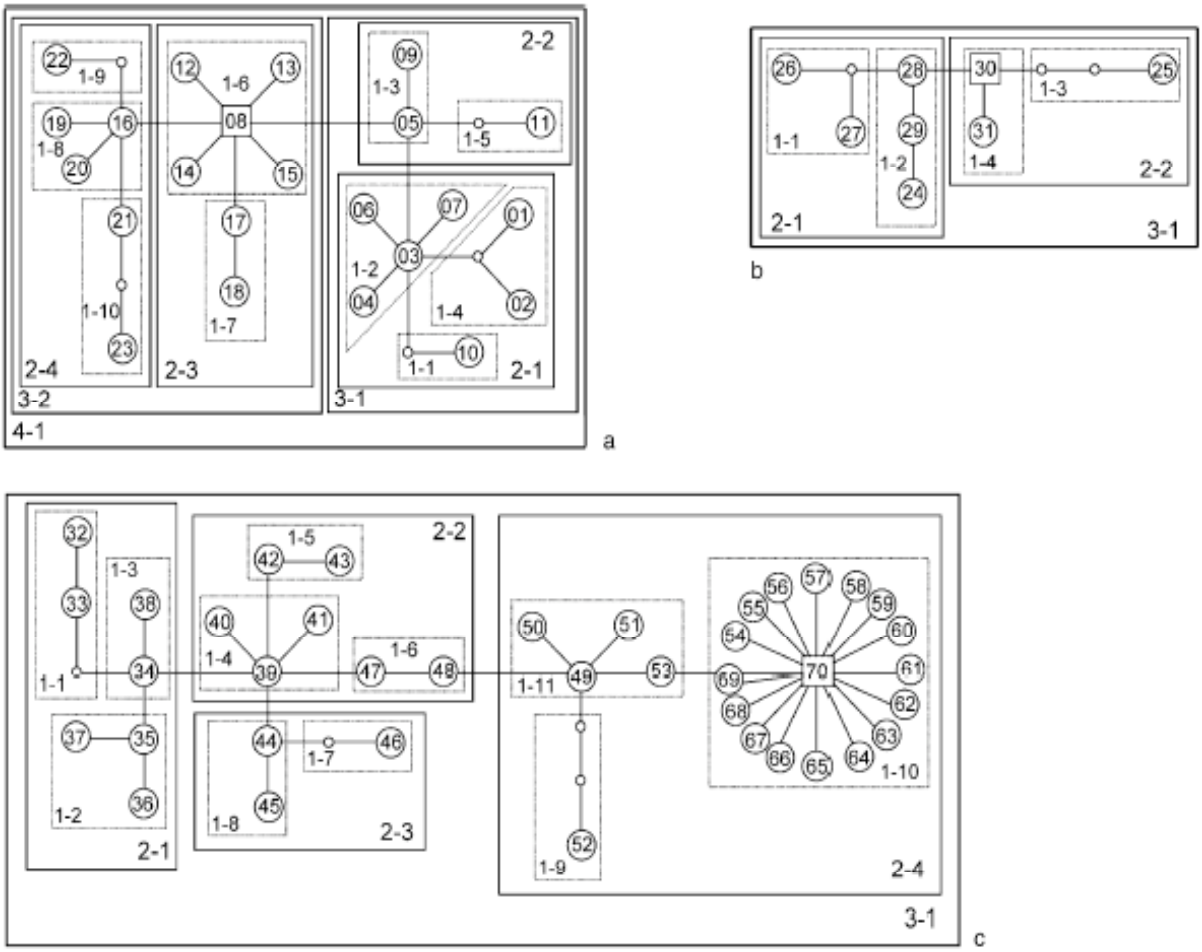


Fig. 4 Haplotype network of all haplotypes detected for *Busseola fusca*. Each haplotype is labelled by its number. Haplotypes not sampled or extinct are designated by small circles. (a) West Africa population; (b) Kenya I population; (c) Kenya II population.

probably due to climatic and/or geological events. Major climatic changes occurred in sub-Saharan Africa during the Pleistocene. The period of climatic instability started 3.3–2.45 Ma, oscillating between hot/humid and cooler, drier periods (Wagner 2002). A shift to arid, open conditions occurred in near 2.8 Ma, 1.7 Ma, and 1.0 Ma (deMenocal 1995). deMenocal (1995) concluded that this alternation of cold, dry periods and warmer, wetter periods led to oscillations in savannah biotope expansion. Considering the approximate nature of molecular clock in insects, and the absence of fossil data, the three major clades may have started their divergence in three bottlenecked populations isolated at one period in the Pleistocene. In comparison to other lepidopteran stem borer pest species, *B. fusca* is less abundant at lower altitude and warmer conditions (Kfir *et al.* 2002). This is consistent with a hypothesis of reduction of its geographic range, isolation of small populations and differentiation in high altitude refuges during a period of warmer conditions. Such climatic changes may also have provided opportunities for some populations of *B. fusca*, as clade *W*, to adapt to warmer conditions. The genetic structure of the *B. fusca* populations resembles the patterns observed for some mammalian herbivores (Alpers *et al.* 2004). For example, the phylogeographic patterns of African mammals are due to alternation between periods of population isolation and genetic differentiation in one or several refuge areas during warmer climatic conditions and periods of population expansion and genetic exchange during colder and drier climatic conditions (Arctander *et al.* 1999; Flagstad *et al.* 2001; Van Hooft *et al.* 2002). Mean divergence age observed for rodent populations from the African genus *Sylvisorex* are also consistent with early Pleistocene differentiation. Increased aridity during this period was regarded as the main factor responsible for population divergence of this group (Quérouil *et al.* 2003). The East Africa region has been described as a mosaic of secondary refuge zones for herbivorous mammals, with periodic exchanges through temporary contact bridges in the East Africa Rift Valley (Livingstone 1982; Arctander *et al.* 1999). Although the East African populations of *B. fusca* are now overlapping, the

Table 4 Result of nested clade analysis for clades in which a significant geographic association was detected. Haplotypes/clades shaded in grey are interior. RGF-IBD: restricted gene flow with isolation by distance, CRE: contiguous range expansion. D_c, average distance of individuals from the clade's geographic centre. D_n, the average distance of individuals from the geographic centre of all members of the nested clade. $(I\text{-}T)_c$, the average distance between interior and tip clades within a given clade, and $(I\text{-}T)_n$, the average distance between interior and tip clades in the nested clade. Superscript L or S denotes significantly large distances and significantly small distances for $P < 0.05$, respectively.

Clade 1-2 (*W*)

Haplotypes	3	4	6	7
D_c	136.7	0	191.3	36.5
D_n	356.8^L	273.4	298.2	111.0^S
$(I\text{-}T)_c$	64.6			
$(I\text{-}T)_n$	185.5^L			

Steps in inference key 1, 2, 11, 17, 4 NO : RGF-IBD

Clade 1-6 (*W*)

Haplotypes	8	12	13	14	15
D_c	181.8	0	5.2	7.7^S	0
D_n	243.6^L	309.6	124.8^S	128.6^S	162.5
$(I\text{-}T)_c$	175.9^L				
$(I\text{-}T)_n$	99.2^L				

Steps in inference key 1, 2, 3, 4 NO : RGF-IBD

Clade 3-2 (*W*)

Clades	2-3	2-4
D_c	192.8	113.2^S
D_n	206.4^L	148.1^S
$(I\text{-}T)_c$	79.5^L	
$(I\text{-}T)_n$	58.2^L	

Steps in inference key 1, 2, 3, 4 NO : RGF-IBD

Clade 2-1 (*KI*)

Clades	1-1	1-2
D_c	35.1^S	108.3
D_n	147.2	107.8
$(I\text{-}T)_c$	73.2^L	
$(I\text{-}T)_n$	−37.4	

Steps in inference key 1, 2, 3, 4 NO : RGF-IBD

Clade 1-3 (*KII*)

Haplotypes	34	38
D_c	115.7	36.1^S
D_n	231.9^L	51.2
$(I\text{-}T)_c$	79.7	
$(I\text{-}T)_n$	180.6^L	

Steps in inference key 1, 2, 3, 4 NO : RGF-IBD

Clade 1-10 (*KII*)

Haplotypes	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70
D_c	0^S	0	0	149.2	149.5	66.9	0	45.2	0	0	0	0	0^S	0	0	0	1063.4^L
D_n	686.7	407.4	472.2	520.2	401.9^S	424.1^S	424.7	496.5	424.7	233.1^S	503.3	472.2	482.7	559.3	2398.2	503.3	976.2^L
$(I\text{-}T)_c$	1022.1^L																
$(I\text{-}T)_n$	419.2^L																

Steps in inference key 1, 2, 3, 4 NO : RGF-IBD

Clade 2-1 (*KII*)

Clades	1-1	1-2	1-3
D_c	33	124.8	80.7
D_n	113.7^S	152.8	104.6
$(I\text{-}T)_c$	−3.2		
$(I\text{-}T)_n$	−30.8^L		

Steps in inference key 1, 2, 11, 17, 4 NO : RGF-IBD

Clade 2-4 (*KII*)

Clades	1-9	1-10	1-11
D_c	0	846.2	120.6
D_n	222.9	825.6^L	273.2^S
$(I\text{-}T)_c$	−716.7^S		
$(I\text{-}T)_n$	−546^S		

Steps in inference key 1, 2, 11, 12 NO : CRE

© 2006 Blackwell Publishing Ltd, *Molecular Ecology*, 15, 407–420

Rift Valley explains most of the molecular variation in Kenya and seems to have worked as an important natural barrier to maintain population structure. The Rift Valley have arisen and expanded from 2.9 to 1.6 Ma (Zeitoun 2000). It works as a natural barrier reinforcing population structure in many species of vertebrates (Arctander *et al.* 1999; Pitra *et al.* 2002). Our results on *B. fusca* suggest this natural barrier plays a similar role on reinforcement of population structure within this species, which in our case were probably initially isolated in climatic refuges. In addition, biogeographic distinction has been observed between the vertebrate faunas of West Africa and Cameroon/East Africa (Alpers *et al.* 2004) and this pattern is what we found for *B. fusca* populations. In summary, the major effect of Rift Valley, the possible importance of Pleistocene climatic cycles and the strong differentiation between Central + East Africa and West Africa suggest that similar climatic and geological processes drove the genetic differentiation of *B. fusca* and herbivorous vertebrate populations. Various demographic indices suggest that each of the three clades of *B. fusca* experienced population expansion after this period of bottlenecks. Within each clade, genetic diversity is now geographically distributed in a pattern suggestive of restricted gene flow with isolation by distance.

The presence of a large population unit, clade *KII*, from Cameroon to Kenya is consistent with the hypothesis of a faunistic link between these two regions (Carcasson 1964; De Jong & Congdon 1993; Bruhl 1997), which are separated by a distance of 3000 km. An eastern origin is a possibility as it was suggested for a butterfly species in Cameroon (De Jong & Congdon 1993). De Jong & Congdon (1993) argue that the low animal species diversity in highland forests of Cameroon suggests that these species originated from long-distance migration from East Africa. However, the nature of this faunistic link was not elucidated by our study and remains unknown for *B. fusca*.

In West Africa, only a few haplotypes were shared between localities. The same was true, but to a lesser extent, in East Africa. Significant fine-scale structure was observed within each biogeographic region, and each geological and ecological category. This fine-scale structure suggests that *B. fusca* is a fairly sedentary species. However, field experiments are required to determine whether this behavioural pattern is the rule in all ecological situations. Kfir *et al.* (2002) have shown that, during the dry season, crop residues contain a large number of *B. fusca* larvae. The authors concluded that the insects present in the crops came from crop residues at the same site rather than migrating from other areas. Moyal (1998) reported large differences in larval population densities between two sites less than 50 km apart, which might be due to demographic isolation and low dispersal capacities. Nevertheless, our recent spurious observations of *B. fusca* in remote poorly suitable sites in northern Kenya, suggest that within this region, *B. fusca* population density is low or that long-distance migration may occur sporadically. The population from western Kenya (*KI*) seems to be more restricted geographically than the other two clades. Future collections of moths in Central African states (such as Uganda and Congo) will be required to improve our estimates of the geographic range of this clade.

The original wild host plant of *B. fusca* has not been identified in the field since this maize stalk borer is uncommon in wild habitats. However, the preference of this species for cultivated sorghum (Khan *et al.* 1997; Haile & Hofsvang 2002) suggests that it was originally associated with a particular species of wild sorghum. Experimental laboratory studies of oviposition choice among a large number of wild and cultivated Poacae, Cyperacae and Typhacae demonstrated that both adults (for oviposition) and larvae (for their development) display a preference for a particular wild sorghum variety (*Sorghum vulgare* pres. var. *sudanense*) (Khan *et al.* 2000; Haile & Hofsvang 2002). Haile & Hofsvang (2002) also showed that *S. vulgare* was the only wild host plant allowing complete development of this insect. The simplest hypothesis is that *B. fusca* switched from wild to cultivated sorghum during the process of domestication.

The domestication of wild sorghum began in northeast Africa, but investigations have shown that genetic exchange occurred locally between this common genetic stock and wild sorghum in all regions during the expansion of agriculture in Africa. Ollitrault *et al.* (1989) revealed that wild sorghum is geographically structured into three main groups, one from West Africa, one from southern Africa and one group associating East and Central Africa. The genetic structure of populations of cultivated sorghum has therefore, to some extent, been influenced by regional hybridization/introgression events with the wild sorghum species and this may have resulted in various regionally distinct genetic groups within cultivated sorghum (Deu & Hamon 1994). For insects associated with wild or cultivated sorghum, the same regional influence seems to be true. *B. fusca* displays a similar geographic pattern, with clades specific to West Africa and clades from East and Central Africa. Further analyses on southern African populations will show us whether they belong to yet another clade. Although we have found a similar pattern, we cannot conclude that the same events structured the insect and its putative wild host plant. We infer the existence of palaeo-refuges for *B. fusca* based on phylogeographic analyses and present ecological preferences of this species. In contrast, the ecological preferences of wild sorghum strains are highly variable (De Wet 1978). This suggests that the refuges and/or fragmentation barriers that have shaped the population structure of the insect may be different for the plant. However, for both the insect and the plant it is likely that ancient events shaped the observed genetic

structure. The influence of ancient genetic events may be even more important for the insect than for the plant: the genetic structure of *B. fusca* populations clearly does not derive from a recent invasion of a common genetic pool originating from the region of sorghum domestication (northeast Africa), but it is instead exclusively inherited from patterns determined by the ancient history of the insect on wild Poaceae in a palaeoclimatic context that favoured the geographic isolation of three different populations in refuge zones corresponding to *B. fusca* ecological preferences. This suggests that the switch from wild to cultivated sorghum occurred several times from local populations of *B. fusca*. This hypothesis of the stability of genetic structure from ancient wild *B. fusca* to actual populations associated to cultivated plants should be tested for other regions of the genome, in particular nuclear genes. This will increase our understanding on the signature of plant domestication on pest population genetic structure. For instance, in the pyralid *Ostrinia nubilalis*, the European corn borer, no genetic differentiation was observed between insects on maize and wild hosts for mitochondrial markers whereas strong differentiation was observed for nuclear allozymes (Thomas *et al.* 2003). Thus, only part of the genome may be affected by domestication and involved in the adaptation to cultivated plants.

In conclusion, based on mitochondrial DNA, three major groups of populations can be distinguished in *B. fusca*: a homogeneous and geographically isolated population from West Africa, and two populations from East and Central Africa with overlapping distributions. Phylogenetic relationships suggest that the three populations differentiated in different regions of Africa during the same period, before human-mediated ecological changes. Analysis of molecular variance and NCPA interpretations suggest that local population differentiation was shaped by restricted gene flow with isolation by distance. The strong genetic structure observed suggests that *B. fusca* has in general limited dispersal capacities.

Although little is known about the biogeography of African phytophagous insects, the factors responsible for the patterns observed in *B. fusca* can be compared with those identified in studies on other herbivores, particularly mammals. The East Africa Rift Valley was also identified as a major factor for genetic structure for several groups of vertebrates in Africa (Arctander *et al.* 1999; Pitra *et al.* 2002). More generally, the study on *B. fusca* suggests that climatic fluctuations in the Pleistocene generated isolated populations and shaped the genetic structures of various groups of herbivores associated with gramineous plant communities in Africa, both for mammals (Arctander *et al.* 1999; Flagstad *et al.* 2001; Alpers *et al.* 2004) and insects. These results for *B. fusca* generate new hypotheses for future investigations of African insect phylogeography.

Acknowledgements

This study was supported by IRD UR072. We would like to thank D. Baille for technical assistance in the development of molecular markers and preliminary phylogeographic analyses, and the members of the ICIPE Noctuid Stem Borer Biodiversity Project for technical assistance during insect collection and insect rearing in the laboratory in Kenya. We are grateful to C. Wanjiru Gitau and L. Corley for their help to the English correction of the manuscript. We are indebted to G. Goergen for his support and Z. Djihou for technical assistance at IITA-Cotonou in Benin. B. Emerson and two anonymous referees also helped greatly in improving this manuscript.

References

Ajayi O (1998) Sorghum: West Africa. In: *African Cereal Stem Borers: Economic Importance, Taxonomy, Natural Enemies and Control* (ed. Polaszek A), pp. 39–45. CABI, Wallingford, UK.

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, **13**, 1771–1784.

Althoff MD, Pellmyr O (2002) Examining genetic structure in bogus yucca moth: a sequential approach to phylogeography. *Evolution*, **56**, 1632–1643.

Arctander P, Johansen C, Coutellec-Vreto M (1999) Phylogeography of three closely related African bovids (tribe *Alcelaphini*). *Molecular Biology and Evolution*, **16**, 1724–1739.

Avise JC (2000) *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351–379.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Bruhl CA (1997) Flightless insects: a test case for historical relationships of African mountains. *Journal of Biogeography*, **24**, 233–250.

Carcasson RH (1964) A preliminary survey of the zoogeography of African butterflies. *East African Wildlife Journal*, **2**, 122–157.

Chastanet M (1998) Introduction et place du maïs au Sahel Occidental (Sénégal-Mauritanie). In: *Plantes et Paysages d'Afrique. Une Histoire à Explorer* (ed. Chastanet M), pp. 251–275. Karthala-CRA, Paris.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Dakouo D, Ratnadass A (1997) Les lépidoptères foreurs de tige du sorgho en Afrique de l'Ouest: fluctuations saisonnières, importance économique et perspectives de lutte intégrée. *Annales de la Société Entomologique de France*, **35** (Suppl.), 463–470.

De Jong R, Congdon TCE (1993) The montane butterflies of the eastern Afrotropics. In: *Biogeography and Ecology of the Rain Forests of Eastern Africa* (eds Lovett JC, Wasser SK), pp. 133–164. Cambridge University Press, Cambridge.

Deu M, Hamon P (1994) Diversité des sorghos: application à la gestion des ressources génétiques et à la sélection. *Agriculture et Développement*, **3**, 25–30.

De Wet JMJ (1978) Systematics and evolution of sorghum section *Sorghum* (Graminae). *American Journal of Botany*, **65**, 477–484.

Doggett H (1988) *Sorghum*, 2nd edn. Longman Scientific and Technical Publishers, London.

Duran S, Giribet G, Turon X (2004) Phylogeography history of the sponge *Crambe crambe* (Porifera, Poecilosclerida): range expansion and recent invasion of the Macaronesian islands from the Mediterranean Sea. *Molecular Ecology*, **13**, 109–122.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Flagstad Ø, Syvertsen PO, Stenseth NC, Jakobsen KS (2001) Environmental change and rate of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 667–677.

Grant SW, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**, 415–426.

Gu X, Fu YX, Li WH (1995) Maximum likelihood estimation of heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution*, **12**, 546–557.

Haile A, Hofsvang T (2001) Survey of lepidopterous stem borers of sorghum, maize and pearl millet in Eritrea. *Crop Protection*, **20**, 151–157.

Haile A, Hofsvang T (2002) Host plant preference of the stem borer *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae). *Crop Protection*, **21**, 227–233.

Harry M, Solignac M, Lachaise D (1998) Molecular evidence for parallel evolution of adaptive syndromes in fig-breeding *Lissocephala* (Drosophilidae). *Molecular Phylogenetics and Evolution*, **9**, 542–551.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **21**, 160–174.

Hassan RM, Corbett JD, Njoroge K (1998) *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya*. CAB International, Wallington, Oxon, UK.

Holloway JD (1998) Noctuidae. In: *African Cereal Stem Borers: Economic Importance, Taxonomy, Natural Enemies and Control* (ed. Polaszek A), pp. 79–86. CABI, Wallingford, UK.

Kfir R, Overholt AW, Khan RZ, Polaszek A (2002) Biology and management of economically important lepidopteran cereal stem borers in Africa. *Annual Review of Entomology*, **47**, 701–731.

Khan ZR, Chiliswa P, Ampong-Nyarko K *et al*. (1997) Utilization of wild gramineous plants for the management of cereal stem borers in Africa. *Insect Science and its Application*, **17**, 143–150.

Khan ZR, Pickett JA, van den Berg J, Wadhams LJ, Woodcock CM (2000) Exploiting chemical ecology and species diversity: stem borer and striga control for maize and sorghum in Africa. *Pest Management Science*, **56**, 957–962.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623–2635.

Kuchta SR, Meyer D (2001) A genealogical view of geographical variation. *Molecular Ecology*, **10**, 2569–2576.

Livingstone DA (1982) Quaternary geography of Africa and refuge theory. In: *Biological Diversification in the Tropics: Proceedings of the 5th International Symposium of the Association for Tropical Biology* (ed. Prance GT), pp. 523–536. Columbia University Press, New York.

Maddison DR, Maddison WP (2002) MACCLADE: *Analysis of Phylogeny and Character Evolution*, Version 4.02. Sinauer Associates, Sunderland, Massachusetts.

Madeira Santos ME, Ferraz Torrao MM (1998) Entre l'Amérique et l'Afrique, les îles du Cap-vert et Sao Tomé: les cheminements des milhos (mil, sorgho et maïs). In: *Plantes et Paysages d'Afrique. Une Histoire à Explorer* (ed. Chastanet M), pp. 69–83. Karthala-CRA, Paris.

deMenocal PB (1995) Plio–Pleistocene African climate. *Science*, **270**, 53–59.

Moyal P (1998) Infestation patterns and parasitism of the maize stalk borer, *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae) in the Ivory Coast. *African Entomology*, **6**, 289–296.

Murty DS, Renard C (2001) Sorgho (*Sorghum bicolor* (L.) Moench). In: *Agriculture en Afrique Tropicale* (ed. Raemaekers RH), pp. 105–125. Direction Générale de la Coopération Internationale, Bruxelles, Belgique.

Ndemah R, Schulthess F, Poelhing M, Borgemeister C (2000) Species composition and seasonal dynamics of lepidopterous stem borers on maize and elephant grass, *Pennisetum purpureum* (Moench) (Poaeceae), at two forest margin sites in Cameroon. *African Entomology*, **8**, 265–272.

Ndemah R, Schulthess F, Korie S, Borgemeister C, Cardwell KF (2001) Distribution, relative importance and effect of lepidopterous borers on maize yields in forest zone and mid-altitude of Cameroon. *Journal of Economic Entomology*, **94**, 1434–1444.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.

Nye IWR (1960) *The Insect Pests of Graminaceous Crops in East Africa*. Colonial Research Studies, 31. Her Majesty's Stationary Office, London, UK.

Ollitrault P, Arnaud M, Chantereau J (1989) Polymorphisme enzymatique des sorghos. II. Organisation génétique et évolutive des sorghos cultivés. *L'Agronomie Tropicale*, **44**, 211–222.

Pitra C, Hansen JA, Lieckfeldt D, Arctander P (2002) An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology*, **11**, 1197–1208.

Polaszek A, Khan ZR (1998) Host plants. In: *African Cereal Stem Borers: Economic Importance, Taxonomy, Natural Enemies and Control* (ed. Polaszek A), pp. 3–10. CABI, Wallingford, UK

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Quérouil S, Verheyen E, Dillen M, Colyn M (2003) Patterns of diversification in two African forest shrews: *Sylvisorex johnstoni* and *Sylvisorex ollula* (Soricidae, Insectivora) in relation to paleoenvironmental changes. *Molecular Phylogenetics and Evolution*, **28**, 24–37.

Ristanovic D (2001) Maïs (*Zea mays* L.). In: *Agriculture en Afrique Tropicale* (ed. Raemaekers RH), pp. 44–70. Direction Générale de la Coopération Internationale, Bruxelles, Belgique.

Schneider SD, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise distances when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Schneider SD, Roessli D, Excoffier L (2000) ARLEQUIN *(version 2000): A software for genetic data analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Schultess F, Bosque-Pérez NA, Shabi-Olaye A, Gounou S, Ndemah R, Goergen G (1997) Exchanges of natural enemies of lepidopteran cereal stem borers between African regions. *Insect Sciences and its Application*, **17**, 97–108.

Simmons RB, Weller SJ (2001) Utility and evolution of cytochrome *b* in insects. *Molecular Phylogenetics and Evolution*, **20**, 196–210.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Sithole SZ (1989) Sorghum stem borers in Southern Africa. In: *International Workshop Sorghum Stem Borers, Patancheru, 1987* (ed. Nwanze KF), pp. 41–47. ICRISAT, Patancheru, India.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and other Methods)*, Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989a) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypothesis about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.

Thomas Y, Bethenod MT, Pelozuelo L, Frérot B, Bourguet D (2003) Genetic isolation between two sympatric host-plant races of the European corn borer, *Ostrinia nubilalis* Hübner. I. Sex pheromone, moth emergence timing, and parasitism. *Evolution*, **57**, 261–273.

Van Hooft WF, Groen AF, Prins HHT (2002) Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Molecular Ecology*, **11**, 267–279.

Wagner T (2002) Late Cretaceous to early Quaternary organic sedimentation in the eastern Equatorial Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **179**, 113–147.

White F (1983) The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. *UNESCO, Natural Resources Research*, **20**, 1–356.

Yang Z (1993) Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution*, **10**, 1396–1402.

Zeitoun V (2000) Adéquation entre changements environnementaux et spéciations humaines au Plio–Pléistocène. *Comptes Rendus de l'Académie Des Sciences*, **330**, 161–166.

Michel Sezonlin is a population geneticist. His research aims to understand the factors that shaped cereal stemborer population heterogeneity in Africa. Stéphane Dupas is a population geneticist studying insect-parasitoid-microorganism coevolutionary dynamics. Bruno Le Rü is an entomologist whose main line of research is to understand plant-insect spatial dynamics in relation to wild host plants use in agricultural systems. Philippe Le Gall is an entomologist working on African insect biogeography. Pascal Moyal is an entomologist studying the systematics, phylogeny and phylogeography of African noctuid stem borers of monocot plants. Paul-André Calatayud is a chemical ecologist whose main line of research is understanding the biochemistry and genetic diversity of host plant choice and suitability for development of cereal stemborers. Isabelle Giffard is a chemist working on insect molecular biology. Nathalie Faure is a molecular ecologist working on insect population genetics and systematics. Jean-François Silvain is an insect systematician and evolutionary biologist. His main line of research is to understand factors that shaped insect diversity.

1.2 Article 2

Phylogeographic pattern and regional evolutionary history of the maize stalk borer *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae) in sub-Saharan Africa

Michel SEZONLIN, Stéphane DUPAS, Bruno LE RÜ, Nathalie FAURE,
Philippe LE GALL and Jean-François SILVAIN

Sous presse dans les Annales de la Société Entomologique de la France

Patron phylogéographique et histoire évolutive régionale du foreur de tiges de maïs *Busseola fusca* (Lepidoptera, Noctuidae) en Afrique Subsaharienne

Ce second article reprend la démarche analytique employée précédemment, mais, le jeu de données analysé a été considérablement complété sur le plan géographique afin de tester les hypothèses issues du premier article.

Contexte de l'étude : afin d'approfondir les connaissances sur la génétique et l'évolution des populations de *B. fusca* et en déduire l'histoire régionale, l'étude à plus grande échelle de ce foreur de tiges a été réalisée. Cette seconde étude devait permettre de compléter et d'affiner le patron phylogéographique observé précédemment et de tester les grandes hypothèses dégagées au cours de la première analyse qui a porté seulement sur six pays (Bénin, Togo, Ghana, Mali, Cameroun et Kenya) et qui a mis en évidence l'existence de trois clades majeurs avec des distributions géographiques et des structurations génétiques caractéristiques. Certaines régions africaines (cas de la région des Grands Lacs) sont connues comme des lieux abritant une importante biodiversité alors qu'une dépression sèche comme celle de la Rift Valley semblait jouer un rôle dans la structuration des espèces. Une meilleure couverture des régions déjà étudiées, mais aussi, celle de distribution de l'aire de *B. fusca* s'avèrent nécessaires étant donné qu'on se pose plusieurs questions concernant la nature des populations présentes dans les autres régions (Afrique Australe et Centrale) non étudiées dans l'étude précédente.

Objectifs de l'étude : Il s'agit (1) de comprendre l'histoire évolutive de *B. fusca* à l'échelle du continent africain, (2) d'identifier le centre d'origine possible des populations mises en évidence et de préciser le patron phylogéographique ayant modelé chaque clade de *B. fusca*. En outre, il convient (3) de déterminer si d'autres clades existent, notamment en Afrique Australe, clades qui pourraient correspondre à des zones refuges différentes de celles mises en évidence précédemment. L'existence d'un clade de *B. fusca* propre à l'Afrique Australe et correspondant au groupe Sud africain de sorgho cultivé doit être, en particulier, recherchée. Il convient enfin (4) de déterminer les limites géographiques de chaque clade identifié.

Méthodes utilisées : l'échantillonnage des individus de *B. fusca* sur le sorgho cultivé couvre la majeure partie de l'aire distribution de ce foreur. Cet échantillonnage est fait parallèlement avec la collecte des individus de *B. fusca* vivant sur maïs. Le même marqueur moléculaire, un fragment du gène codant pour le cytochrome *b*, a été séquencé chez tous les individus collectés. L'analyse des données moléculaires est faite en combinant les différentes méthodes utilisées en phylogéographie et présentées au chapitre 3.

Principaux résultats et points discutés : cette étude confirme l'existence des trois clades principaux d'haplotypes (*W*, *KI*, *KII*) chez *B. fusca*. Il n'y a pas de nouveau clade spécifique à l'Afrique Australe. Les analyses démographiques et phylogénétiques indiquent une expansion démographique locale s'accompagnant des phénomènes de restriction de flux de gènes avec isolement par la distance. Le clade *W* reste séparé géographiquement des clades *KI* et *KII* qui sont partiellement sympatriques. La distribution géographique du clade *KII*, connue de l'Afrique de l'Est à l'Afrique Centrale, s'étend jusqu'en Afrique Australe. De même, celle du clade *KI* connue uniquement au Kenya remonte jusqu'en Erythrée en passant par Ethiopie, Ouganda et Ruanda. L'analyse de mismatch distribution et les valeurs négatives de l'indice D de Tajima sont bien en accord avec l'hypothèse d'une expansion démographique de ces trois clades. Des différenciations génétiques significatives sont révélées aux différents niveaux géographiques par l'analyse moléculaire de la variance (AMOVA). L'estimation des indices de diversité haplotypique et nucléotidique à travers les différentes zones écologiques ou géographiques (regroupement des pays) définies a permis de proposer des centres probables d'origine géographique de chaque clade : la région forestière d'Afrique de l'Ouest (Ghana) pour le clade *W*, la région érythréenne pour le clade *KI* et la région Est-Sud africaine (Kenya) pour le clade *KII*.

Abstract *Busseola fusca* (Fuller) (Lepidoptera, Noctuidae) is one of the major cereal pests in sub-Saharan Africa. Previous phylogeographic investigations on samples collected in Kenya, Cameroon and West-Africa showed the presence of three main clades (*W*, *KI*, *KII*) originated from populations isolated in West and East Africa around one million years ago. Demographic and phylogenetic analyses suggested that this event was followed by local demographic expansion and isolation by distance. These hypotheses were tested by a more comprehensive sampling across *B. fusca*'s geographic range in Africa. Comparisons of sequences of partial mitochondrial DNA gene (cytochrome *b*) from 489 individuals of 98 localities in southern, central, eastern and western African countries confirmed the presence of the three main clades. Phylogenetic, F-statistics, demographic parameters and nested clade phylogeographic analyses confirmed that the clades experienced geographic and demographic expansion with isolation by distance after their isolation in three refuge areas. The geographic range of clade *KII*, already known from East to Central sub-Saharan Africa was extended to Southern Africa. Mismatch distribution analysis and the negative values of Tajima's D index are consistent with a demographic expansion hypothesis for these three clades. Significant genetic differentiations were revealed at various hierarchical levels by analysis of molecular variance (AMOVA). Hypotheses about the geographic origin of the three main clades are detailed.

Scénario phylogéographique et histoire évolutive régionale du foreur de graminées *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae) en Afrique sub-saharienne.

Résumé *Busseola fusca* (Fuller) (Lépidoptère, Noctuidae) est l'un des ravageurs majeurs des cultures céréalières en Afrique Subsaharienne. Une première étude phylogéographique portant sur des individus échantillonnés au Kenya, au Cameroun et en Afrique de l'Ouest a montré l'existence de trois clades principaux (*W*, *KI*, *KII*) issus de populations isolées à l'Ouest et à l'Est de l'Afrique il y a environ un million d'années. Les analyses démographiques et

phylogénétiques indiquent que cet événement a été suivi d'une expansion démographique locale avec des phénomènes d'isolement par la distance. Ces hypothèses ont été testées à plus grande échelle grâce à un échantillonnage des populations de *B. fusca* couvrant désormais la majeure partie de son aire de distribution. Le séquençage d'un fragment du gène mitochondrial codant pour le cytochrome *b* chez 489 individus provenant de 98 localités des pays sud, centre, est et ouest africains confirme l'existence des trois clades observés précédemment. Les résultats des analyses phylogénétiques, les paramètres démographiques, les statistiques de Wright ainsi que les analyses des clades emboîtés confirment que ces trois populations, après avoir été isolées dans des aires refuges différentes, ont connu une expansion démographique et géographique avec un isolement par la distance. La distribution géographique du clade *KII*, connue de l'Afrique l'Est à l'Afrique centrale, s'étend jusqu'en Afrique Australe. L'analyse de 'mismatch distribution' et les valeurs négatives de l'indice D de Tajima sont bien en accord avec l'hypothèse d'une expansion démographique de ces trois clades. Des différenciations génétiques significatives ont été révélées aux différents niveaux hiérarchiques par l'analyse moléculaire de la variance (AMOVA). Les hypothèses sur l'origine géographique des trois clades sont précisées.

Keywords: Stem borer, population genetics, cytochrome *b*, Pleistocene, Africa.

Introduction

The most important cereal crops in sub-Saharan Africa are maize, millet, rice and sorghum. Most of these cereal crops serve as host plants for many stem borer insects among which *Busseola fusca* (Fuller 1901) (Lepidoptera: Noctuidae) is one of the most economically important pests (Polaszek & Khan 1998). *B. fusca* is an endemic species in sub-Saharan Africa with wide geographical distribution (Ajayi 1998; Moyal 1998; Ndemah *et al*. 2001; Haile & Hofsvang 2001; Kfir *et al*. 2002). It is a major pest of sorghum (*Sorghum bicolor* [L.] Moench) and maize (*Zea mays* L.) (Poaceae). The domestication of sorghum probably began some 5,000 years ago in North-East Africa (Dogett 1988; Murty & Renard 2001) whereas maize was introduced more recently, at the end of the 16th Century (Madeira Santos & Ferraz Torrao 1998; Chastanet 1998). *Busseola fusca* varies in its ecological preference across its geographical range. It is more adapted to lowland in West Africa than in East and Southern Africa (Kfir *et al*. 2002). Previous study on the genetic structure of *B. fusca* shows that this ecological preference is associated with major differences in partial DNA sequence of the cytochrome *b* mitochondrial gene (Sezonlin *et al*. 2006). According to that study, *B. fusca* populations are differentiated into three major clades of mitochondrial haplotypes, one located in West African region (*W*), one restricted to East Africa (*KI*) and one found in Central and East Africa (*KII*). The origin of these clades is likely related to Pleistocene climatic events. Partial geographic overlap was observed only between clades *KI* and *KII*. Biogeographic barriers likely corresponding to *B. fusca* ancient history on wild Poaceae have been shown to be the major factors of differentiation of this species. These barriers, namely the Cameroon Volcanic Line (CVL) region between Central and West Africa and the Rift Valley in East Africa, appear to be similar to those that shaped geographic differentiation of phytophagous mammals and rodents (Sezonlin *et al.* 2006).
However, no signature of sorghum domestication and maize introduction has been detected yet on the genetic structure of *B. fusca*, despite the expected important demographic consequences of this switch to cultivated plants.

The regional evolutionary history and the centres of origin of the mitochondrial clades remain unknown. To answer these questions and to more accurately estimate geographical distribution of each clade, the sampling of *B. fusca* was completed to include most of its geographic range from Western, Eastern to Southern Africa. Samples from Eritrea, Ethiopia, Malawi, Mozambique, Rwanda, Republic of South Africa, Uganda, Zambia and Zimbabwe

were added. Partial sequences of the gene coding for cytochrome *b*, informative at the intrageneric level in Lepidoptera (Simmons & Weller 2001; Sezonlin *et al*. 2006) have been used. A sequential approach combining several phylogeographic and evolutionary methods (Bernatchez 2001) was used to analyse the molecular data and infer the demographic history of *B. fusca* populations in more details. Such analyses in sequential approach that start from phylogeny to evolutionary history via demography and genetic structure allow us to move from testing deeper phylogenetic splits to inferring recent patterns of population structure. This also may highlight the centres of origin of *B. fusca* populations and may elucidate the regional evolutionary history that has produced this genetic structure.

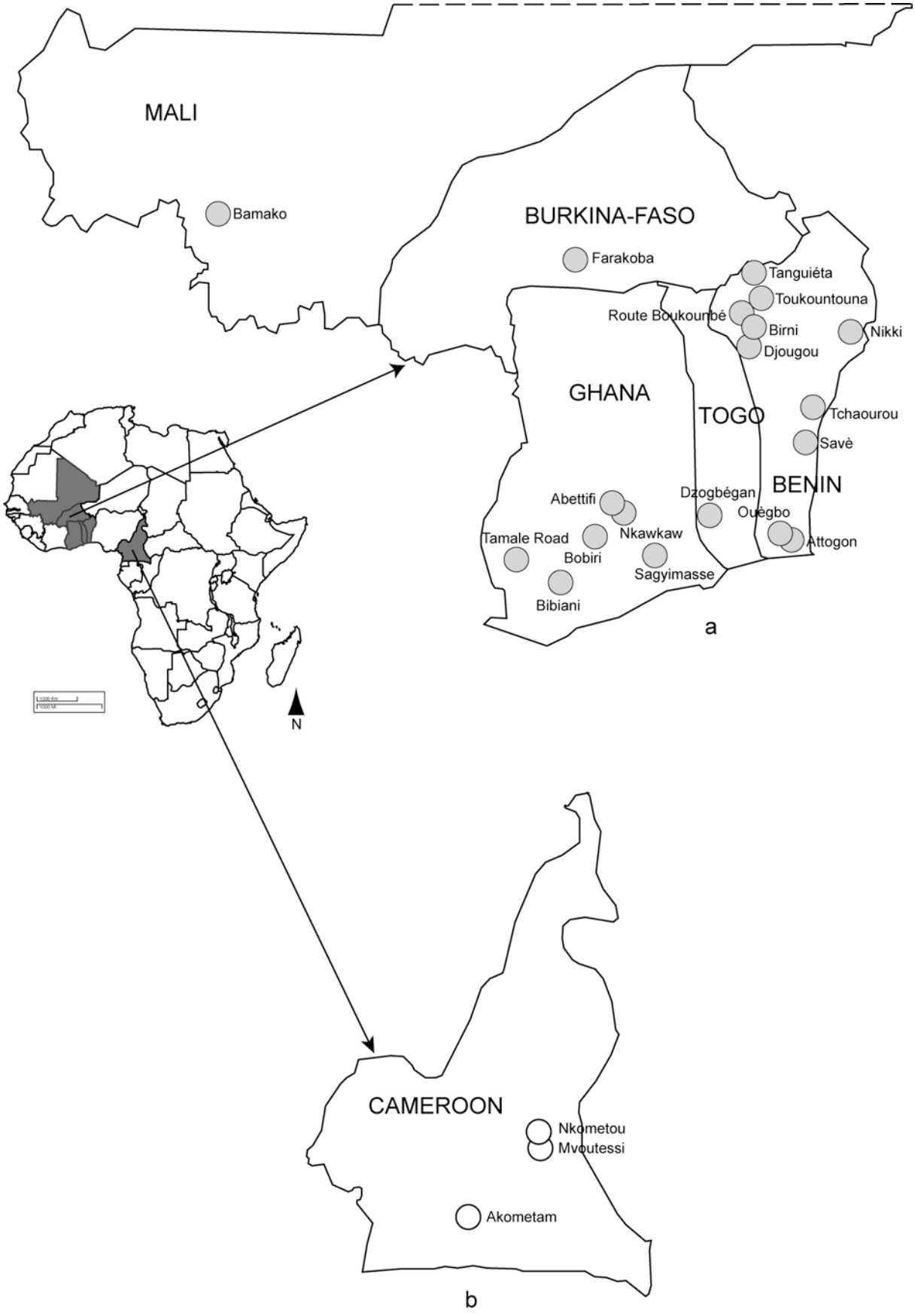
Material and methods

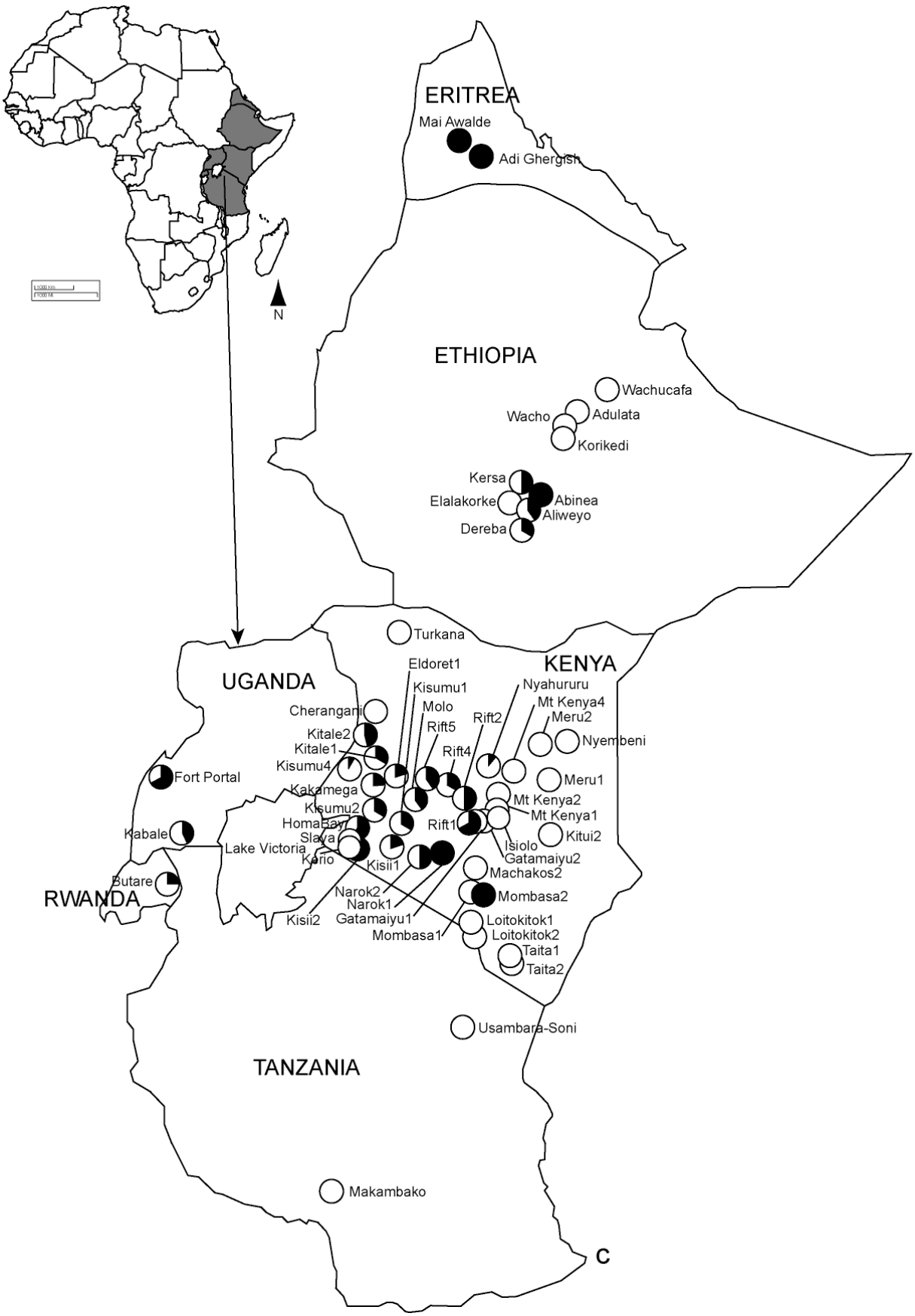
Moth sampling

The sampling of *Busseola fusca* individuals was carried out between 2001 and 2004. *B. fusca* individuals were sampled in West Africa (19 localities from Benin, Togo, Ghana, Mali and Burkina-Faso), in Central Africa (3 localities from Cameroon), in East Africa (55 localities from Kenya, Uganda, Rwanda, Tanzania, Ethiopia and Eritrea) and in Southern Africa (21 localities from Malawi, Mozambique, Zambia, Zimbabwe and Republic of South Africa) (figs.1a, 1b, 1c, 1d).

Moths rearing and conservation

Larvae and pupae collected were brought to the laboratories (IITA - Cotonou for Central and West Africa and ICIPE - Nairobi for Southern and Eastern Africa) to be reared to adulthood on semi-natural medium made of fresh stems of maize and cultivated sorghum (IITA) and artificial medium (ICIPE). The rearing of larvae allows the morphological identification of *B. fusca* moths among other stem borers species. The moths were killed just after emergence and preserved in absolute ethanol before DNA extraction.


MALI
Bamako
BURKINA-FASO
Farakoba
Tanguiéta
Toukountouna
Route Boukoumbé
Birni
Nikki
Djougou
GHANA
TOGO
Tchaourou
Savè
BENIN
Abettifi
Dzogbégan
Ouègbo
Tamale Road
Nkawkaw
Attogon
Bobiri
Sagyimasse
Bibiani
a
N
CAMEROON
Nkometou
Mvoutessi
Akometam
b


ERITREA
Mai Awalde
Adi Ghergish
ETHIOPIA
Wachucafa
Wacho
Adulata
Korikedi
Kersa
Elalakorke
Abinea
Aliweyo
Dereba
N
Turkana
KENYA
UGANDA
Eldoret1
Kisumu1
Nyahururu
Mt Kenya4
Meru2
Molo
Rift5
Rift2
Cherangani
Kitale2
Kitale1
Kisumu4
Rift4
Nyembeni
Meru1
Fort Portal
Kakamega
Kisumu2
Mt Kenya2
Mt Kenya1
Homabay
Rift1
Kitui2
Kabale
Siaya
Lake Victoria
Kano
Kisii1
Isiolo
Gatamaiyu2
Machakos2
Butare
RWANDA
Narok2
Mombasa2
Narok1
Kisii2
Gatamaiyu1
Loitokitok1
Mombasa1
Loitokitok2
Taita1
Taita2
Usambara-Soni
TANZANIA
Makambako
C

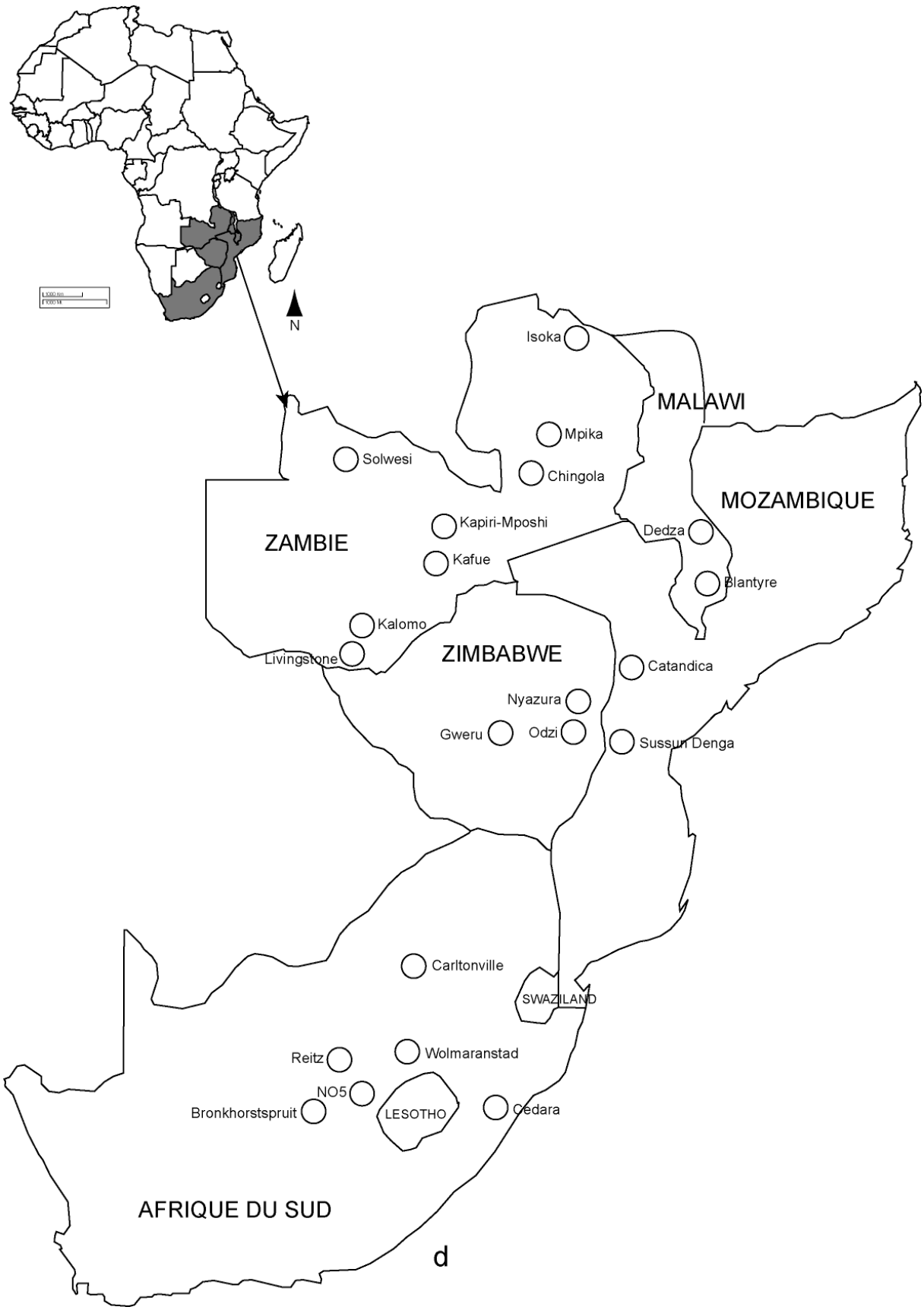


Figure 1. Geographic distribution of localities where *B. fusca* was sampled and the percentage of each clade in West Africa (a), Central Africa (b), East Africa (c) and Southern Africa (d). Grey dots: haplotypes belonging to the clade *W*, black dots: haplotypes belonging to the clade *KI*, white dots: haplotypes belonging to the clade *KI.*

Molecular analysis

Total DNA was extracted from insect thoraxes, using the DNeasy tissue kit (Qiagen GmbH, Germany). The number of individuals of *B. fusca* analyzed for each locality ranged from one to 13. The molecular marker used is a fragment of the gene coding for the cytochrome *b* for which approximately 1000 bp were amplified by PCR. The same primers and PCR protocol as described by Sezonlin *et al.* (2006) were used for all samples. Amplified PCR products were purified with the Quick protocol (Promega Wizard SV Gel and PCR Clean-Up System) and directly sequenced on an automated sequencer ABI prism 377 using the amplification primers in both directions. The consensus sequences obtained were aligned manually using MacClade 4.06 (Maddison & Maddison 2002).

Haplotype phylogeny

Phylogenetic relationships were estimated by means of maximum parsimony (MP) and Neighbour-Joining (NJ) using Maximum-Likelihood distances with PAUP* 4b10 (Swofford 2002). MP analyses were performed using a heuristic search strategy starting with stepwise addition trees replicated 10 times, using a random input order of sequences to get the initial tree for each replicate. Robustness of MP topologies was assessed by bootstrap with 100 replicates (full heuristic search) of 10 random stepwise addition replicates each, for all analyses.

Although the systematics of African lepidopteran stem borers is still rather confused (Holloway 1998), recent studies showed that *Busseola phaia* Bowden 1956 (Lepidoptera: Noctuidae), collected from various regions of East Africa, is the sister species of *B. fusca* (Moyal, *pers. com*.). Therefore, *B. phaia* was chosen as outgroup.

MODELTEST version 3.07 (Posada & Crandall 1998) was used to select the substitution model(s) that best describe the data. This software performs a hierarchical test of likelihood fits under 56 different models of character variation.

Diversity indices and demographic history of *B. fusca* clades

Haplotypic (h) and nucleotide (π) diversity values and all demographic parameters were performed with ARLEQUIN 2.000 software (Schneider *et al.* 2000) for groups of haplotypes in order to estimate their level of polymorphism and to localize the centre of origin of the different clades. These diversity indices are useful to examine the demographic history of a

lineage (Grant & Bowen 1998) because their value does not depend on the length of the DNA fragment, nor on the sample size (Nei & Li 1979; Nei 1987). The centre of origin of each *B. fusca* clade was established by comparing the genetic diversity for different groups defined within this main population. In this case, populations were grouped according to geography in Central, South and East Africa and phytogeographic zones (White 1983) in West Africa. In West Africa three phytogeographic zones concerned our study, forest (drier types) region, forest and secondary grassland region and savannah region whereas in the rest of Africa, four geographic zones were retained, namely East – North, East – South, Austral and Cameroon. Centres of origin would have higher haplotype and nucleotide diversity than more recently founded populations (Althoff & Pellmyr 2002). This analysis could allow us to track the recent geographical expansions of these populations. The distribution of pairwise differences between individual sequences was analyzed by means of mismatch distribution analysis (Slatkin & Hudson 1991; Schneider & Excoffier 1999). A unimodal distribution would be expected for populations in expansion or for populations that have undergone a recent bottleneck, and a multimodal distribution for populations at demographic equilibrium (Slatkin & Hudson 1991). The raggedness index of the observed distribution (*r*) representing the modality of the distribution, and the sum of square deviation from the mismatch expected from a model of sudden population expansion (SSD) were calculated. Since the nucleotide substitution models selected by hierarchical likelihood ratio tests (hLRTs) (HKY + I + G) and Akaike information criterion (AIC) (K81uf + I + G) were not available in the ARLEQUIN 2.000 software, the *r* and SSD indices were calculated by using pairwise differences. The significance of these statistics was tested as implemented in ARLEQUIN. Finally, Tajima's D index was calculated with ARLEQUIN. This index can provide information about demographic history with demographic expansion leading to negative values, and subdivided populations leading to positive values (Tajima 1989a, b).

Genetic structure of *B. fusca* populations

In order to test for genetic differentiation, hierarchical levels of genetic divergence between various groups were calculated with the fixation index Φ_{ST} (Excoffier *et al*. 1992), an estimator that includes information on haplotype frequency and molecular distance. The significance of Φ_{ST} for population comparisons was assessed using 1000 permutations. The Φ_{ST} values and permutations were computed in ARLEQUIN 2.000 (Schneider *et al*. 2000).

Nested clade phylogeographic analysis (NCPA)

NCPA was performed as described by Templeton (1998, 2004). The genealogic relationships are represented through a haplotype parsimony network to define a series of nested clades. The probabilities of haplotype connections were calculated according to coalescent theory using TCS1.21 software (Clement *et al*. 2000) and the network with probabilities above the parsimony threshold (0.95) was selected. The hypothesis of random geographic distributions is tested through permutation tests for each clade and subclade components. These statistical analyses of geographical distances within and between clades were carried out with GeoDis 2.1 (Posada *et al*. 2000). GPS coordinates of all sampling localities were used. The geographical distances between centres of distributions of clades were tested for significance in permutation tests, within clade (*Dc*, the average distance of individuals from the clade's geographical centre), with nested clade centre (*Dn*, the average distance of individuals from the geographical center of all members of the nested clade) or between interior and tip at each level (*(I-T)Dc*, the average distance between interior and tip clades within a given clade and *(I-T)Dn*, the average distance between interior and tip clades in the nested clade). Significant geographic patterns were interpreted in terms of population history, using the latest inference key from Templeton (2004) from http://darwin.uvigo.es.

Results

Phylogenetic reconstruction

A fragment of 965 bp encoding cytochrome *b* was sequenced from 489 individuals of *Busseola fusca* across its geographic range from Western, Central, Southern and Eastern Africa. We observed 108 different haplotypes (GenBank accession numbers AY769536 to AY769605 and DQ284857 to DQ284895). The haplotypes with GenBank accession number AY769536 to AY769605 were re-used whereas those with accession number DQ284857 to DQ284895 were news. 123 nucleotide sites were variable (12.75%) and 58 were informative in parsimony analysis (6.01%). Parsimony analysis generated 438 equiparsimonious trees (length = 342, CI = 0.371, RI = 0.808). All trees were divided into the same three clades: a clade grouping sequences from the West African region only (*W*), a Kenya I clade (*KI*) and a Kenya II clade (*KII*), which also contained sequences from Cameroon and southern African countries (fig. 2). Discrepancies between these 438 equiparsimonious topologies concerned

only the apical nodes. It was therefore possible to construct a majority rule consensus of the most parsimonious trees (fig. 2).

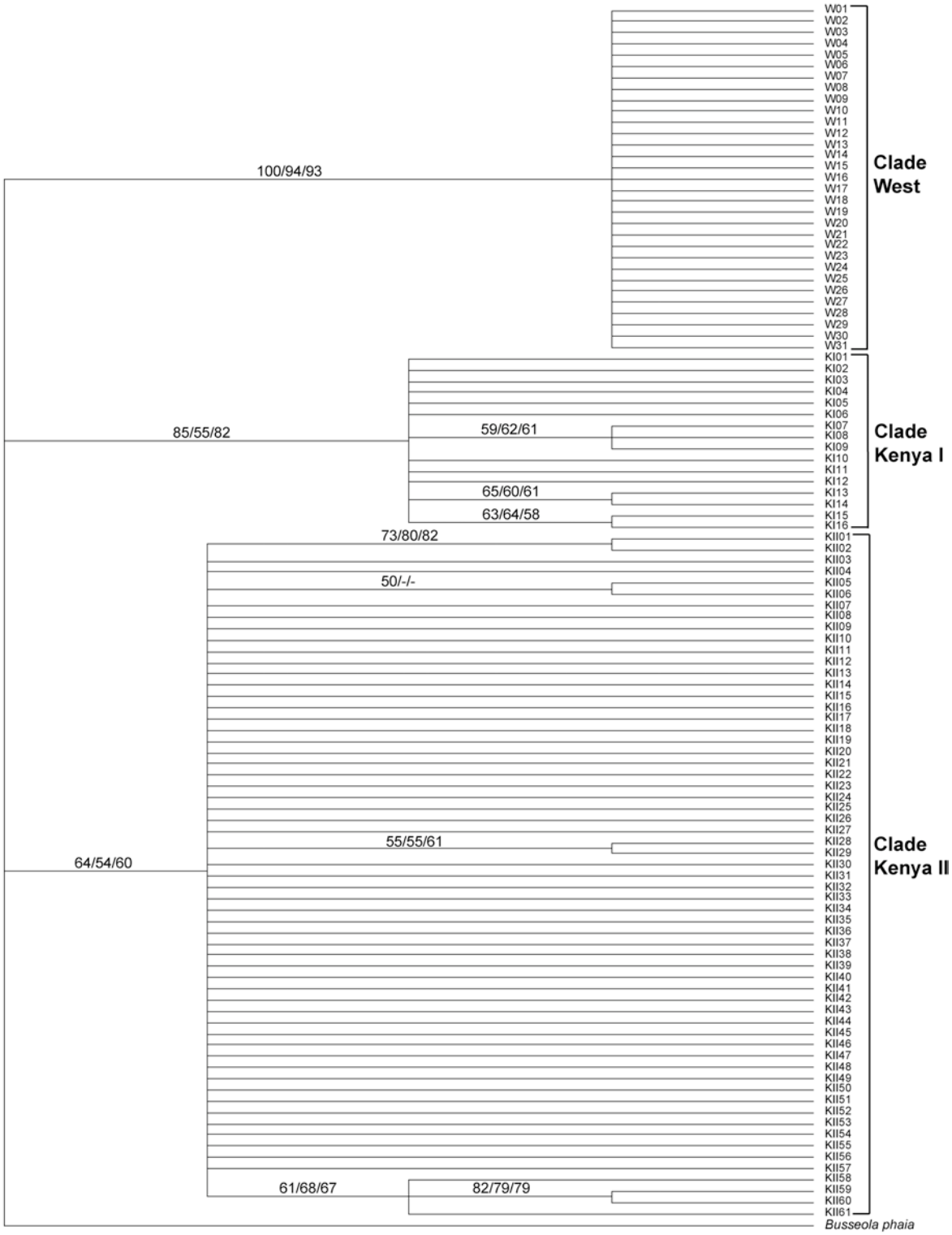


Figure 2. Majority consensus of the most parsimonious trees. Bootstrap support of >50% in both MP (first number), NJ with hLRT (second number), and NJ with AIC (third number) searches in 1000 replicates are given for the relevant nodes. *Busseola phaia* was used as the outgroup taxon.

The model selected by the Akaike information criterion of the maximum likelihood (ML) was the K81uf + I + G (-LnL=2479.6731) (Kimura 1981). The parameters inferred from this substitution model were: A=0.337, C=0.140, G=0.106, T=0.417; [AC substitution rate] = 1.000, [AG] = 22.5462, [AT] = 0.3998, [CG] = 0.3998, [AT] = 22.5462, [GT] = 1.000 with a certain proportion of invariable sites (I=0.6918) and heterogeneous rate of substitution following a gamma distribution with alpha shape α=0.7750. According to the hierarchical likelihood ratio tests (hLRT), the HKY + I + G model of evolution (-LnL=2480.8875) (Hasegawa *et al*. 1985; Yang 1993; Gu *et al*. 1995) was selected. The parameters estimations were I=0.6904 and α=0.7714. Neighbour-Joining analyses of the ML distances obtained using the parameter estimates derived from each substitution model were performed. Bootstrap values for each NJ analysis were obtained from 1000 replications. Topology of NJ tree obtained using ML distances was similar to the one derived from MP analyses (fig. 2).

As already pointed out by Sezonlin *et al*. (2006), the three conspicuous clades of individuals and haplotypes are supported by bootstrap values exceeding 50% in both MP an NJ analyses. The smallest clade *KI* comprised 16 haplotypes and 70 individuals, all of which came from East Africa. The clade *KII* comprised 61 haplotypes and 280 individuals and had the largest distribution from East to Central Africa via Southern Africa. Finally, the clade *W* comprised 31 haplotypes and 139 individuals and was found only in West Africa (fig. 1). No haplotype was shared between West African populations and East-Central-Southern African populations. Both in MP and NJ analyses, *W* and *KI* were supported by high bootstrap values whereas *KII* was supported by lower bootstrap values. The phylogenetic relationships between these three major clades remain unresolved. The sister group status of clades *W* and *KI* was observed only in NJ analyses using the substitution models selected by hLRT and AIC criterion. In both NJ analyses, the bootstrap values remain low (55% and 56% respectively for hLRT and AIC).

Genetic structure of the *Busseola fusca* populations

Most of the molecular variation was accounted for by the differentiation between the three clades highlighted by phylogenetic analyses with Φ_{ST} = 0.868 ($P<10^{-5}$). At fine scale, the genetic structure was observed within local populations in each major clade. The different values were 0.226 ($P<10^{-5}$), 0.238 ($P<10^{-5}$), 0.344 ($P<10^{-5}$) respectively for *W*, *KI* and *KII* clades.

Diversity and demographic history of *Busseola fusca*

Haplotype and nucleotide diversity were calculated for different groups considered within each clade identified (tab. 1). The values of these indices vary greatly within each clade. For clade *W*, the forest region has the highest haplotype diversity. The East-North region is more diverse for clade *KI* than for other regions. Finally for *KII*, the East-South has the highest haplotype diversity. Both the variance (SSD) and raggedness index (*r*) tests suggested that the curves (figs 3a, 3b, 3c) do not significantly differ from the distribution under a model of population expansion (P_{SSD} = 0.41 and P_r = 0.69 for *W*; P_{SSD} = 0.10 and P_r = 0.27 for *KI*; P_{SSD} = 0.87 and P_r = 0.94 for *KII*). Similarly, the negative values obtained for Tajima's D index for each clade (-1.62105; -1.63863; -1.49025 for clades *W*, *KI* and *KII*, respectively) are all consistent with the hypothesis of population expansion since the origin of the clades. The current geographic distribution of *B. fusca* with centre of origin of each clade is illustrated fig. 4.

Table 1. Estimates of haplotype and nucleotide diversity for different population groupings of *B. fusca*. Forest region populations have the highest diversity for clade *W*, East-North region populations have the highest diversity for clade *KI*, East-South region populations have the highest diversity for clade *KII*.

Clade	Region	Haplotype diversity	Nucleotide diversity (%)
W	Forest (drier types)	0.879 +/- 0.024	0.254 +/- 0.155
	Forest and secondary grassland	0.772 +/- 0.094	0.209 +/- 0.137
	Savannah	0.659 +/- 0.072	0.154 +/- 0.105
KI	East – North	0.897 +/- 0.041	0.209 +/- 0.138
	East - South	0.789 +/- 0.032	0.166 +/- 0.110
KII	East - North	0.495 +/- 0.151	0.256 +/- 0.165
	East – South	0.862 +/- 0.021	0.379 +/- 0.213
	Austral	0.720 +/- 0.069	0.108 +/- 0.080
	Cameroon	0.199 +/- 0.112	0.029 +/- 0.020

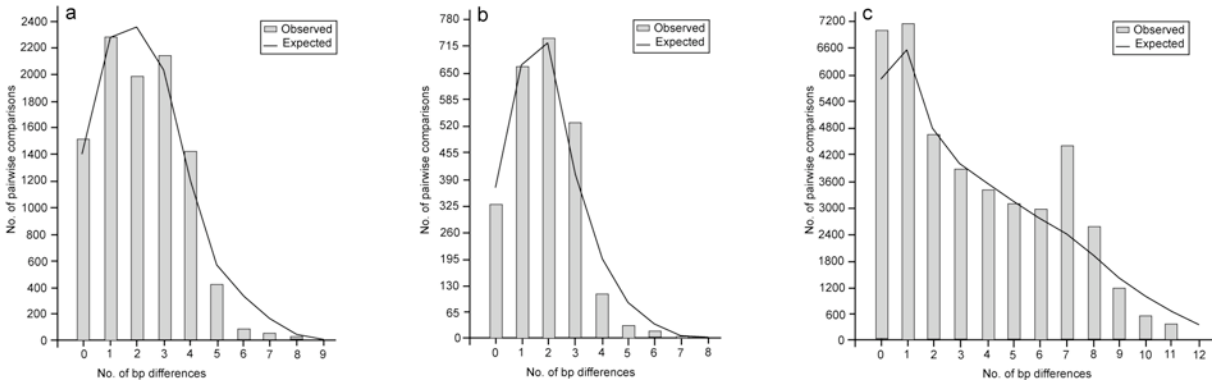


Figure 3. Mismatch distribution analysis showing histogram of observed and expected mismatch frequencies (a: West Africa population; b: Kenya I population; c: Kenya II population).

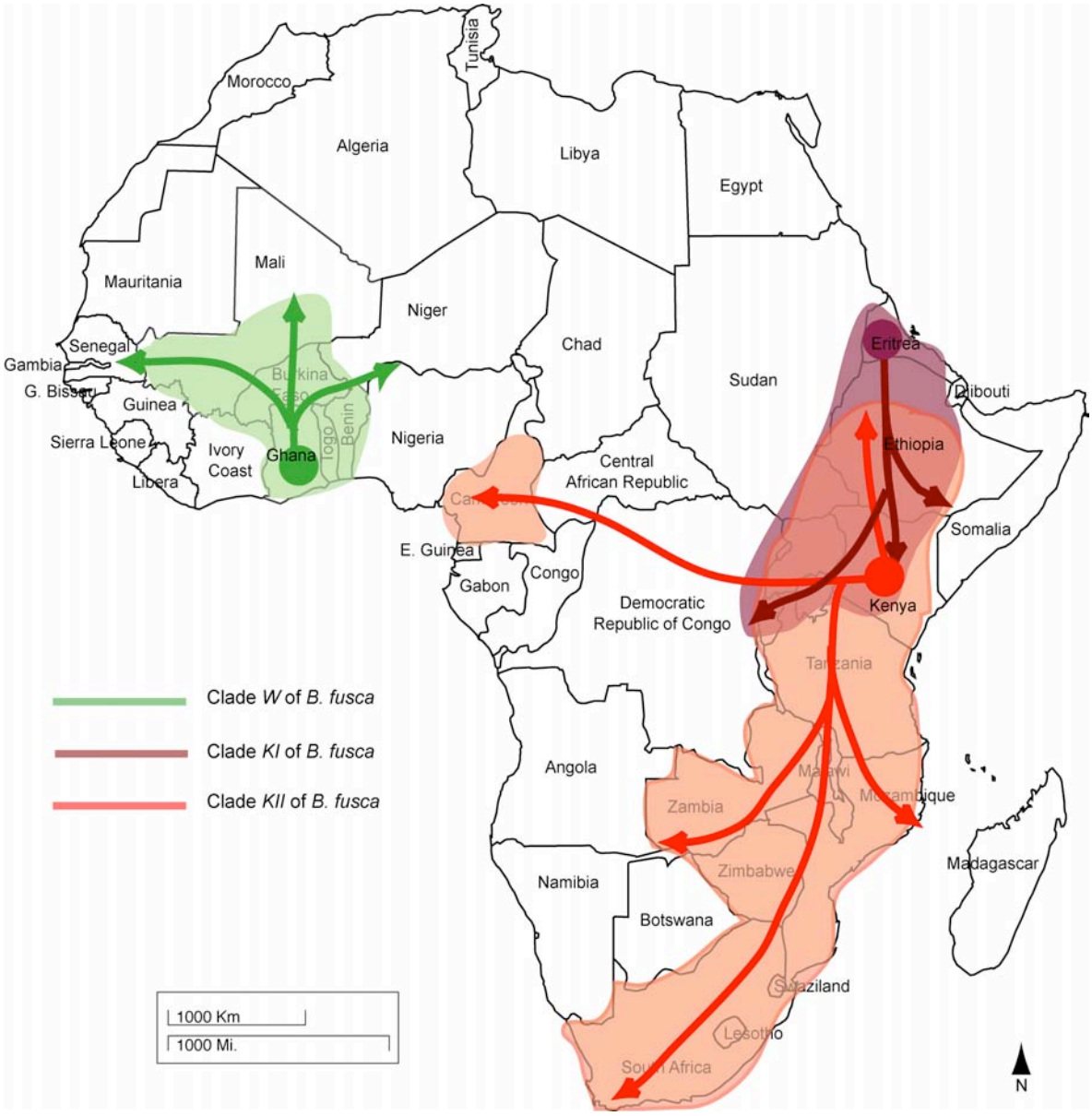


Figure 4. Current geographic distribution of *B. fusca* and the different putative centres of origin of its three main clades highlighted by phylogenetic analyses.

Nested clade phylogeographic analysis

The NCPA network calculation identified the same three clades revealed by MP and NJ analyses. The networks of these three clades were represented by figs 5a, 5b, 5c. The West African clade contained the 31 haplotypes observed exclusively in this region and four hypothetical intermediate haplotypes, hierarchically grouped into 11 'one-step clades', four 'two-step clades', and two 'three-step clades'. The *KI* clade contained the 16 observed haplotypes and four hypothetical haplotypes: 8 'one-step clades' and 2 'two-step clades'. Finally, the *KII* clade, with 61 observed haplotypes and 10 hypothetical is organized into 20 'one-step' clades, 6 'two-step clades' and 2 'three-step clades'.

Nested contingency analysis on the haplotype network revealed significant geographic associations in the three major networks at all clade levels. These significant values were interpreted using Templeton's (2004) inference key (tab. 2). Most of the clades displaying geographic associations were interpreted by restricted gene flow with isolation by distance although some of them were interpreted as restricted gene flow or dispersal but with some long distance dispersal and one by contiguous range expansion.

Table 2. Inference chain results of geographical distance analysis from Fig. 5a, 5b, 5c.

Clade	Chain of inference	Inference
1 – 1 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 2 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 4 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 6 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 10 (*W*)	1-2-11-17: NO	Inconclusive outcome
2 – 1 (*W*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal
2 – 2 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
2 – 3 (*W*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
2 – 4 (*W*)	1-2-11-12: NO	Contiguous range expansion
3 – 2 (*W*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal
1 – 2 (*KI*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal
2 – 1 (*KI*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal
1 – 3 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 5 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 18 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
1 – 20 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
2 – 2 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
2 – 3 (*KII*)	1-2-11-17: NO	Inconclusive outcome
2 – 6 (*KII*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal
3 – 1 (*KII*)	1-2-3-4: NO	Restricted gene flow with isolation by distance
3 – 2 (*KII*)	1-2-3-5-6-7: YES	Restricted gene flow / Dispersal but with some long distance dispersal

Discussion

The goal of this study was to reconstruct the phylogeographic pattern of *Busseola fusca* across the whole geographic range of the species, to evaluate the current geographic distribution of each clade and to determine their centres of origin.

Genetic structure and phylogeographic pattern of *B. fusca*

All phylogenetic analyses confirm the separation of *B. fusca* into three major clades corresponding to three geographical units: one localized in West African region (*W*), one restricted to East Africa (*KI*) and one found from Central to East Africa via Southern Africa (*KII*). Partial geographic overlap was observed only between the clades *KI* and *KII*. The genetic distances between the clades suggest that the differentiation occurred during the Pleistocene (Sezonlin *et al*. 2006). Major climatic changes occurred in sub-Saharan Africa during the Pleistocene. The period of climatic instability started 3.3 to 2.45 Ma, oscillating between hot/humid and cooler, drier periods (Wagner 2002). A shift to arid, open conditions occurred in near 2.8 Ma, 1.7 Ma, and 1.0 Ma (de Menocal 1995). de Menocal (1995) concluded that this alternation of cold, dry periods and warmer, wetter periods led to oscillations in savannah biotope expansion. Pleistocene events also played an important role in differentiation of African vertebrates (Quérouil *et al.* 2003). The present study confirmed that *B. fusca* populations are differentiated in three clades that were isolated during Pleistocene in three different refuges (Sezonlin *et al.* 2006). This hypothesis was also supported by genetic structure analyses that show high values of the fixation index between the three clades. The fragmentation hypothesis can be inferred between West African and other clades that do not overlap at all, but is questionable between the *KI* and *KII* clades that overlap in wide areas. According to Templeton *et al*. (1995), only mainly non-overlapping populations can clearly be inferred as a product of past fragmentation.

All the mitochondrial variation revealed significant and historical separation between clades *W*, *KI* and *KII*. The monophyly of the western populations allows the designation of at least one Evolutionary Significant Unit (ESU) (Alpers *et al*. 2004) within the *B. fusca* species. It is the same for *KI* and *KII* that can be considered as ESUs. If these clades are confirmed by nuclear polymorphism analyses (microsatellites data), pheromones and behavioural ecology, therefore different biological strategies must be used to control *B. fusca* populations.

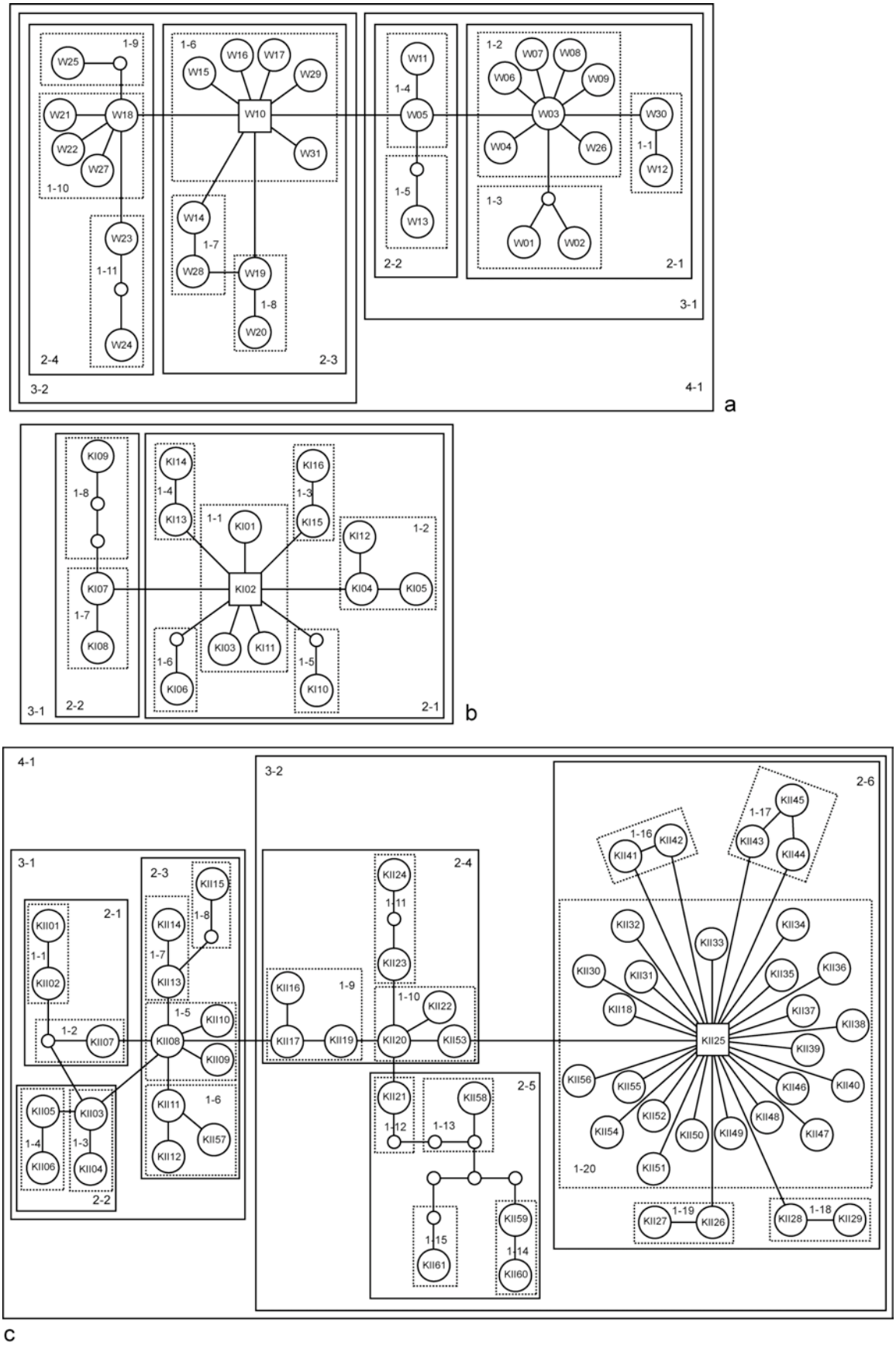


Figure 5. Haplotype network of all haplotypes detected for *Busseola fusca*. Each haplotype is labelled by its number. Hypothetical haplotypes are designated by small circles (A: West Africa population; B: Kenya I population; C: Kenya II population).

Origin and evolutionary history of each clade

In the present *B. fusca* study, the population of forest region has higher haplotype and nucleotide diversity for clade *W* than other phytogeographic regions of West Africa. It would likely be the centre of origin for this clade. Alpers *et al*. (2004) found also a center of origin in Ghana for the West African populations of the roan antelope (*Hippotragus equinus*) (Desmarest 1804), a herbivorous species that has a sub-Saharan distribution similar to that of *B. fusca*. Concerning the *KI* clade, the eritrean region, where only *KI* individuals were found was more diverse than other regions. Accordingly, we will consider it as the possible/putative centre of origin for clade *KI*. Finally for clade *KII*, the highest haplotype and nucleotide diversity was found among population of South-East region. The centre of origin of this clade is likely in this geographic area. Consequently, we can suggest that the centres of origin were likely localized in forest, East-North and East-South regions respectively for clades *W*, *KI* and *KII*. We observed that the strongest haplotype diversity was associated with the lowest levels of nucleotide diversity. This accumulation of haplotypes suggests that the clades experienced bottlenecks at their origins, followed by major population demographic expansion (Grant & Bowen 1998; Avise 2000). However, Petit *et al*. (2003) have shown that for some European tree species the highest diversity is not observed in the centre of origin but rather in secondary contact zones. Therefore in *B. fusca* case, further studies will be necessary to confirm the reality of these possible centres of origin. The fact that clades *KI* and *KII* originate from East Africa gives support to some previous studies (Livingstone 1982, Arctander *et al*. 1999). Indeed, these authors have described the East African region as a mosaic of secondary refuge zones for herbivorous mammals, with periodic exchanges between refuge zones through temporary contact bridges in the East African Rift Valley. Although the East African populations of *B. fusca* are now overlapping, the Rift Valley was pointed out as one of the main factors that explain most of the molecular variation in East Africa (Sezonlin *et al*. 2006). The Rift Valley seems to act as an important natural barrier to maintain population structure for other African species in East Africa (Arctander *et al*. 1999, Pitra *et al*. 2002).

The current distribution of *B. fusca* populations can be explained by contiguous range expansion or by dispersal with some long distance dispersal as it is highlighted by NCPA inferences. This pattern is probably linked to the expansion of its wild host plant (*Sorghum arundinaceum* (Desv. Stapf) (Poaceae) (Haile & Hofsvang 2002) during the Pleistocene. The presence of individuals of the large population unit, clade *KII* in Central Africa (Cameroon) is consistent with the hypothesis of a faunistic link between these two regions (Bruhl 1997),

which are separated by a distance of 3000 km. An eastern origin of central populations is a possibility as it was suggested for a butterfly species in Cameroon (De Jong & Congdon 1993). De Jong & Congdon (1993) argued that the low animal species diversity in highland forests of Cameroon suggests that these species originated from long distance migration from East Africa. The faunistic link between Eastern and Central Africa also exists for some vertebrate species (Pitra *et al*. 2002). However, the nature of this faunistic link was not elucidated by our study and remains unknown for *B. fusca*. The present study shows that the clade *KII* is also present in Southern Africa. This geographic expansion toward Southern Africa is consistent with the patterns highlighted by the study of some African vertebrates (Faulkes *et al*. 2004). Climatic and topographic differences between major biogeographic African regions (de Menocal 1995) might explain the different processes that govern the current geographic distribution of each *B. fusca* clade.

Evolution of host plant specialization in *B. fusca*

Many ecological studies have shown that the introduction of exotic plants can lead to host plant shifts in oligo- or monophagous insect species. This is the case for *Rhagoletis pomonella* (Walsh 1867) (Diptera: Tephritidae) which in North America switched from hawthorn to introduced apple tree (Bush 1994); the bug *Jadera haematoloma* (Herrich-Schaeffer 1847) (Hemiptera: Rhopalidae), which added to its host plant spectrum an introduced ornamental plant of the family Sapindaceae (Carroll & Boyd 1992). Other examples of such host switches are the nymphalid *Euphydrias editha* (Boisduval 1852) (Lepidoptera: Nymphalidae) which switched to *Plantago lanceolata* L. (Plantaginaceae), introduced by North American breeders (Singer *et al*. 1993), the groundnut beetle *Caryedon serratus* (Olivier 1790) (Coleoptera: Bruchidae) which added groundnut, *Arachis hypogea* L. (Fabaceae), an introduced Papilionoideae to its native host plant range (Delobel 1995) and finally African cereal stem borers such as *B. fusca* and *Sesamia calamistis* Hampson 1910 (Lepidoptera, Noctuidae) which shifted to maize after its introduction into Africa.

Host plant shifts appear as a major factor promoting ecological specialization in phytophagous insects. Within an insect population on a given host plant, some genotypes procuring better fitness to their bearers can be selected. This phenomenon may occur within the various populations of an insect species associated with different host plants and leads to genetic differentiation between these populations. A study by Via *et al*. (2000) on the aphid *Acyrthosiphon pisum* (Harris 1776) (Hemiptera: Aphididae) that feed on two different legume

species shows that direct selection against migrants and hybrids in each parental environment could favour the evolution of more precise or efficient habitat choices. Moreover, these authors argued that the differential phenology of host plants strengthens ecological specialization by increasing the reproductive isolation between populations. This phenomenon promotes mating within the same habitat. As is well established for some phytophagous insects, such host plant shifts can lead to ecological segregation of populations with appearance of host races and genetic differentiation. This is well known with *R. pomonella* (Feder *et al*. 1988; McPheron *et al.* 1988), *R. cerasi* (L. 1758) (Diptera: Tephritidae) (Schwartz *et al*. 2003), the European corn borer *Ostrinia nubilalis* (Hübner 1796) (Lepidoptera: Pyralidae) (Thomas *et al*. 2003), *Spodoptera frugiperda* (Smith & Abbott 1797) (Lepidoptera: Noctuidae) (Pashley Prowell *et al*. 2004), *C. serratus* (Sembène 2000) and the pea aphid *A. pisum* (Via *et al*. 2000; Caillaud & Via 2000). Genetic differentiation of such host races can occur within a limited period of time. For example, it has taken a few centuries for *R. pomonella* and *C. serratus* populations to differentiate, several tens of years for *J. haematoloma* and less than 20 years for *E. editha*. Considering the age of sorghum domestication and of maize introduction in Africa, it is thus likely that these major agricultural events could have been at the origin of such a phenomenon of genetic divergence among *B. fusca* populations and possibly at the origin of host race appearance in this species. The fact that *B. fusca*, considered as oligophagous (Le Rü *et al.* 2006) now preferentially uses maize and cultivated sorghum in most part of its distribution areas suggests that these host plants offer this species very suitable resources. The preference for these cultivated plants (Kfir *et al*. 2002, Le Rü *et al*. 2006) would represent one of the factors that might have led *B. fusca* to start ecological specialization between populations. This ecological segregation between *B. fusca* populations exploiting wild and cultivated host plants or cultivated sorghum and maize could be highlighted by further molecular studies. However, a preliminary population genetic study using mitochondrial marker (Sezonlin *et al.* 2006) has not demonstrated any genetic structure between cultivated sorghum and maize *B. fusca* populations. An extensive survey carried out in East and West Africa indicates that *B. fusca* is now mostly associated with maize and sorghum crops and is generally uncommon in the wild habitat (Le Rü *et al*. 2006; Sezonlin *et al*. unpublished) Therefore, we have not been able until now to test ecological segregation between wild host plants and cereal crops, but the recent discovery of *B. fusca* populations associated to *Phragmites mauritianus* Kunth. (Poacaea) in Ethiopia and Eritrea and to *Setaria megaphylla* (Steud.) (Poaceae) T. Duran & Schinz in Kenya (Le Rü *et al*. 2006) will soon help us to test this hypothesis.

Acknowledgements

This study was supported by IRD UR072. We are grateful to Dr. G. Goergen, Collection Manager for IITA Biodiversity Centre - Station Cotonou in Benin, for his support. We would like to thank the members of the ICIPE Noctuid Stem Borer Biodiversity Project (NSBBP) mainly L. Ngala, B. Musyoka, J. Mungai and H. Njambi for technical assistance during insect collection and insect rearing in the laboratory in Kenya. Anonymous referee also helped greatly in improving this manuscript.

References

Ajayi O. 1998. Sorghum: West Africa, p. 39-45 *in*: **Polaszek A. (ed.)**, *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. CABI, Wallingford, UK.

Alpers D.L., Van Vuuren B.J., Arctander P., Robinson T.J. 2004. Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology* **13**: 1771-1784.

Althoff M.D., Pellmyr O. 2002. Examining genetic structure in bogus yucca moth: A sequential approach to phylogeography. *Evolution* **56**: 1632-1643.

Arctander P., Johansen C., Coutellec-Vreto M. 1999. Phylogeography of three closely related African bovids (Tribe *Alcelaphini*). *Molecular Biology and Evolution* **16**: 1724-1739.

Avise J.C. 2000. *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts, 447 p.

Bernatchez L. 2001. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution* **55**: 351-379.

Bruhl C.A. 1997. Flightless insects: a test case for historical relationships of African mountains. *Journal of Biogeography* **24**: 233-250.

Bush G.L. 1994. Sympatric speciation in animals: new wine in old bottles. *Trends in Ecology and Evolution* **9**: 285-288.

Caillaud C.M., Via S. 2000. Specialized feeding behavior influences both ecological specialization and assortative mating in sympatric host races of pea aphids. *The American Naturalist* **156**: 607-621.

Carroll S.P., Boyd C. 1992. Host race radiation in the soapberry bug: Natural history with the history. *Evolution* **46**: 1052-1069.

Chastanet M. 1998. Introduction et place du maïs au Sahel Occidental (Sénégal-Mauritanie), p. 251-275 *in*: **Chastanet M. (ed.)**, *Plantes et paysages d'Afrique. Une histoire à explorer*. Karthala-CRA, Paris.

Clement M., Posada D., Crandall K.A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657-1660.

De Jong R., Congdon T.C.E. 1993. The montane butterflies of the eastern Afrotropics, p. 133-164 *in*: **Lovett J. C., Wasser S. K. (eds.)**, *Biogeography and ecology of the rain forests of eastern Africa*. Cambridge University Press, Cambridge.

Delobel A. 1995. The shift of *Caryedon serratus* (Ol.) from wild Caesalpiniaceae to groundnuts took place in West Africa (Coleoptera: Bruchidae). *Journal of Stored Products Research* **31**: 101-102.

de Menocal P.B 1995. Plio-Pleistocene African climate. *Science* **270**: 53-59.

Doggett H. 1988. *Sorghum.* 2nd ed, Longman Scientific and Technical Publishers, London, 512 p.

Excoffier L., Smouse P.E., Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479-491.

Faulkes C.G., Verheyen E., Verheyen W., Jarvis J.U.M., Bennett N.C. 2004. Phylogeograpical patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae). *Molecular Ecology* **13**: 613-629.

Feder J.L., Chilcote C.A., Bush G.L. 1988. Genetic differentiation between sympatric host races of the apple maggot fly, *Rhagoletis pomonella*. *Nature* **336**: 61-64.

Grant S.W., Bowen B.W. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89**: 415-426.

Gu X., Fu Y.X., Li W.H. 1995. Maximum likelihood estimation of heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution* **12**: 546-557.

Haile A., Hofsvang T. 2001. Survey of lepidopterous stem borers of sorghum, maize and pearl millet in Eritrea. *Crop Protection* **20**: 151-157.

Haile A., Hofsvang T. 2002. Host plant preference of the stem borer *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae). *Crop Protection* **21**: 227-233.

Hasegawa M., Kishino H., Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **21**: 160-174.

Holloway J. D. 1998. Noctuidae, p. 79-86 *in*: **Polaszek A. (ed)**, *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. CABI, Wallingford, UK.

Kfir R., Overholt A.W., Khan R.Z., Polaszek A. 2002. Biology and management of economically important lepidopteran cereal stem borers in Africa. *Annual Review of Entomology* **47**: 701-731.

Kimura M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences of the USA* **78**: 454-458.

Le Rü B.P., Ong'amo G.O., Moyal P., Muchugu E., Ngala L., Musyoka B., Abdullah Z., Kauma-Matama T., Lada V.Y., Pallangyo B., Omwega C.O., Schulthess F., Calatayud P.-A., Silvain J.-F. 2006. Major ecological characteristics of East African noctuid stemborers. *Annales de la Société Entomologique de France* 42 (3-4) XXX-XXX.

Livingstone D.A. 1982. Quaternary geography of Africa and refuge theory, p. 523-536 *in*: **Prance G. T. (ed.)**, *Biological diversification in the tropics: Proceedings of the 5th international symposium of the association for tropical biology*. Columbia University Press, New York.

Maddison D.R., Maddison W.P. 2002. *MacClade: Analysis of phylogeny and character evolution.* Version 4.02. Sinauer Associates, Sunderland, Massachusetts.

Madeiras Santos M.E., Ferraz Torrao M.M. 1998. Entre l'Amérique et l'Afrique, les îles du Cap-vert et Sao Tomé: les cheminements des milhos (mil, sorgho et maïs), p. 69-83 *in*: **Chastanet M. (ed.)**, *Plantes et paysages d'Afrique. Une histoire à explorer*. Karthala-CRA. Paris.

McPheron B.A., Smith C.D., Berlocher S.H. 1988. Genetic differences between host races of *Rhagoletis pomonella*. *Nature* **336**: 64-66.

Moyal P. 1998. Infestation patterns and parasitism of the maize stalk borer, *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae) in the Ivory Coast. *African Entomology* **6**: 289-296.

Murty D.S., Renard C. 2001. Sorgho (*Sorghum bicolor* (L.) Moench), p. 105-125 *in*: **Raemaekers R. H. (ed.),** *Agriculture en Afrique Tropicale*. Direction Générale de la Coopération Internationale, Bruxelles, Belgique.

Ndemah R., Schulthess F., Korie S., Borgemeister C., Cardwell K.F. 2001. Distribution, relative importance and effect of lepidopterous borers on maize yields in forest zone and mid-altitude of Cameroon. *Journal of Economic Entomology* **94**: 1434-1444.

Nei M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 p.

Nei M., Li W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the USA* **76**: 5269-5273.

Pashley Prowell D., McMichael M., Silvain J.-F. 2004. Multilocus genetic analysis of host use, introgression, and speciation in host strains fall armyworm (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America* **97**: 1034-1044.

Petit R.J, Aguinagalde I., de Beaulieu J-L., Bittkau C., Brewer S., Cheddadi R., Ennos R., Fineschi S., Grivet D., Lascoux M., Mohanty A., Müller-Starck G., Demesure-Musch B., Palmé A., Martín J.P., Rendell S., Vendramin G.G. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563-1565.

Pitra C., Hansen J.A., Lieckfeldt D., Arctander P. 2002. An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology* **11**: 1197-1208.

Polaszek A., Khan Z.R. 1998. Host plants, p. 3-10 *in*: **Polaszek A. (ed.)**, *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. CABI, Wallingford, UK.

Posada D., Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.

Posada D., Crandall K.A., Templeton A.R. 2000. GeoDis: A program for the cladistic nested analysis of geographical distribution of genetic haplotypes. *Molecular Ecology* **9**: 487-488.

Quérouil S., Verheyen E., Dillen M., Colyn M. 2003. Patterns of diversification in two African forest shrews: *Sylvisorex johnstoni* and *Sylvisorex ollula* (Soricidae, Insectivora) in relation to paleo-environmental changes. *Molecular Phylogenetics and Evolution* **28**: 24-37.

Schneider S.D., Excoffier L. 1999. Estimation of past demographic parameters from the distribution of pairwise distances when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**: 1079-1089.

Schneider S.D., Roessli D., Excoffier L. 2000. *Arlequin V. 2.000: a software for genetic data analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Schwartz D., McPheron A.B., Hartl B.G., Foller F.E., Hoffmeister S.T. 2003. A second case of genetic host races in *Rhagoletis*? A population genetic comparison of sympatric host populations in the European cherry fruit fly, *Rhagoletis cerasi*. *Entomologia Experimentalis et Applicata* **108**: 11-17.

Sembène M. 2000. *Variabilité de l'Espaceur Interne transcrit (ITS1) de l'ADN ribosomique et polymorphisme des locus microsatellites chez la bruche d'arachide Caryedon serratus (Olivier) : différenciation en races d'hôtes et infestation de l'arachide au Sénégal*. Thèse de doctorat d'Etat ès sciences, Université Cheik Anta Diop de Dakar, Dakar, 214p.

Sezonlin M., Dupas S., Le Rü B., Le Gall P., Moyal P., Calatayud P.-A., Giffard I., Faure N., Silvain J.-F. 2006. Phylogeography and population genetics of the maize stalk borer *Busseola fusca* (Lepidoptera, Noctuidae) in sub-Saharan Africa. *Molecular Ecology* **15**: 407-420.

Simmons R.B., Weller S.J. 2001. Utility and evolution of cytochrome *b* in insects. *Molecular Phylogenetics and Evolution* **20**: 196-210.

Singer M.C., Thomas C.D., Parmesan C. 1993. Rapid human-induced evolution of insect-host associations. *Nature* **366**: 681-683.

Slatkin M., Hudson R.R. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555-562.

Swofford D.L. 2002. PAUP*. *Phylogenetic Analysis Using Parsimony (*and other methods)*. Sinauer Associates, Sunderland, Massachusetts.

Tajima F. 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.

Tajima F. 1989b. The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597-601.

Templeton A.R., Routman E., Phillips C.A. 1995. Separating population structure from population history: a cladistic analysis of the geographical dispersion of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140**: 767-782.

Templeton A.R. 1998. Nested clade analysis of phylogeographic data: testing hypothesis about gene flow and population history. *Molecular Ecology* **7**: 381-397.

Templeton A.R. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**: 789-809.

Thomas Y., Bethenod M.T., Pelozuelo L., Frérot B., Bourguet D. 2003. Genetic isolation between two sympatric host-plant races of the European corn borer, *Ostrinia nubilalis* Hübner. I. Sex pheromone, moth emergence timing, and parasitism. *Evolution* **57**: 261-273.

Via S., Bouk C.A., Skillman S. 2000. Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution* **54**: 1626-1637.

Wagner T. 2002. Late Cretaceous to early Quaternary organic sedimentation in the eastern Equatorial Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology* **179**: 113-147.

White F. 1983. The vegetation of Africa, a descriptive memoir to accompany the UNESCO / AETFAT / UNSO vegetation map of Africa. *UNESCO, Natural Resources Research* **20**: 1-356.

Yang Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution* **10**: 1396-1402.

Section II : Etudes complémentaires des populations de *B. fusca* du Cameroun

Le Cameroun est une zone de contact entre deux unités populationnelles apparemment très hermétiques au niveau des lignées mitochondriales de *B. fusca*. Elle mérite une étude plus approfondie. L'obtention à la fin de l'année 2005 de 101 individus supplémentaires de *B. fusca* provenant de plusieurs zones écologiques du Cameroun a permis d'analyser à nouveau la diversité moléculaire de certaines populations et d'étudier de manière plus fine la structure génétique des populations de *B. fusca* dans les 3 pays où l'effort d'échantillonnage a été le plus grand. Il s'agit du Bénin (79 individus de *B. fusca*), du Cameroun (120 individus de *B. fusca*) et du Kenya (230 individus de *B. fusca*).

2.1. La diversité moléculaire

L'analyse de l'ensemble des échantillons camerounais disponibles avec le programme TCS 1.21 a mis en évidence 11 haplotypes dont 8 correspondent à des haplotypes déjà identifiés dans le clade *KII* et 3 à des haplotypes nouveaux (KII62 ; KII63 ; KII64). Le tableau 4-1 fait la synthèse sur ces 11 haplotypes camerounais. Ils se regroupent avec les autres haplotypes du clade *KII*. Cette analyse confirme également la présence du clade *KII* en Afrique Centrale et l'ampleur de son extension géographique.

L'haplotype KII25 identifié comme haplotype ancestral du clade *KII* dans les études antérieures est reparti géographiquement dans toutes les localités camerounaises. Les localités Kaiy et Tatum, correspondant pratiquement à un même site d'échantillonnage et possédant le plus grand nombre d'haplotypes mitochondriaux, sont les plus diversifiées génétiquement. Ces deux localités appartiennent aux régions d'altitude élevée.

L'intégration de ces nouvelles données moléculaires dans le jeu de données totales a permis d'estimer à nouveau avec le programme ARLEQUIN v.2.000 la diversité haplotypique et nucléotidique de *B. fusca* pour le clade *KII* en considérant les grandes régions géographiques retenues au cours de l'analyse du deuxième article. Les résultats sont consignés dans le tableau 4-2. Comme on peut le remarquer à travers la lecture ce tableau, par rapport à la majorité de régions considérées, la diversité haplotypique et nucléotidique reste relativement faible au Cameroun alors que le nombre de localités échantillonnées et le nombre d'individus analysés a augmenté. Ces résultats confirment l'hypothèse selon laquelle

le centre d'origine de ce clade se situerait dans la région kenyane. La colonisation du Cameroun, situé géographiquement à l'Ouest de l'Afrique Centrale et environ à 3.000 km du Kenya, apparaît comme un événement récent. Des analyses complémentaires avec des échantillons de *B. fusca* en provenance des sites successifs de l'Est à l'Ouest de la République Démocratique du Congo sont nécessaires pour étayer la perte progressive de la diversité génétique, révélatrice d'une expansion géographique continue à partir du Kenya vers le Cameroun. Cette étude a, en outre, révélé l'expansion géographique considérable (présence en Afrique Orientale, Australe et Centrale) du clade *KII* qui se comporterait comme une population "invasive" au détriment des clades *KI* et *W* bien que les limites géographiques de ce dernier restent encore imprécises. Malgré le défaut d'échantillonnage de ce foreur de céréales dans certains pays Ouest africains (Nigeria, Côte-d'Ivoire, Liberia, Sierra Leone, Guinée-Bissau, Guinée Conakry, Sénégal), l'aire de répartition du clade *KII* reste très nettement supérieure à celle des clades *W* (localisé seulement en Afrique Occidentale) et *KI* (confiné sur une partie d'Afrique de l'Est).

Tableau 4-1 : Le nombre d'individus de *B. fusca* analysés et les haplotypes identifiés par localité camerounaise.

Pays	Localités	Individus analysés	Haplotypes identifiés
Cameroun	Akometam	3	KII25
	Bali	2	KII25
	Bilik Bindzik	1	KII25
	Binshua	4	KII25
	Kaiy	18	KII25 ; KII28 ; KII32 ; KII62 ; KII63 ; KII64
	Kikaikilaiki	21	KII25 ; KII28 ; KII37
	Kitiwum	16	KII25 ; KII28 ; KII63
	Mah	11	KII25 ; KII28 ; KII37 ; KII51
	Mvoutessi	3	KII25
	Njavnyuy	2	KII25 ; KII28
	Nkar	2	KII25
	Nkolmelen	1	KII25
	Nkometou	17	KII25 ; KII42
	Ntongee	4	KII25 ; KII37
	Takui	8	KII25 ; KII28 ; KII40
	Tatum	6	KII25 ; KII32 ; KII37 ; KII53 ; KII63
	Wainama	1	KII25

Tableau 4-2 : Estimation de la diversité haplotypique et nucléotidique des différents groupes d'individus constituant le clade *KII*. La région kenyane a toujours les plus fortes valeurs de diversité haplotypique et nucléotidique alors la région camerounaise présente les plus faibles valeurs de ces indices de diversité.

Clade	Régions	Diversité haplotypique	Diversité nucléotidique (%)
KII	Est - Nord	0,495 +/- 0,151	0,256 +/- 0,165
	Est - Sud	0,862 +/- 0,021	0,379 +/- 0,213
	Australe	0,720 +/- 0,069	0,108 +/- 0,080
	Cameroun	0,483 +/- 0,054	0,073 +/- 0,061

2.2. Structure génétique fine des populations de *B. fusca*

La disponibilité d'un grand nombre d'échantillons de *B. fusca* récoltés dans de nombreuses localités du Bénin, du Cameroun et du Kenya rend possible cette analyse. Elle a conduit à étudier et à comparer le degré de différenciation des populations locales de ces trois pays. Les différentes valeurs des indices de diversité génétique sont consignées dans le tableau 4-3.

Tableau 4-3 : Les différentes valeurs des indices de diversité génétique des populations de *B. fusca* au Bénin, au Cameroun et au Kenya. Φ_{ST} indique la différenciation génétique entre les populations locales considérées, Φ_{SC} indique le degré de différenciation génétique au sein des groupes phytogéographiques ou agro-écologiques considérés alors que le Φ_{CT} indique la différenciation génétique entre ces groupes.

Pays	Φ_{ST}	Φ_{SC}	Φ_{CT}
Bénin	0,145 *($P<10^{-3}$)*	0,128 *($P<10^{-3}$)*	0,019 *($P<10^{-3}$)*
Cameroun	0,034 *($P<10^{-3}$)*	-0,017 *($P<10^{-3}$)*	0,051 *($P<10^{-3}$)*
Kenya	0,352 *($P<10^{-3}$)*	0,256 *($P<10^{-3}$)*	0,129 *($P<10^{-3}$)*

Il existe une différenciation génétique entre les populations locales béninoises et kenyanes de *B. fusca*. Ce ravageur de céréales présente au Kenya une différenciation génétique au sein et entre les groupes agro-écologiques considérés. Par contre au Bénin, on observe qu'une structuration au sein des groupes phytogéographiques considérés. Toujours dans ce pays, il n'y a pratiquement pas de différenciation génétique entre ces groupes phytogéographiques. On constate que les différentes valeurs des indices (Φ_{ST}, Φ_{SC}, Φ_{CT}) de diversité génétique sont plus élevées au Kenya qu'au Bénin. Au Cameroun, on n'observe aucune différenciation génétique aux différents niveaux hiérarchiques considérés (populationnel, au sein et entre groupes agro-écologiques).

2.3. Analyse phylogéographique des clades emboîtés de *B. fusca*

Cette nouvelle analyse a permis d'affiner les inférences obtenues jusqu'ici. Comparativement au jeu de données de l'article 2, deux nouveaux clades supplémentaires (1-10 et 1-17 en rouge sur la figure 4-1) se sont ajoutés à ceux du rang 1. Ces clades sont bien intégrés au seul clade 2-4. Les trois nouveaux haplotypes propres au Cameroun sont figurés en vert. L'interprétation des clades au différents niveaux hiérarchiques à l'aide de la clé d'inférence la plus récente (key, 11/ 11/ 2005) a permis de déduire que trois clades (1-22 ; 2-4 ; 3-2) s'interprètent par le phénomène de restriction de flux géniques ou, alternativement, par la dispersion avec la colonisation à longue distance alors que le cladogramme total s'interprète toujours par le phénomène de la restriction de flux génique avec isolement par la distance. Les anciens clades 1-11 et 2-4 ont désormais trouvé leur interprétation via cette analyse qui prend en compte un échantillonnage plus important. Toutes les inférences obtenues pour cette analyse du clade *KII* sont consignées dans le tableau 4-4.

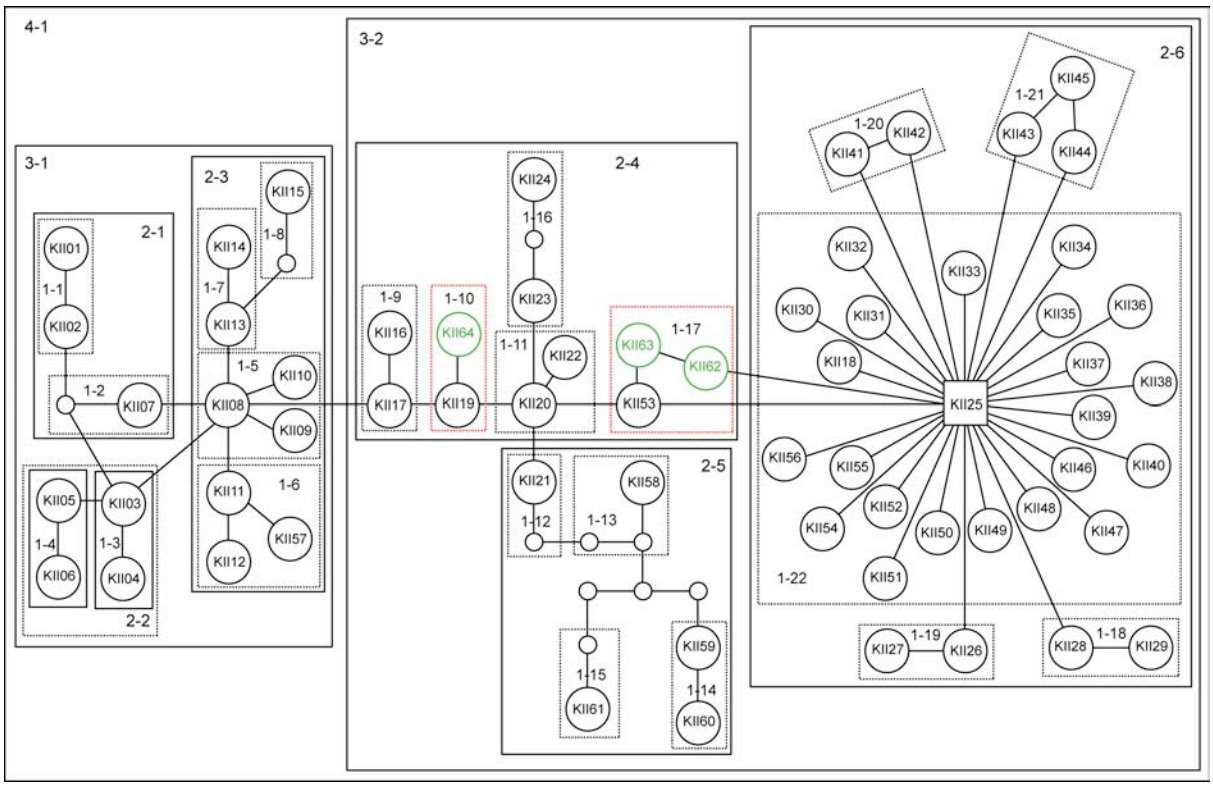


Figure 4-1 : Nouveau réseau parcimonieux des haplotypes mitochondriaux y compris tous ceux du Cameroun (de l'échantillonnage complémentaire) trouvés chez les individus de *B. fusca* du clade *KII*. Chaque haplotype est marqué par son nombre. Les trois nouveaux haplotypes propres au Cameroun sont mis en vert. Les haplotypes hypothétiques sont représentés par les petits cercles. Les nouveaux clades formés sont mis en rouge.

Tableau 4-4 : Résultats de la chaîne d'inférence de l'analyse des distances géographiques de la population mitochondriale *KII*.

Clade	Chaîne d'inférence	Inférence
1 - 3	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
1 - 5	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
1 -11	1-2-11-17-4 : NON	Restriction de flux de gènes avec isolement par la distance
1 - 18	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
1 - 22	1-2-3-5-6-7 : OUI	Restriction de flux de gènes / Dispersion avec colonisation à longue distance
2 - 2	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
2 - 3	1-2-11-17 : NON	Sans résultat concluant
2 - 4	1-2-3-5-6-7 : OUI	Restriction de flux de gènes / Dispersion avec colonisation à longue distance
2 - 6	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
3 - 1	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance
3 - 2	1-2-3-5-6-7 : OUI	Restriction de flux de gènes / Dispersion avec colonisation à longue distance
KII	1-2-3-4 : NON	Restriction de flux de gènes avec isolement par la distance

Cette analyse complémentaire confirme, en grande partie, les premiers résultats publiés dans les articles. La diversité génétique reste faible dans les zones récemment colonisées comme cela a été observé avec les populations camerounaises qui ne présentent aucune différenciation génétique. En général, *B. fusca* se caractérise par des phénomènes de restriction de flux de gènes avec isolement par la distance même si d'autres sont possibles.

Chapitre 5
Discussion générale et perspectives

La discussion générale sur l'étude phylogéographique de *B. fusca* apportera plus de précision et de détail sur les aspects qui n'ont pas été bien développés au cours des différentes analyses successives conduisant à la rédaction des articles. Elle abordera aussi les autres aspects non étudiés dans les articles tels que l'importance des informations biologiques déduites de l'étude d'un nouvel échantillon important (101 individus de *B. fusca*) provenant de Cameroun et les nouvelles hypothèses qui en découleront. Nous insisterons sur l'apport du modèle *B. fusca* à la compréhension de la structure génétique et de la phylogéograhie des ravageurs en général et ceux africains en particulier. Cette partie se terminera par les perspectives de recherche.

***Relations phylogénétiques entre les haplotypes et la structuration génétique des populations de* B. fusca**

Toutes les analyses phylogénétiques prenant en compte les différents haplotypes mitochondriaux identifiés chez *B. fusca* mettent clairement en évidence trois clades majeurs correspondant à des populations occupant des entités géographiques bien définies : un clade présent en Afrique de l'Ouest (*W*), un clade limité à quelques pays d'Afrique de l'Est (*KI*), un clade *KII* à large distribution géographique regroupant les pays d'Afrique de l'Est, du Centre et Australe. Les différentes analyses successives n'ont pas révélé de clade nouveau même avec un échantillonnage plus important et recouvrant une large partie de l'aire de distribution de *B. fusca*. La monophylie du clade *W* est confirmée par le soutien statistique fort (bootstrap) qui n'est pas sensible à l'ajout de nouvelles populations. Le clade *W* est séparé géographiquement des clades *KI* et *KII* qui sont eux partiellement sympatriques. On peut expliquer ce résultat, en première analyse, par le fait que les barrières naturelles telles que le massif forestier d'Afrique Centrale ou la Ligne Volcanique du Cameroun (LVC, ou CVL en anglais : il s'agit d'une chaîne de montagne de près de 1.700 km de long culminant à 4095 m d'altitude) réduiraient ou empêcheraient les flux géniques entre la population d'Afrique de l'Ouest et les populations d'Afrique Centrale, de l'Est et du Sud.

L'étude de la structure génétique fine des populations du *B. fusca* du Bénin, du Cameroun et du Kenya montre un fort contraste entre l'homogénéité des populations camerounaises (Φ_{ST} = 0,034) et la différenciation hautement significative des populations au sein des deux autres pays (Φ_{ST} = 0,145 pour le Bénin; Φ_{ST} = 0,352 pour le Kenya). Pour ces derniers, il s'agit d'une structuration génétique ancienne qui permet d'exclure l'hypothèse d'une introduction récente. On observe également une structuration génétique au sein des

différents groupes agro-écologiques ou phytogéographiques considérés (Φ_{SC} = 0,128 pour les groupes phytogéographiques béninois et Φ_{SC} = 0,256 pour les groupes agro-écologiques kenyans). Enfin, seuls les groupes agro-écologiques kenyans présentent une structuration génétique entre eux avec Φ_{CT} = 0,129 qui montre que cette différenciation, quoique significative, reste moyenne. Ces résultats indiquent clairement que les massifs forestiers africains ne joueraient qu'un rôle secondaire dans la mise en place de la structure génétique des populations de *B. fusca* : la structuration génétique présentée par les populations de ce ravageur de tiges dépend d'autres facteurs que celui des formations végétales existant en Afrique. La structuration génétique entre les zones agro-écologiques kenyanes pourrait aussi s'expliquer par l'existence d'autres facteurs comme la Rift Valley dont l'effet a été déjà souligné. L'absence de structuration génétique entre les populations camerounaises de *B. fusca,* et au sein et entre les groupes phytogéographiques définis dans ce pays, associée à leur faible diversité génétique sont en faveur d'une expansion géographique récente. Leur diversité haplotypique réduite représente un sous-groupe des haplotypes mitochondriaux du clade *KII* : cela confirme l'origine Est africaine de ces populations. Elles sont vraisemblablement issues de l'expansion vers l'Ouest de la population mitochondriale *KII*, expansion qui s'est accompagnée d'une perte de diversité génétique comme c'est cas de *Prodoxus decipiens* (Lepidoptera, Prodoxidae) en Amérique du Nord (Althoff & Pellmyr 2002). L'analyse moléculaire des échantillons provenant de localités distribuées de l'Est à l'Ouest de la République Démocratique du Congo serait nécessaire pour confirmer cette hypothèse.

L'estimation de la distance génétique entre les trois clades principaux et la prise en compte d'un taux d'évolution moyen pour l'ADN mitochondrial des insectes de l'ordre de 2,2 % par million d'années situe la période de leur isolement autour d'un million d'années, une période bien antérieure à la domestication et à l'extension de la culture du sorgho en Afrique. Cette période est également très antérieure à l'introduction de maïs en Afrique (Chastanet 1998). Elle correspond au Pléistocène. De grands changements climatiques, qui ont profondément modifié les biotopes, se sont produits en Afrique Subsaharienne au cours de cette période Pléistocène. Wagner (2002) souligne que l'instabilité climatique avait commencé autour de 3,3 à 2,45 Ma avec des oscillations entre les périodes humides et chaudes et froides et sèches. deMenocal conclut que ces oscillations entre les différentes périodes ont conduit à des oscillations dans l'expansion des biotopes savanicoles, ce qui a pu avoir des conséquences sur la distribution des consommateurs de graminées. La structure génétique des populations de *B. fusca* rappelle le patron phylogéographique observé pour certains mammifères herbivores (Alpers *et al*. 2004). Celui-ci découle de l'alternance des

phases d'isolement et de différenciation génétique dans une ou plusieurs zones refuges au cours des périodes climatiques chaudes et des phases d'expansion et d'échange génétique entre les populations pendant les périodes climatiques froides et sèches (Arctander *et al*. 1999; Flagstad *et al*. 2001; Van Hooft *et al*. 2002). Plusieurs auteurs (Livingstone 1982; Arctander *et al*. 1999) décrivent l'Afrique de l'Est comme une mosaïque de zones de refuges secondaires pour les mammifères herbivores avec des échanges périodiques temporaires à travers la Rift Valley. L'existence des clades *KI* et *KII* dont les centres d'origine sont situés en Afrique de l'Est confirme la nature mosaïque des refuges de cette région géographique. Bien que les populations *KI* et *KII* soient partiellement sympatriques, nos analyses permettent de conclure que la Rift Valley explique l'essentiel de la variation moléculaire dans cette partie de l'Afrique. Cette dépression représente une barrière naturelle qui assure le maintien et le renforcement de la structure génétique des populations comme cela a été signalé pour de nombreuses espèces de vertébrés africains (Arctander *et al*. 1999, Pitra *et al*. 2002).

Le patron phylogéographique mis en évidence par l'analyse moléculaire de *B. fusca* indique aussi une fracture biogéographique entre l'Afrique de l'Ouest et le reste du continent africain. Une telle fracture biogéographique entre la faune Ouest africaine et celle présente dans l'ensemble Cameroun / Afrique de l'Est est observée chez l'antilope chevaline (Alpers *et al*. 2004). L'Afrique de l'Ouest se différencie géologiquement de l'Afrique de l'Est par l'absence de massif montagneux, ce qui permet plus facilement l'adaptation aux conditions de basse altitude et a conduit à l'apparition d'un clade *W* de *B. fusca* présentant une niche écologique de basse altitude. Des adaptations écologiques s'étaient vraisemblablement produites en favorisant une structuration génétique locale des populations de cette région géographique comme l'indique la valeur de Φ_{ST} observée, par exemple, au Bénin. Ce sont plus les conditions écologiques et l'absence de refuge altitudinal qui auraient entraîné l'adaptation d'une race de basse altitude qui s'était étendue lors des périodes de refroidissement. L'absence de refuge montagneux a donc favorisé le développement d'un écotype de basse altitude. L'analyse moléculaire de la variance a révélé que l'essentiel de la variation moléculaire est expliqué par la structuration entre l'Afrique de l'Ouest et le reste du continent avec un Φ_{CT} = 0,708 ($P < 10^{-3}$). La Ligne Volcanique du Cameroun est vraisemblablement le facteur qui explique cette fracture biogéographique Ouest /Est africain. Elle constitue la barrière géographique qui a contribué par réduction les flux géniques entre Est et Ouest africains et à la différenciation génétique de la population Ouest africaine de *B. fusca*. La partie continentale de cette chaîne volcanique est constituée de trois unités montagneuses anciennes (les Monts Bambouto et Oku et le Plateau Ngaoundere) et d'une

unité montagneuse (l'ensemble formé par le Mont Cameroun et du Manengouba) récemment mise en place (moins de 3 à 1 million d'années) (Marzoli *et al*. 2000). Selon ces auteurs, ces unités montagneuses sont de différents âges : le début de la mise en place de la LVC est de l'Oligocène (figure 5-1). Elles appartiennent à une suite d'appareils volcaniques et subvolcaniques. L'activité volcanique, commencée il y a 42 millions d'années, continue de se produire jusqu'à nos jours comme en témoigne la dernière éruption volcanique (1999) qui s'est manifestée aux abords du Mont Cameroun (Marzoli *et al*. 1999). Les différentes parties constituant chaque unité montagneuse sont apparues progressivement dans le temps. La formation de l'unité montagneuse constituée du Mont Cameroun et de Manengouba serait le facteur principal qui a certainement réduit davantage les flux géniques entre les populations Ouest et Est africaines de *B. fusca*. La LVC, considérée comme une grande zone de biodiversité, joue en même temps le rôle de barrière naturelle pour certaines espèces en favorisant ainsi la séparation des blocs biogéographiques Ouest/Est africains et pour d'autres celle de refuge. Elle est composée d'une proportion importante de plantes avec des espèces végétales endémiques et géographiquement limitées (Küper *et al*. 2005).

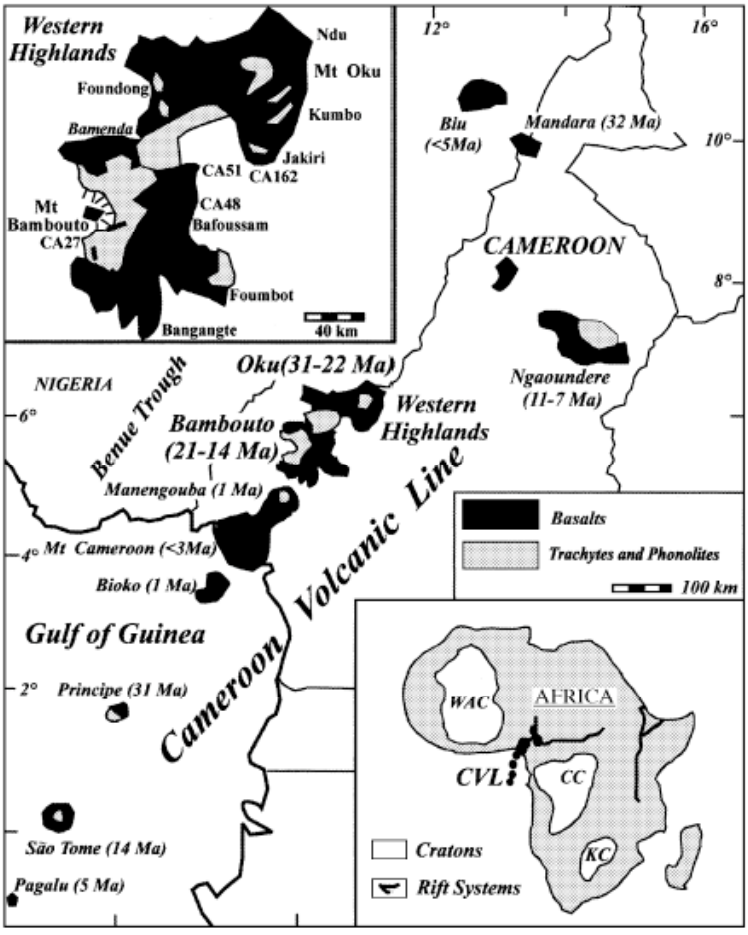


Figure 5-1 : Carte de la Ligne Volcanique Camerounaise (LVC). Les âges marqués pour les différentes unités de la LVC correspondent au volcanisme du stratoïde du secteur continental et au début du volcanisme basaltique des îles océaniques. Les différents cratons sont : Craton Ouest Africain (en anglais, West African Craton : WAC), Craton Congolais (en anglais, Congo Craton : CC) et le Craton de Kalahari (en anglais, Kalahari Craton : KC). (d'après Marzoli *et al*. 2000).

On peut conclure que la Rift Valley, la Ligne Volcanique Camerounaise et les zones désertiques ont modelé la structure génétique des populations de *B. fusca*. A ces facteurs, s'ajoute l'ampleur des cycles climatiques du Pléistocène qui ont vraisemblablement joué un grand rôle dans la différenciation des populations de *B. fusca* et des vertébrés herbivores africains associés au peuplement graminéen. En effet, les dépressions sèches constituent des barrières naturelles lors des périodes chaudes et des lieux d'échange lors des périodes froides. Dans les zones où le refuge montagneux est possible comme Afrique de l'Est et du Centre, il ne se développe pas de race de *B. fusca* de basse altitude car les populations au lieu de s'adapter sont probablement remontées en altitude. Les analyses préliminaires des données microsatellites tendent à confirmer la différenciation Est/Ouest africaine des populations de *B. fusca* (Silvain & Faure comm. pers.)

Les inférences de la NCPA renforcent en partie ces conclusions. Même si la plupart des clades s'interprètent par l'hypothèse d'une restriction de flux de gènes avec isolement par la distance, il en existe néanmoins quelques-uns (clade *W* : 2-1, 3-2 ; clade *KI* : 1-2, 2-1 ; clade *KII* : 1-22, 2-4, 3-2) qui s'interprètent par une hypothèse de restriction de flux de gènes ou, alternativement, par la dispersion avec la colonisation à longue distance. Le clade *W* : 2-4 s'interprète par le phénomène d'expansion continue de l'aire de distribution. Les dispersions à longue distance peuvent être occasionnées par les femelles ou les mâles ou les deux à la fois. En tenant compte de la nature cytoplasmique de nos données moléculaires (comparaison des séquences d'un gène mitochondrial), on peut penser qu'au moins certaines femelles dispersent à longue distance même si des observations écologiques supplémentaires sont nécessaires pour conclure à la généralité du phénomène. Cependant, cela n'exclut pas que les mâles puissent aussi se disperser. Une telle dispersion ne peut être révélée avec un marqueur moléculaire mitochondriale et hérité maternellement. Nos résultats suggèrent cependant que les déplacements des individus de *B. fusca* sont relativement rares. Ils peuvent correspondre aux phases de recolonisation qui font suite au rétablissement de conditions climatiques favorables en zone de basse altitude ou simplement à une colonisation de nouvelles aires de distribution suite à l'extension de la plante hôte. Pour illustrer un de ces phénomènes, nous considérons les deux sous-clades (1-2 et 2-1) de la population *KI* qui s'interprètent par la restriction de flux géniques ou, alternativement, par la dispersion avec la colonisation à longue distance. Dans le sous-clade 1-2, on a deux regroupements géographiques : (Mai Awalde) et (HomaBay, Kisii1, Kisumu1, Kisumu4, Kitale1, Molo, Mombasa2, Narok1, Nyahururu, Rift1, Rift2, Rift5) séparés l'un de l'autre d'une distance d'environ 1.500 km. Bien que plusieurs localités intermédiaires entre ces deux regroupements géographiques aient

été échantillonnées, on n'observe aucun individu de *B. fusca* possédant les haplotypes mitochondriaux qui caractérisent ce sous-clade. En poursuivant la même analyse avec le sous-clade 2-1, on note quatre regroupements géographiques des localités concernées : (Mai Awalde, Adi Ghergish) ; (Abinea, Aliweyo, Kersa, Dereba) ; (Fort Portal, Kabale, Butare) ; (Eldoret1, Kitale1, Kitale2, Kisumu1, Kisumu2, Kisumu4, Kakamega, HomaBay, Kisii1, Kisii2, Narok1, Mombasa2, Rift1, Rift2, Rift5, Nyahururu). Les distances géographiques séparant ces regroupements géographiques sont en moyenne de 400 km. Les deux premiers regroupements géographiques sont séparés par plusieurs localités intermédiaires échantillonnées en *B. fusca* qui ne présentent aucun haplotype mitochondrial propre au sous-clade 2-1. Il se dégage de ces analyses que même si le phénomène de la restriction de flux géniques avec isolement par la distance paraît plus vraisemblable et général, celui de la dispersion avec la colonisation à longue distance ne peut être complètement exclu.

Les résultats préliminaires de l'analyse des microsatellites associés à ceux des données mitochondriale permettent d'affirmer que la diversité génétique de *B. fusca* est plus importante en Afrique de l'Est dans la zone regroupant Kenya, Ouganda, Tanzanie comparativement à d'autres régions africaines (Silvain & Faure comm. pers.). On peut faire l'hypothèse que la population ancestrale de ce ravageur de céréales est d'origine Est africaine. D'autres analyses complémentaires en écologie chimique comme l'étude de la composition chimique des phéromones provenant des différentes populations permettront de vérifier cette hypothèse.

***Démographie de* B. fusca**

Chacun des trois clades majeurs identifiés montre une forte diversité haplotypique associée à une faible diversité nucléotidique. Cette accumulation d'haplotypes confirmée par les valeurs des différents indices démographiques calculés suggère que chaque clade identifié chez *B. fusca* a connu à l'origine un goulot d'étranglement suivi d'une expansion démographique des populations (Grant & Bowen 1998 ; Avise 2000). Une telle expansion démographique peut être liée aux conditions climatiques redevenues favorables. Ultérieurement, les populations de *B. fusca* se sont adaptées aux conditions écologiques actuelles (présence dans les milieux forestiers et savanicoles). Plus récemment, le développement et l'intensification de l'agriculture en Afrique Subsaharienne qui se traduisent par une adaptation remarquable et une préférence du *B. fusca* pour le maïs et le sorgho cultivé. Les populations camerounaises, qui présentent une faible diversité moléculaire, pourraient constituer un argument à l'appui de cette hypothèse. Cette diversité mitochondriale, quoique faible, va de pair avec des

populations de taille importante comme le témoigne l'abondance des individus de *B. fusca* dans les différentes zones écologiques camerounaises (Ndemah *et al*. 2001). On aurait pu penser aussi que cette expansion démographique soit due à l'agriculture, mais, ce n'est pas le cas car les taux de substitution de l'ADN mitochondrial sont trop faibles pour mettre en évidence des phénomènes démographiques aussi récents.

***Plantes hôtes et spécialisation écologique chez* B. fusca**

Le changement de plante hôte est un phénomène courant chez les insectes phytophages. Les processus de l'adaptation des ravageurs aux nouvelles ressources alimentaires restent très mal connus. On peut se demander pourquoi une espèce de ravageur et pas une autre est passée sur les céréales cultivées sur toute l'aire d'expansion de la culture du sorgho et du maïs. On peut également se poser la question de savoir pourquoi ce sont uniquement le sorgho cultivé et le maïs qui sont les plantes hôtes préférentielles de *B. fusca*. La plante hôte originelle de l'espèce reste inconnue étant donné que *B. fusca* est rare dans les habitats sauvages (Kfir *et al*. 2002, Le Rü *et al*. in press). Cependant, la préférence manifestée par ce foreur vis-à-vis du sorgho cultivé a amené certains auteurs (Khan *et al*. 1997; Haile & Hofsvang 2002) à suggérer que *B. fusca* était originellement associé à une espèce particulière de sorgho. Les recherches écologiques réalisées sur plusieurs années en Afrique de l'Est indiquent que *B. fusca* est vraisemblablement oligophage, c'est-à-dire qu'il ne consomme qu'un nombre limité de plantes hôtes appartenant à une même famille botanique (Le Rü *et al*. in press), contrairement aux conclusions initialement avancées par certains écologistes qui l'associaient à un large éventail de plantes hôtes. Ces travaux récents conduisent à considérer que les recensements antérieurs des plantes hôtes sauvages de *B. fusca* étaient erronés du fait de la méconnaissance des chenilles de chaque espèce de foreurs. Ils ont, par ailleurs, montré que la préférence pour les plantes sauvages varie suivant les régions géographiques. Pour *B. fusca*, seules quatre plantes hôtes sauvages appartenant toutes à la famille botanique des Poaceae ont été identifiées : *Pennisetum purpureum*, *Arundo donax*, *Sorghum arundinaceum* et *Setaria megaphylla*. De plus, les auteurs constatent qu'en Erythrée et en Ethiopie, en dehors du sorgho cultivé et du maïs, *B. fusca* retrouvé dans les habitats sauvages est majoritairement et géographiquement associé à *A. donax* alors que dans les milieux sauvages kenyans, il manifeste plutôt sa préférence pour *S. megaphylla* avec une distribution géographique très limitée. Les études d'écologie comportementale conduites au laboratoire sur le choix d'oviposition sur un éventail des plantes hôtes sauvages et cultivées appartenant aux Poaceae, Cyperaceae et Typhaceae ont montré qu'aussi bien les adultes (pour l'oviposition) que les

chenilles (pour leur développement) manifestent une préférence pour une variété particulière de sorgho sauvage (*Sorghum vulgare* pres. var. *sudanense*) (Khan *et al*. 2000; Haile & Hofsvang 2002). Haile & Hofsvang (2002) considéraient que *S. vulgare* était la seule plante hôte sauvage permettant le développement complet de l'insecte dans cette série de plantes étudiées. L'hypothèse la plus vraisemblable et la plus parcimonieuse est que *B. fusca* est passé de sorgho sauvage, et probablement d'autres plantes hôtes sauvages non encore confirmées par les études écologiques, au sorgho cultivé au cours des processus de domestication de cette graminée. L'étude comparée des plantes hôtes sauvages montre que parmi les quatre hôtes sauvages de *B. fusca* identifiées par Le Rü *et al*. (in press), seul *P. purpureum* est pris en compte dans les analyses de Haile & Hofsvang (2002). Des études complémentaires d'écologie comportementale en milieu contrôlé avec les trois autres plantes hôtes sauvages (*A. donax*, *S. arundinaceum*, *S. megaphylla*) sont nécessaires pour savoir le rôle réel de ces dernières.

La spécialisation écologique est une caractéristique des insectes phytophages. Dans le cas de *B. fusca*, les analyses conduites avec le marqueur moléculaire que nous avons utilisé ne mettent pas en évidence de différenciation génétique entre le sorgho cultivé et le maïs. Les haplotypes observés chez les individus vivant sur sorgho cultivé sont identiques à ceux des individus présents sur le maïs. Il est donc très peu vraisemblable que l'on assiste dans le futur à la formation de races d'hôte impliquant ces deux espèces végétales. Si toutefois une ségrégation écologique existe entre les individus exploitant ces deux plantes hôtes, sa signature génétique est à rechercher dans le compartiment nucléaire avec les marqueurs hypervariables étant donnée qu'elle n'existe pas au niveau mitochondrial. Le temps permettant l'apparition d'une ségrégation écologique chez *B. fusca* est certainement supérieur à l'ancienneté de la domestication du sorgho et à l'introduction du maïs en Afrique.

L'effet de la domestication des sorghos sauvages et des évènements historiques

La domestication des sorghos sauvages a commencé au Nord-Est de l'Afrique, mais, les différentes recherches effectuées ont montré qu'il y a eu localement des échanges génétiques entre le stock génétique commun au sorgho cultivé et les sorghos sauvages à travers toutes les régions africaines au cours de l'expansion de l'agriculture. La structuration génétique observée chez *B. fusca* est antérieure à la domestication du sorgho. Elle a eu lieu sur la (les) plante(s) hôte(s) sauvage(s). La principale plante hôte d'origine serait le sorgho sauvage (certainement plusieurs variétés de cette graminée). Les analyses moléculaires (Ollitrault *et al*. 1989) conduites avec les alloenzymes sur les populations de sorgho ont

montré que les sorghos sauvages sont géographiquement structurés en trois groupes principaux : un groupe originaire d'Afrique de l'Ouest, un autre provenant de l'Afrique Australe et un autre originaire du bloc biogéographique Est-Centre d'Afrique. La structure génétique des populations de sorgho cultivé a été influencée par les événements régionaux d'hybridation/introgression avec les espèces de sorghos sauvages. Cela a pu conduire à des groupes génétiques distincts (les cultivars) au sein du sorgho cultivé (Deu & Hamon 1994). Biologiquement, on devrait s'attendre également à observer une même influence régionale pour les insectes associés spécifiquement aux sorghos sauvages et cultivés. *B. fusca* montre un patron géographique partiellement similaire à celui des sorghos sauvages avec un clade *W* spécifique à l'Afrique de l'Ouest alors que le clade *KII* présente une large distribution géographique regroupant l'Afrique de l'Est, du Centre et du Sud. En prenant en compte la préférence différentielle manifestée par *B. fusca* pour les plantes hôtes sauvages suivant les zones écologiques, il n'est pas exclu que d'autres espèces de plantes sauvages botaniquement proches des sorghos sauvages aient servi d'hôtes à ce foreur de tiges en influençant en partie sa structure génétique. Sur la base des analyses phylogéographiques et des préférendums écologiques actuels, l'existence de zones refuges pour *B. fusca* est suggérée. En contraste, De Wet (1978) montre que les préférendums écologiques pour des groupes génétiques de sorghos sauvages sont hautement variables d'une zone écologique à une autre. Compte tenu de l'existence de quelques espèces de plantes hôtes sauvages, on peut penser que les refuges ou les barrières de fragmentation ayant modelé la structure des populations de l'insecte sont vraisemblablement différents, au moins en partie, pour les sorghos sauvages étant donné que *B. fusca* ne serait pas uniquement lié à cette espèce végétale. La distribution géographique actuelle de plusieurs espèces monophages (qui consomment une seule espèce végétale) ou oligophages ne peut pas être expliquée seulement par la distribution de ses plantes hôtes (Le Rü *et al*. in press). L'influence des événements passés (périodes de glaciation, extension des savanes) est vraisemblablement plus marquée pour *B. fusca* que pour le sorgho étant donné que les deux espèces n'ont pas la même biologie. L'isolement et la différenciation (commencée il y a un million d'années et se continuant jusqu'à nos jours) des trois clades mitochondriaux identifiés s'expliquent par des événements anciens correspondant à une époque géologique bien antérieure à la domestication de sorgho à partir de son pool génétique commun du Nord-Est de l'Afrique. Le patron phylogéographique de *B. fusca* est vraisemblablement hérité de l'histoire ancienne de l'insecte associé aux graminées sauvages dans un contexte paléoclimatique qui a favorisé l'isolement géographique de ses populations en des zones refuges correspondant aux préférendums écologiques de ce foreur et *a fortiori*

différents de ceux des sorghos sauvages. Le passage des populations de *B. fusca* des sorghos sauvages aux cultivés s'est vraisemblablement produit à plusieurs reprises et de façon indépendante à partir des populations locales dans les différentes zones biogéographiques. *B. fusca* est aujourd'hui essentiellement inféodé au maïs et au [sorgho] cultivé alors que l'histoire récente de ces plantes n'a pas eu d'influence sur la structure génétique de l'insecte. Ce qui est justement intéressant avec le modèle *B. fusca* est que sa structure génétique n'indique pas une influence des activités humaines alors que l'écologie des insectes actuels est étroitement dépendante de l'existence des agro-systèmes céréaliers.

Lutte biologique et unités évolutives

La présente étude phylogéographique sur *B. fusca* a été motivée par le fait que l'on supposait l'existence de préférendums écologiques différents pour les populations Ouest et Est africaines de ce ravageur de céréales. Même si nos résultats ne nous permettent pas de faire un lien direct entre la diversité génétique neutre et l'adaptation écologique, ils tendent néanmoins à confirmer la différenciation d'une population Ouest africaine très homogène. Ce que renforcent les différentes analyses mitochondriales qui ont montré une séparation nette et ancienne entre les trois populations *W*, *KI*, *KII*. La distribution géographique du clade *W* associée à sa monophylie révélée au cours des analyses successives permet de le considérer comme une unité évolutive significative. Les clades *KI* et *KII* peuvent être considérés comme des unités évolutives significatives malgré le recouvrement géographique partiel existant entre eux. Cependant, des analyses moléculaires avec d'autres marqueurs (les microsatellites) et des approches d'écologie chimique (les phéromones) et comportementale sont nécessaires pour apporter plus d'information sur la biologie et la structure des populations de *B. fusca*. Si la monophylie de chaque clade mitochondrial est confirmée, l'adaptation des stratégies de lutte biologique à l'écologie des insectes faisant partie de ces populations doit être envisagées. L'identification et le recensement des différentes espèces de parasitoïdes associés à *B. fusca* suivi de l'étude écologique des relations coévolutives entre *B. fusca* et ses parasitoïdes et l'analyse génétique des populations de ces antagonistes permettront de déterminer quelle espèce de parasitoïde s'avère la plus virulente vis-à-vis d'une population donnée du foreur. On pourra rechercher des populations particulières de parasitoïdes susceptibles de s'adapter ou ayant développé une adaptation spécifique à une population donnée de *B. fusca* et aux conditions écologiques dans lesquelles se développe cette dernière. Il semble d'ailleurs assez évident que le cortège d'antagonistes est différent suivant les grandes zones biogéographiques africaines. La comparaison des complexes d'ennemis naturels connus en Afrique Orientale et

Occidentale permet de mettre en évidence que le scélionidé parasitoïde des œufs *Telenomus isis* est caractéristiques des populations de *B. fusca* et de *S. calamistis* d'Afrique de l'Ouest alors que le braconidé parasitoïde de larves *Cotesia sesamiae* est commun en Afrique Orientale et Australe (Schulthess *et al.* 1997). Le braconidé parasitoïde des œufs *Chelonus curvimaculatus* et l'ichneumonidé parasitoïde des chrysalides *Procerochasmias nigromaculatus* sont connus sur les populations de *B. fusca* en Afrique Centrale, Orientale et Australe alors qu'ils sont complètement absents en Afrique de l'Ouest (van Achterberg & Walker 1998 ; Zwart 1998). *B. fusca* reste l'hôte principal de *P. nigromaculatus* même s'il est reconnu que d'autres foreurs de tiges comme *S. calamistis*, *C. partellus*, *E. saccharina* servent lui également d'hôtes (Zwart 1998). Sur la base de la différence de faunes d'antagonistes entre Ouest et Est de l'Afrique, certains des parasitoïdes cités ci-dessus sont déjà proposés pour être utilisé après introduction contre *B. fusca* et *S. calamistis* (Schulthess *et al*. 1997).

***Diversité génétique neutre et adaptation écologique des populations de* B. fusca**

Au plan biogéographique, les barrières naturelles empêchent les échanges génétiques entre populations en créant d'importantes différences moléculaires, et parfois morphologiques, entre celles-ci. Les refuges naturels de ces populations constituent sans aucun doute des centres de diversité génétique des espèces. On observe une diversité écologique associée à une diversité mitochondriale des populations de *B. fusca*.

La différenciation mitochondriale mise en évidence chez *B. fusca* n'est pas, par contre, associée une différenciation des traits d'histoire de vie de ces populations comme l'ont montré les travaux de Calatayud *et al*. (in press). Ces auteurs ont analysé quelques échantillons du Kenya et ils sont parvenus à montrer que les périodes d'émergence, de comportement d'appel et de ponte restent identiques chez les individus de *B. fusca* quel que soit leur clade mitochondrial.

Bien que Roelofs (1979) reconnaisse que la structure chimique des phéromones de Lépidoptères est très proche au sein d'un même groupe systématique, des travaux menés sur des Tordeuses (Baltensweiler & Priesner 1988) et sur des Noctuelles (Foster *et al*. 1991) démontrent que le polymorphisme phéromonal au sein d'une même espèce n'est pas rare chez les Lépidoptères et qu'il conduit dans la plupart des cas à un isolement reproducteur entre les races phéromonales. Le système de reconnaissance du partenaire sexuel (SMRS - Specific Mate Recognition System) (Paterson & McEvey 1992) et la spécialisation à une plante-hôte peuvent participer à l'isolement reproducteur entre les populations ou les espèces. Ces deux points sont considérés comme des moteurs de la structuration et de la différenciation des

populations. L'existence du polymorphisme phéromonal au sein des populations d'une même espèce peut être aussi liée à une spécialisation à des plantes-hôtes différentes (Emelianov *et al*. 2001), phénomène qui pose la question de la notion de race d'hôte voire de spéciation sympatrique (Emelianov *et al*. 2003 ; Pelozuelo *et al*. 2004). Les races phéromonales pourraient constituer la première étape d'un processus d'isolement reproducteur qui, associé à des divergences écologiques, conduirait à l'apparition de nouvelles espèces (Dres & Mallet 2002). L'analyse de la composition chimique de la phéromone produite par des femelles de *B. fusca* issues de la nature, prélevées dans les zones géographiques différentes correspondant aux trois types mitochondriaux, montre un polymorphisme phéromonal portant sur des variations de pourcentage ou de ratio des constituants aux niveaux intra- et inter-populationnels (Félix & Frérot comm. pers.). Par contre, bien que les trois grandes populations (*W*, *KI*, *KII*) soient séparées par des distances génétiques non négligeables, il ne semble pas exister de relation entre la différenciation populationnelle observée au niveau mitochondrial et la variabilité des réponses des mâles à la phéromone sexuelle femelle ou aux phéromones synthétiques. Ce résultat n'est pas étonnant puisque l'ADN mitochondrial ne donne qu'une image de l'histoire passée des populations. Du fait de l'absence de recombinaison, l'ADN mitochondrial ne renseigne pas, du moins au niveau populationnel, sur les échanges actuels ou récents entre les individus des trois types mitochondriaux (*W*, *KI* et *KII*). Les résultats de ces analyses phéromonales sont en faveur de l'hypothèse de la non divergence du SMRS ou de son absence de renforcement chez *B. fusca*. L'ADN mitochondrial n'est certainement pas informatif sur ces aspects de la biologie et de l'écologie de ce ravageur de tiges. Les marqueurs moléculaires nucléaires s'avèrent ici nécessaires. En effet, le génome nucléaire est transmis de génération en génération par les deux sexes. Il est soumis aux mêmes lois génétiques (transmission biparentale, recombinaison génique) que la plupart des caractères adaptatifs à la différence du mitochondrial. Dans le cas où les deux sexes ont des aptitudes différentes pour la dispersion (caractère contrôlé par le sexe), la structuration observée à partir d'un marqueur mitochondrial peut différer de celle du génome nucléaire. Tout cela justifie l'utilisation des marqueurs moléculaires nucléaires. Actuellement, l'étude des huit (08) locus microsatellites est cours afin de tirer les informations biologiques complémentaires que pourrait apporter le génome nucléaire. Ils font l'objet d'une étude de génétique des populations sur les mêmes individus ayant servi à faire la présente analyse mitochondriale.

Les insectes ravageurs et la biologie moléculaire

Les capacités évolutives et adaptatives des insectes nuisibles ne doivent pas être négligées. Il s'agit de l'une des raisons qui font que les risques entomologiques en agriculture n'ont significativement pas diminué avec la modernisation des méthodes de lutte. On peut citer la résistance aux insecticides liée à l'adaptation des insectes phytophages aux substances chimiques fréquemment utilisées. Dans le cas des insectes foreurs de tiges, les chenilles vivent à l'intérieur des tiges qui les protègent ainsi contre l'effet nocif des insecticides. Il est donc vraisemblable que les insecticides restent sans effet sur ces foreurs comme cela a été signalé pour la pyrale *E. saccharina* (King *et al*. 2002). La compréhension des mécanismes de l'évolution des insectes est un enjeu majeur. Une lutte biologique efficace contre une espèce ravageuse des plantes cultivées doit passer par la recherche des variétés résistantes de plantes hôtes, la gestion des habitats, le recensement et l'identification des parasitoïdes associés et surtout l'analyse moléculaire des populations afin de les caractériser et de mettre en évidence la structure génétique qu'elles présentent. A l'échelle du Kenya, (entre 3 populations de l'Est de ce pays et 3 autres populations de l'Ouest du même pays) une faible différenciation a été observée chez *B. fusca* pour son niveau de résistance à *C. sesamiae*. Cependant, les possibilités d'adaptation locale à certains génotypes hôtes restent minimes : l'essentiel de la variation dans le succès parasitaire est déterminé par le parasitoïde (Gitau 2006).

En Afrique, on connaît encore moins qu'ailleurs les facteurs de structuration populationnelle, d'isolement et de différenciation qui constituent avec la sélection naturelle les principales forces qui gouvernent l'évolution des espèces. Cela montre toute utilité de l'application des analyses moléculaires à la connaissance des organismes biologiques et à l'élaboration des méthodes de lutte contre les ravageurs de cultures.

Les analyses moléculaires ont apporté un aperçu majeur à la connaissance des mécanismes de l'écologie et de la structure génétique des populations de ravageurs. La plupart des études réalisées se rapportent à des insectes américains ou européens. Les quelques études de génétique des populations portées sur des espèces africaines ravageuses de cultures concernent *Eldana saccharina* (King *et al*. 2002) ; *Ceratitis rosa* et *Ceratitis fasciventris* (Diptera, Tephritidae) (Baliraine *et al*. 2004) ; *Thaumatotibia batrachopa* et *Cryptophlebia peltastica* (Lepidoptera, Tortricidae) (Timm *et al*. 2006). Même si l'intégration des analyses moléculaires dans les études de la génétique des populations et de la phylogéographie des insectes phytophages ne répond pas à toutes les thématiques de la lutte biologique, la résolution de certains problèmes de ravageurs passe en Afrique comme ailleurs par l'utilisation d'outils modernes d'investigation comme ceux apportés par les marqueurs

moléculaires. La mise en évidence des races d'hôtes chez les insectes phytophages est révélée par les analyses moléculaires comme cela a été souligné dans le premier chapitre de cette présente étude. Les centres d'origine, les voies de dispersion et l'histoire évolutive récente comme ancienne des populations sont connus ou approfondis en détail grâce aux analyses moléculaires. La présente étude sur *B. fusca* a eu le mérite de confirmer que les différences de préférendums écologiques entre l'Afrique de l'Ouest et de l'Est qu'on soupçonnait au début de nos analyses sont corroborées par les résultats de l'analyse de la diversité moléculaire au niveau mitochondrial. Cette analyse permet de conclure que *B. fusca* présente de différenciations génétiques en fonction des régions biogéographiques. Il reste à voir et à approfondir ce qui se passe au niveau nucléaire.

L'étude phylogéographique du modèle *B. fusca* permet de prendre en compte dans les études futures d'autres modèles biologiques africains, insectes ravageurs ou non, des facteurs naturels comme la LVC ou la Rift Valley qui paraissent jouer un rôle important dans la structuration des populations. Elle confirme l'ampleur de la diversité génétique dans certaines zones comme l'Afrique de l'Est et le rôle de refuge joué par les reliefs élevés de cette région. Elle souligne aussi la particularité de l'adaptation écologique des populations animales présentes en Afrique de l'Ouest. La présente analyse a apporté un appui fort à la biogéographie africaine en ce sens que les différences de préférendums écologiques entre l'Est et l'Ouest sont congruentes avec la différenciation génétique révélée par l'utilisation d'un marqueur moléculaire.

Perspectives

L'objectif principal actuel de l'équipe IRD - UR072 est de caractériser génétiquement avec les marqueurs microsatellites les individus de *B. fusca* ayant fait l'objet d'analyse moléculaire mitochondriale. Les analyses viseront à déterminer le nombre de populations qui pourront être différenciées en utilisant ces marqueurs et à rattacher chaque individus à ces populations. La localisation géographique de ces populations permettra de compléter notre connaissance des barrières aux flux de gènes qui interviennent dans la structuration des populations de l'espèce. Des corrélations seront recherchées entre l'appartenance populationnelle des individus et les différences de comportements ou de composition des phéromones. Les patrons de structuration révélés par les marqueurs mitochondriaux et microsatellites seront comparés. Dans un second temps, et pour l'Afrique de l'Est au sens large, les patrons de structuration de l'insecte ravageur et de son parasitoïde principal, *C. sesamiae*, seront aussi comparés. Parallèlement, on va aborder la génétique des

populations de *T. isis*, parasitoïde caractéristique des populations de *B. fusca* d'Afrique de l'Ouest.

Ces travaux porteront sur le jeu de données actuellement disponible. Cependant, une meilleure connaissance de la phylogéographie de *B. fusca* et de sa structuration génétique nécessiteront la réalisation de nouvelles missions de collecte. Nous avons vu que des différences existent au plan écologique entre les populations d'Afrique de l'Ouest et de l'Est. Cette dernière partie de l'Afrique est caractérisée par l'abondance de zones d'altitude où *B. fusca* domine. En Afrique de l'Ouest, il est possible de retrouver des zones d'altitude en Guinée, région qui n'a pu être explorée pendant la période couverte par cette thèse. Un échantillonnage dans cette zone permettrait à la fois d'estimer la limite Ouest de l'aire de distribution de *B. fusca* d'une part et d'autre part d'évaluer si cette région a pu dans le passé constituer une zone refuge pour l'espèce. En Afrique Centrale, le rôle de la LVC a été souligné. Un projet de recherche associant l'IRD et le MNHN (Muséum National d'Histoire Naturelle, Paris) devrait permettre d'y échantillonner les foreurs à la fois dans le milieu naturel et dans les zones cultivées. L'objectif visé pour *B. fusca* sera de préciser les limites de l'aire de distribution du clade *KII*. On se demande si celle-ci se situe sur le piedmont Est ou sur le piedmont Ouest de l'autre côté de la LVC.

La compréhension des phénomènes adaptatifs qui se cachent derrière tout ce qui a été montré au niveau de *B. fusca* est importante. Pour aborder le lien entre la génétique et l'adaptation écologique, on peut envisager à l'avenir de s'intéresser au protéome (ensemble des protéines exprimées par une cellule ou un tissu à un instant donné) d'un tel modèle. Bien sûr, la protéomique est une méthode lourde et financièrement coûteuse, mais, son utilisation sur des modèles appliqués comme *B. fusca* peut constituer de bons outils pour des approches plus fondamentales. Les intérêts de l'analyse du protéome par rapport à ceux de l'analyse du génome ou du transcriptome résident dans les caractéristiques suivantes : (1) un même génome peut donner naissance à beaucoup de protéomes différents ; (2) le niveau d'expression des protéines n'est pas prédictible à partir du niveau d'expression des messagers ; (3) les messagers peuvent présenter une importante diversité, en raison de l'utilisation de promoteurs distincts, ou d'épissages alternatifs ; (4) les protéines peuvent être traduites à partir de différents codons d'initiation, et sont fréquemment modifiées après la traduction ; (5) il est très difficile de prédire, à partir de l'analyse des séquences d'ADN ou d'ADNc, dans quelles conditions un gène est transcrit et traduit, à quel stade du développement, dans quel tissu ; (6) le nombre de protéines exprimé dans les tissus est de l'ordre de plusieurs milliers (Bernot 2003). Les travaux de la protéomique permettraient

éventuellement de révéler à ce niveau des différences, probablement adaptatives, entre certaines populations et donc de déterminer quels gènes interviennent dans les phénomènes adaptatifs de ce foreur de tiges. Des études sur la drosophile ont montré des changements dans la charge alaire en fonction de l'altitude : les populations d'altitude présentant une plus faible charge alaire (Morin *et al*. 1999). Les gènes sous exprimés dans les muscles alaires des populations d'altitude et impliqués dans ces adaptations pourraient être identifiés par des méthodes de protéomique comparée.

L'élargissement de l'étude physio-comportementale (les phéromones, les tunnels de vol, les périodes d'appel des femelles) à des échantillons plus importants de *B. fusca* sur toutes ses zones biogéographiques de distribution afin de révéler les petites différences. Les substances chimiques caractéristiques des populations sont à rechercher afin d'approfondir les connaissances sur les traits d'histoire de vie de ce ravageur de plantes cultivées.

Après le recensement et l'identification des parasitoïdes spécifiques à chaque population identifiée ou susceptibles de contrôler efficacement les populations de *B. fusca*, l'analyse moléculaire de la structure génétique de ces populations de parasitoïdes est une étude primordiale pour comprendre le rôle des flux géniques dans les interactions coévolutives entre les parasitoïdes et l'hôte.

L'approfondissement de l'analyse des marqueurs microsatellites en complément à celle mitochondriale présentée ici, associée à la connaissance de la biologie et de l'écologie à la fois des populations de *B. fusca* et de ses parasitoïdes, pourrait conduire à améliorer les chances de réussite de la lutte biologique contre ce ravageur et fournir par extension des enseignements pour améliorer la lutte contre d'autres insectes africains ravageurs de cultures agricoles.

Bibliographie (hors articles)

Ajayi O (1991) Yield reduction in sorghum caused by *Busseola fusca* (Fuller) and *Sesamia calamistis* Hampson. Paper presented at the 9th conference of the African Association of Insect Scientists, Accra, Ghana, 23-27 September, 1991.

Al-Barrak M, Loxdale DH, Brookes PC, Dawah AH, Biron GD, Alsagair O (2004) Molecular evidence using enzyme and RAPD markers for sympatric evolution in Britic-sh species of *Tetramesa* (Hymenoptera: Eurytomidae). *Biological Journal of the Linnean Society*, **83**, 509-525.

Alberti AC, Rodriguero MS, Gomez Cendra P, Saidman BO, Vilardi JC (2002) Evidence indicating that Argentine populations of *Anastrepha fraterculus* (Diptera: Tephritidae) belong to a single biological species. *Annals of the Entomology Society America*, **95**, 505-512.

Alvarez N, Hossaert-McKey M, Rasplus JY, Mckey D, Mercier L, Soldatt L, Aebi A, Shani T, Benrey B (2005) Sibling species of bean bruchids: a morphological and phylogenetic study of *Acanthoscelides obtectus* Say and *Acanthoscelides obvelatus* Bridwell. *Journal of Zoological Systematics and Evolutionary Research*, **43**, 29-37.

Andreev D, Breilid H, Kirkendall L, Brun LO, ffrench-Constant RH (1998) Lack of nucleotide variability in a beetle pest with extreme inbreeding. *Insect Molecular Biology*, **7**, 197-200.

Andrews CM, Callaghan A, Field ML, Williamson SM, Moores DG (2004) Identification of mutation conferring insecticide-insentitive AChE in the cotton-melon aphid, *Aphis gossypii* Glover. *Insect Molecular Biology*, **13**, 555-561.

Appert J (1970) *Insects Harmful to Maize in Africa and Madagascar.* Madagascar Institute of Agronomic Research, 71pp.

Arber W, Linn S (1969) DNA modification and restriction. *Annual Review of Biochemistry*, **38**, 467-500.

Avise JC (1996) Conservation genetics: case histories from nature. Chapman & Hall, New York, NY, USA.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics,* **18**, 489-522.

Baker DA, Loxdale HD, Edwards OR (2003) Genetic variation and founder effects in the parasitoid wasp, *Diaeretiella rapae* (M'intosh) (Hymenoptera: Braconidae: Aphidiidae), affecting its potential as a biological control agent. *Molecular Ecology*, **12**, 3303-3311.

Baliraine FN, Bonizzoni M, Guglielmino CR, Osir SE, Lux A, Mulaa FJ, Gomulski LM, Zheng L, Quilici S, Gasperi G, Malacrida AR (2004) Population genetics of the potentially invasive African fruit fly species, *Ceratitis rosa* and *Ceratitis fasciventris* (Diptera: Tephritidae). *Molecular Ecology*, **13**, 683-685.

Baltensweiler W, Priesner E (1988) A study of pheromone polymorphism in *Zeiraphera diniana* Gn. (Lep, Tortricidae). 3. Specificity of attraction to synthetic pheromone sources by different male response types from two host races. *Journal of Applied Entomology*, **106**, 217-231.

Barrow MR (1987) The effect of first generation maize stalk borer, *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae), on yield of different maize genotypes. *Journal of the Entomological Society of South Africa*, **50**, 291-298.

Batley J, Edwards KJ, Barker JHA, Dawson KJ, Wiltshire CW, Glen DM, Karp A (2004) Population structure of the beelte pests *Phyllodecta vulgatissima* and *P. vitellinae* on UK plantations. *Insect Molecular Biology*, **13**, 413-421.

Baumann R, Schubert R, Heitland W, Auger-Rozenberg MA, Faivre-Rimpant P, Müller-Starck G (2003) Genetic diversity within and among populations of *Diprion pini* (Hymenoptera, Diprionidae) determined by random amplified polymorphic DNA-polymerase chain reaction of haploid males. *Journal of Applied Entomology*, **127**, 258-264.

Becerra JX (1997) Insect on plants: macroevolutionary chemical trends in host use. *Science*, **276**, 253-256.

Beckmann JS, Weber JL (1992) survey of human and rat microsatellites. *Genomics*, **12**, 627-631.

Berkov A (2002) The impact of redefined species limits in *Palame* (Coleoptera: Cerambycidae: Lamiinae: Acanthocinini) on assessments of host, seasonal, and stratum specificity. *Biological Journal of the Linnean Society*, **76**, 195-209.

Berlocher SH (1998) Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence? In Howard, DJ and SH Berlocher, Eds., *Endless Froms: Species and Speciation*, pp. 99-113. New York, NY: Oxford University Press.

Bermingham E, Lamb T, Avise JC (1986) Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. *Journal of Heredity*, **77**, 249-252.

Bernays EA, Chapman RF (1994) Host-Plant selection by phytophagous insects. Chapman & Hall: New York, NY, 312pp.

Bernot A (2003) Analyse de génomes, transcriptomes et de protéomes. 3è édition, Biotech Info, Dunod.

Borror DJ, Triplehorn CA, Johnson NF (1989) An introduction to the study of Insects. 6th edition. New York, Saunders College Publishers.

Bland R, Jaques H (1978) How to know the insects. Third edition. BWm.C. Brown Company Oublishers 409pp.

Bosque-Pérez NA, Schulthess F (1998) Maize: West and Central africa. p. 11-24 *in*: Polaszek A. (ed.), *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. Wallingford, UK: CABI.

Brower AV, Freitas AV, Lee M, Silva-Brandào KL, Whinnett A, Willmott KR (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**, 288-301.

Brower AVZ, DeSalle R (1994) Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Annals of the Entomological Society of America*, **87**, 702-716.

Burban C, Petit JR, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Mastsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology*, **8**, 1593-1602.

Bush GL (1975) Modes of animal speciation. *Annual Review of Ecology and Systematics*, **6**, 339-364.

Calatayud PA, Guénégo H, Le Rü B, Silvain JF, Frérot B Temporal patterns of emergence, calling behaviour and oviposition period of the maize stem borer, *Busseola fusca* (Fuller 1901) (Lepidoptera: Noctuidae). *Annales de la Société Entomologique de France* (in press).

Cardwell KF, Schulthess F, Ndemah R, Ngoko (1997) A systems approach to assess crop health and maize yield losses due to pests and diseases in Cameroon. *Agriculture, Ecosystems and Environment*, **65**, 33-47.

Carisio L, Cervella P, Palestrini C, DelPero M, Rolando A (2004) Biogeographical patterns of the genetic differentiation in dung beetles of the genus *Trypocopris* (Coleoptera, Geotrupidae) inferred from mtDNA and AFLP analyses. *Journal of Biogeography*, **31**, 1149-1162.

Carlson JE, Tulsieram LK, Glaubitz JC, Luk VWK, Kauffeldt C, Rutledge R (1991) Segretation of random amplified DNA makers in F_1 progeny of conifers. *Theoretical and Applied Genetics*, **83**, 194-200.

Carew ME, Pettigrove V, Hoffmann AA (2005) The utility of DNA markers in classical taxonomy: using cytochrome oxidase I markers to differentiate Australian Cladopelma (Diptera: Chironomidae) midges. *Annals of the Entomological Society of America*, **98**, 587-594.

Caskey CT, Pizzuti A, Fu YH, Fenwick RGJ, Nelson DL (1992) Triplet repeat mutations in human disease. *Science*, **256**, 784-788.

Caterino MS, Cho S, Sperling FAH (2000) The current state of insect molecular systematics: a thriving tower of babel. *Annual Review of Entomology*, **45**, 1-54.

Chen Y, Giles KL, Payton ME, Greenstone MH (2000) Identification key cereal aphid predators by molecular gut analysis. *Molecular Ecology*, **9**, 1887-1898.

Choudhary M, Strassman JE, Solis CR, Queller DC (1993) Microsatellite variation in a social insect. *Biochemical Genetics*, **31**, 87-96.

Cognato AI, Sun J, Anducho-Reyes MA, Owen DR (2005) Genetic variation and origin of red turpentine beetle (*Dendroctonus valens* LeConte) introduced to the People's Republic of China. *Agricultural and Forest Entomology*, **7**, 87-94.

Cunningham M, Moritz C (1998) Genetic effects of forest fragmentation on a rainforest restricted lizard (Scincidae: Gnypetscincus queenslandiae). *Biological Conservation*, **83**, 19-30.

De Barro PJ (2005) Genetic structure of the whitefly *Bemisia tabaci* in the Asia-Pacific region revealed using microsatellite markers. *Molecular Ecology*, **14**, 3695-3718.

DeSalle R, Hunt JA (1987) Molecular evolution in Hawaiian drosophilids. *Trends in Ecology and Evolution*, **2**, 212-216.

Dres M, Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 357, 471-492.

Emelianov I, Simpson F, Narang P, Mallet J (2003) Host choice promotes reproductive isolation between host races of the larch budmoth *Zeiraphera diniana*. *Journal of Evolutionary Biology*, **16**, 208-218

Emelianov I, Dres M, Baltensweiler W, Mallet J (2001) Host-induced assortative mating in host races of the larch budmoth. *Evolution*, **55**, 2002-2010.

Ebenebe AA, Van den Berg J, Van den Linde TC (1999) Effect of planting date of maize on damage and yield loss caused by the stalk borer, *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae) in Lesotho. *South African Journal of Plant and Soil*, **16**, 180-185.

Enayati AA, Ranson H, Hemingway J (2005) Insect glutathione transferases and insecticide resistance. *Insect Molecular Biology*, **14**, 3-8.

Endersby NM, Mckechnie SW, Ridland PM, Weeks AR (2006) Microsatellites reveal a lack of structure in Australian populations of the diamondback moth, *Plutela xylostella* (L.). *Molecular Ecology*, **15**, 107-118.

Estoup A, Solignac M, Cornuet JM (1994) Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies. *Proceedings of the Royal Society, London, Serie* B 258, 1-7.

Evans JD (1993) Parentage analysis in ant colonies using single sequence loci. *Molecular Ecology*, **2**, 89-93.

Farrell BD (1998) "Inordinate fondness" explained: why are there so many beetles? *Science*, **281**, 555-559.

Faure N, Silvain JF (2005) Characterization of eight microsatellite loci in the maize stalk borer *Busseola fusca* Fuller, 1901 (Lepidoptera: Noctuidae). *Molecular Ecology Notes*, **5**, 846-848.

Feder JL, Chilcote CA, Bush GL (1989) Are the apple maggot, *Rhagoletis pomonella*, and blueberry maggot, *R. mendax*, distinct species? Implications for sympatric speciation. *Entomologia Experimentalis et Applicata*, **51**, 113-123.

Fisk JH, Daly JC (1989) Electrophoresis of *Helicoverpa armigera* (Hübner) and *H. punctigera* (Wallengren) (Lepidoptera: Noctuidae): Genotype eexpression in eggs and allozyme variations between life stages. *Journal of the Australian Entomoological Society*, **28**, 191-192.

Foster SP, Dugdale JS, White CS (1991) Sex pheromones and the status of greenheaded and brownheaded leafroller moths in New Zealand. *New Zealand Journal of Zoology*, **18**, 63-74.

Franck P, Garnery L, Loiseau A, Oldroyd BP, Hepburn HR, Solignac M, Cornuet JM (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. *Heredity*, **86**, 420-430.

Friedlander TP, Regier JC, Mitter C, Wagner DL (1996) A nuclear gene for higher level phylogenetics: phosphoenolpyruvate carboxykinase tracks Mesozoic-Age divergences within Lepidoptera (Insecta). *Molecular Biology and Evolution*, **13**, 594-604.

Friedlander TP, Regier JC, Mitter C (1994) Phylogenetic information content of five nuclear genes sequences in animals: initial assessment of character sets from concordance and divergence studies. *Systematic biology*, **43**, 511-525.

Friedlander TP, Regier JC, Mitter C (1992) Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. *Systematic Biology*, **41**, 483-490.

Funk DJ, Futuyma DJ, Orti G, Meyer A (1995) A history of host association and evolutionary diversification for *Opharella* (Coleoptera: Chrysomelidae): new evidence from mitochondrial DNA. *Evolution*, **49**, 1008-1017.

Gallego D, Galian J (2001) The internal transcribed spacers (ITS1 and ITS2) of the rDNA differentiates the bark beetle forest pests *Tomicus destruens* and *T. piniperda*. *Insect Molecular Biology*, **10**, 415-420.

Gitau ACW (2006) *Geographic variation in the development of* Cotesia sesamiae *(Hymenoptera: Braconidae) on* Busseola fusca *(Lepidoptera: Noctuidae) in Kenya: Polydnaviruses and Co-evolutionary dynamics. Ph.D., Thesis. Kenyatta University. Nairobi. Kenya.*

Gómez-Zurita J, Vogler AP, Funk DJ (2004) Diagnosing an overlooked North American taxon: biological observations and mitochondrial insights on *Calligrapha suturella* (Coleoptera, Chrysomelidae). *Annals of the Entomological Society of America*, **97**, 28-36.

Goodisman MAD, Matthews RW, Crozier RH (2001) Hierarchical genetic structure of the introduced wasp *Vespula germanica* in Australia. *Molecular Ecology*, **10**, 1423-1432.

Gounou S, Schulthess F, Shanower T, Hammond WNO, Braima H, Olaleye I, Cudjoe AR, Adjakloe R, Antwi KK (1994) Stem and ear borers of maize in Ghana. Plant Health Management Research Monograph No. 4. International Institut of Tropical Agriculture (IITA): Ibadan, Nigeria.

Grapputo A, Boman S, Lindström L, Lyytinen A, Mappes J (2005) The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. *Molecular Ecology*, **14**, 4207-4219.

Hall HG (1990) Parental analysis of introgressive hybridization between African and European honeybees using nuclear DNA RFLPs. *Genetics*, **125**, 611-621.

Hamada H, Petrino MG, Kakunaga T (1982) Molecular structure and evolutionary origin of human cardiac muscle actin gene. *Proceedings of the National Academy of Science*, **79**, 5901-5905.

Haring E, Aspöck U (2004) Phylogeny of the Neuropterida: a first molecualr approach. *Systematic Entomology*, **29**, 415-430.

Harper GL, Maclean N, Goulson D (2003) Microsatellite markers to assess the influence of population size, isolation and demography change on the genetic structure of the UK butterfly *Polyommatus bellargus*. *Molecular ecology*, **12**, 3349-3357.

Harris H (1966) Enzyme polymorphisms in man. *Proceedings of the Royal Society of London, B*. **164**, 298-310.

Harris KM (1962) Lepidopterous stemborers of cereals in Nigeria. *Bulletin of Entomological Research*, **53**, 139-171.

Harris KM, Nwanze KF (1992) *Busseola fusca* (Fuller), the African maize stalk borer: a handbook of information. *Information Bulletin* 33. ICRISAT. Patancheru, India and CABI Oxon U.K. 84 pp

Harrison RG, Rand DM, Wheeler WC (1987) Mitochondrial DNA size variation in field crickets across a narrow hybrid zone. *Molecular Biology and Evolution*, **4**, 144-158.

Harry M (2001) Génétique moléculaire et évolutive. Maloine, pp: 326.

Hartl DL (1994) Génétique des populations. Médecine-Sciences, Flammarion, pp: 305.

Haymer DS, McInnis DO, Arcangeli L (1992) Genetic variation between strains of the Mediterranean fruit fly, *Ceratitis capitata*, detected by DNA fingerprinting. *Genome*, **35**, 528-533.

Heckel DG, Gahan LJ, Daly JC, Trowell S (1998) A genomic approach to understanding *Heliothis* and *Helicoverpa* resistance to chemical and biological insecticides. *Phylosophical Transactions of the Royal Society: Biological Sciences* B 353: 1751-1756.

Hillis DM, Mable BK, Larson A, Davis SK, Zimmer EA (1996) Nucleic acids IV: Sequencing and cloning. P: 321-381 in *Molecular Systematics* (Hillis DM, Moritz C, Mable BK). 2nd Edition, Sinauer Associates Inc. Sunderland, MA, USA.

Hughes CR, Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, **2**, 131-137.

Hunter RL, Markert CL (1957) Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science*, **125**, 1294-1295.

Imbert G, Kretz C, Johnson K, Mandel JL (1993) Origin of the expansion mutation in myotonic dystrophy. *Nature Genetics*, **4**, 72-76.

Irwin DM, Kocher DK, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of molecular Evolution*, **32**, 128-144.

Javed N, Viner R, Williamson SM, Field ML, Devonshire LA, Moores DG (2003) Characteriszation of acetylcholinesterases, and their genes, from the hemipteran species *Myzus persicae* (Sulzer), *Aphis gossypii* (Glover), *Bemisia tabaci* (Gennadius) and *Trialeurodes vaporariorum* (Westwood). *Insect Molecular Biology*, **12**, 613-620.

Jeffrey LF, Williams SM, Berlocher SH, McPheron BA, Bush GL (1999) The population genetics of the apple maggot fly, *Rhagoletis pomonella* and the snowberry maggot, *R. zephyria*: implications for models of sympatric speciation. *Entomologia Experimentalis et Applicata*, **90**, 9-24.

Jepson WF (1954) A critical review of the world literature on the lepidopterous stalk borers of tropical graminaceous crops. London, Commonwealth Institut of Entomology,127pp.

Jost MC, Shaw KL (2006) Phylogeny of Ensifera (Hexapoda: Orthoptera) using three ribosomal loci, with implications for the evolution of acoustic communication. *Molecular Phylogenetics and Evolution*, **38**, 510-530.

Kelley SC, Farrell BD, Mitton JB (2000) Effects of specialization on genetic differentiation in sister species of bark beetles. *Heredity*, **84**, 218-227.

Kergoat GJ, Delobel A, Silvain JF (2004) Phylogeny and host-specificity of European seed beetles (Coleoptera, Bruchidae), new insights from molecular and ecological data. *Molecular Phylogenetics and Evolution*, **32**, 855-865.

Kfir R (1991) Non-chemical control of grain stalk borers. *Plant Protection News*, **24**, 4

Kfir R (1988) Hibernation by the lepidopteran stalk borers, *Busseola fusca* and *Chilo partellus* on grain sorghum. *Entomologia Experimentalis et Applicata*, **48**, 31-36.

Kfir R (1997) Competitive displacement of *Busseola fusca* (Lepidoptera: Noctuidae) by *Chilo partellus* (Lepidoptera: Pyralidae). *Annals of the Entomological Society of America*, **90**, 620-624.

Kfir R (1997) Natural control of the cereal stem borers *Busseola fusca* and *Chilo partellus* in South Africa. *Insect Science and its Application*, **17**, 61-68.

Kim SK, Sappington WT (2004) Genetic structuring of boll weevil populations in the US based on RAPD markers. *Insect Molecular Biology*, **13**, 293-303.

King H, Conlong DE, Mitchell A (2002) Genetic differentiation in *Eldana saccharina* (Lepidoptera: Pyralidae): evidence from the mitochondrial cytochrome oxidase I and II genes. *Proceedings of the South African Sugar Technology Association*, **76**, 321-328.

Knoll S, Rowell-Rahier M, Mardulyn P, Pasteel JM (1996) Spatial genetic structure of leaf beetle species with special emphasis on alpine populations. In *Chrysomelidae Biology*, Vol. 1: *The Classification, Phylogeny, and Genetics* (eds Jolivet PHA, Cox ML), pp. 379-388. SPB Academic Publishing, Amsterdam.

Kox FEL, van den Beld EH, Lindhout IB, de Goffau WJL (2005) Identification of economically importabt *Liriomyza* species by PCR-RFLP analysis. *Bulletin*, **35**, 79-85.

Kronauer DJC, Schöning C, Pedersen JS, Boomsma JJ, Gadau J (2004) extreme queen-mating frequency and colony fission in African army ants. *Molecular Ecology*, **13**, 2381-2388.

Kumar R (1991) La lutte contre les insectes ravageurs. CTA-Karthala 311p

Küper W, Sommer JH, Lovett JC, Mutke J, Linder HP, Beentje HJ, Sarvan R, Chatelain C, Sosef M, Barthlott W (2005) Africa's hotspots of biodiversity redefined. *Annals of the Missouri Botanical Garden*, **91**, 525-535.

Le Rü BP, Ong'amo GO, Moyal P, Ngala L, Musyoka B, Abdullah Z, Cugala D, Defabachew B, Haile TA, Matama TK, Lada VY, Negassi B, Pallangyo B, Ravolonandrianina J, Sidumo A, Omwega CO, Schulthess F, Calatayud PA, Silvain JF (2006) Diversity of lepidopteran stem borers on monocotyledonous plants in eastern Africa and the islands of Madagascar and Zanzibar revisited. *Bulletin of Entomological Research* (in press).

Levy HC, Garcia-Maruniak A, Maruniak JE (2002) Strain identification of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) insects and cell line: PCR-RFLP of cytochrome oxydase *c* subunit I gene. *Florida Entomolgist*, **85**, 186-190.

Lin CP Danforth BN (2004) How do insect nuclear and mitochondrial genes substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Molecular Phylogenetics and Evolution*, **30**, 686-702.

Lewontin RC, Hubby JL (1966) A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. *Genetics*, **54**, 595-609.

Llewellyn KS, Loxdale HD, Harrington R, Brookes CP, Clark SJ, Sunnucks P (2003) Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. *Molecular Ecology*, **12**, 21-34.

Loukas M, Krimbas CB, Mavragani-Tsipidou P, Kastritsis CD (1979) Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. *Journal of Heredity*, **70**, 17-26.

Loxdale HD, den Hollander J (1989) *Electrophoretic studies on agricultural pests*. Systematics Association Special volume No. 39, Oxford, Clarendon Press.

Loxdale HD, Lushai G (1998) Molecular markers in entomology. *Bulletin of Entomological Research*, **88**, 577-600.

Loxdale HD, Tarr IJ, Weber CP, Brookes CP, Digby PGN, Castanera P (1985) Electrophoretic study of enzymes from cereal aphid population. III. Spatial and temporal genetic variation of populations of *Sitobion avenae* (F.) (Hemiptera: Aphididae). *Bulletin of Entomological Research*, **75**, 121-141.

Lu YJ, Adang MJ, Isenhour DJ, Kochert GD (1992) RFLP of genetic variation in North American populations of the fall armyworm moth *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Molecular Ecology*, **1**, 199-207.

Lunt HD, Hyman CB (1997) Animal mitochondrial DNA recombination. *Nature*, **387**, 247-.

Maa WCJ, Terriere LC (1983) age-dependent variation in enzymatic and electrophoretic properties of house fly (*M. domestica*) carboxylesterases. *Compparative Biochemistry and Physiology*, **74C**, 461-467.

Mallatt JM, Garey JR, Shultz JW (2004) Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and kin. *Molecular Phylogenetics and Evolution*, **31**, 178-191.

Mallarino R, Bermingham E, Willmott KR, Whinnett A, Jiggins CD (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Molecular Phylogenetics and Evolution*, **34**, 625-644.

Mally CW (1920) The maize stalk borer *Busseola fusca* (Fuller). *Bulletin the Department Agriculture of South Africa* 3, 111 pp.

Mardulyn P, Milinkovitch MC, Pasteels JM (1997) Phylogenetic analyses of DNA and allozymes data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. *Systematic Biology*, **46**, 722-747.

Marzoli A, Renne PR, Piccirillo EM, Francesca C, Bellieni G, Melfi AJ, Nyobe JB, N'ni J (1999) Silicic magmas from the continental Cameroon Volcanic Line (Oku, Bambouto and Ngaoundere): ^{40}Ar-^{39}Ar dates, petrology, Sr-Nd-O isotopes and their petrogenetic significance. *Contributions to Mineralogy and Petrology*, **135**, 133-150.

Marzoli A, Piccirillo EM, Renne PR, Bellieni G, Lacumin M, Nyobe JB, Tongwa AT (2000) The Cameroon Volcanic Line Revisited petrogenesis of continental basaltic magmas from lithospheric and asthenospheric mantle sources. *Journal of Petrology*, **41**, 87-109.

Mason ML, Lawson FA (1982) Biology of the American aspen beetle [Coleoptera: Chrysomelidae: *Gonioctena americana* (Schaeffer)] in the Medicine Bow National Forest, Wyoming. *Journal of the Kansas Entomological Society*, **55**, 779-788.

Mathee JJ (1974) Pests of graminaceous crops in South Africa. *Entomology Memoir 40, Departement of Agricultural Technical Services, Republique of South Africa*, 24 pp.

Mayaux P, Richards T, Janodet E (1999) A vegetation map of Central Africa derived from satellite imagery. *Journal of Biogeography*, **25**, 353-366.

McCauley DE, Wade MJ, Breden FJ, Wohltman M (1998) Spatial and temporal variation in group relatedness: evidence from the imported willow leaf beetle. *Evolution*, **42**, 184-192.

McFarane J (1990) Damage assessment and yield losses in sorghum due to the stem borer *Busseola fusca* (Fuller) (Lepidoptera: Noctuidae) in northern Nigeria. *Tropical Pest Management* **36**, 131-137.

Meglecz E, Petenian F, Danchin E, Cœur D'Acier A, Rasplus J-Y, Faure E (2004) High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. *Molecular Ecology*, **13**, 1693-1700.

Meglécz E, Solignac M (1998) Microsatellite loci for *Parnassius mnemosyne* (Lepidoptera). *Hereditas*, **128**, 179-180.

Mercer-Quarshie H (1970) Sorghum and millet improvement in Ghana, *African soils/Sols Africans*, **15**, 175-184.

Meselson M, Yuan R (1968) DNA restriction enzyme from *E. Coli*. *Nature*, **217**, 1110-1114.

Meusel MS, Moritz RFA (1993) Transfer of paternal mitochondrial DNA in fertilization of honeybees (*Apis mellifera* L.) eggs. *Current Genetics*, **24**, 539-543.

Mitchell A, Mitter C, Regier JC (2006) Systematics and evolution of the cutworm moths (Lepidoptera: Noctuidae): evidence from two protein-coding nuclear genes. *Systematic Entomology*, **31**, 21-46.

Mohyuddin AI, Greathead DJ (1970) An annotated list of the parasites of graminaceous stem borers in East Africa, with discussion of their potential in biological control. *Entomophaga*, **15**, 241-274.

Morin JP, Moreteau B, Petavy G, David JR (1999) Divergence of reaction norms of size characters between tropical and temperate populations of *Drosophila melanogaster* and *D. simulans*. *Journal of Evolutionary Biology,* **12**, 329-339.

Moriyama EN, Powell JR (1997) Codon usage bias and tRNA abundance in *Drosophila*. *Journal of Molecular Evolution*, **45**, 514-523.

Moya A, Guirao P, Cifuentes D, Beitia F, Cenis JL (2001) Genetic diversity of Iberian populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on random amplified polymorphic DNA-polymerase chain reaction. *Molecular Ecology*, **10**, 891-897.

Mulaa MA (1995) Evaluation of factors leading to rational pesticide use for the control of the maize stalk borer *Busseola fusca* in Trans-Nzoia district, Kenya. PhD thesis, University of Wales, Cardiff.

Müller C, Barker A, Boevé JL, De Jong PW, De Vos H, Brakefield PM (2004) Phylogeography of two parthenogenetic sawfly species (Hymenoptera: Tenthredinidae): relationship of population genetic differentiation to host plant distribution. *Biological Journal of the Linnean Society*, **83**, 219-227.

Muyango S (1987) Les chenilles foreuses des tiges du maïs et du sorgho : écologie et lutte. In *Séminaire sur les maladies et les ravageurs des principales cultures vivrières d'Afrique Centrale*, pp 117-126. Ede/Wageningen : Centre Technique de Coopération Agricole et Rurale (CTA).

Nagoshi RN, Meagher RL (2003) *FR* tandem-repeat sequence in fall armyworm (Lepidoptera: Noctuidae) host strains. *Annals of the Entomology Society America*, **96**, 329-335.

Nardi F, Carapelli A, Dallai R, Roderick GK, Frati F (2005) Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). *Molecular Ecology*, **14**, 2729-2738.

Nève G, Meglécz E (2000) Microsatellite frequencies in different taxa. *Tree*, **15**, 376-377.

Nice CC, Anthony N, Gelembiuk G, Raterman D, Ffrench-Constant R (2005) The history and geography of diversication within the butterfly genus *Lycaeides* in North America. *Molecular Ecology*, **14**, 1741-1754.

Ohshima I, Yoshizawa K (2006) Multiple host shift between distantly related plants, Juglandaceae and Ericaceae, in the leaf-minig moth *Acrocercops leucophaea* complex (Lepidoptera: Gracillariidae). *Molecular Phylogenetics and Evolution*, **38**, 231-240.

Olivo PD, Van De Walle MJ, Laipis JL, Hauswirth WW (1983) Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. *Nature*, **306**, 400-402.

Palo J, Varvio SL, Hanski I, Väinölä R (1995) Developing microsatellite markers for insect population structure: complex variation in a checkerspot butterfly. *Hereditas*, **123**, 295-300.

Parsons MY, Shaw LK (2001) Species boundaries and genetic diversity among Hawaiian crickets of the genus *Laupala* identified using amplified fragment length polymorphism. Molecular Ecology, **10**, 1765-1772.

Pasteur N, Stordeur E (1976) L'α-glycérophosphate deshydrogenase du moustique *Culex pipens* : génétique formelle, linkage et étude des populations, *Genetica*, **46**, 319-326.

Paterson HEH McEvey SF (1992) Evolution and the Recognition Concept of Species: Collected Writings (Hardcover): *The Johns Hopkins University Press*

Pelozuelo L, Malosse C, Genestier G, Guenego H, Frérot B (2004) Host-plant specialization in pheromone strains of the European corn borer *Ostrinia nubilalis* in France. *Journal of Chemical Ecology*, **30**, 335-352.

Pinto JD, Kazmer DJ, Platner GR, Sassaman CA (1992) Taxonomy of the *Trichogramma minutum* complex (Hymenoptera: Trichogrammatidae): allozymic variation and its relationship to reproductive and geographic data. *Annals of the Entomological Society of America*, **85**, 413-422.

Poinar GO, Polaszek A (1998) Nematoda, Fungi, Protozoa, Bacteria and viruses, p. 283-293 *in*: Polaszek A. (ed.), *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. Wallingford, UK: CABI.

Pramual P, Kuvangkadilok C, Baimai V, Walton C (2005) Phylogeography of the black fly *Simulium tani* (Diptera: Simuliidae) from thailand as inferred from mtDNA sequences. *Molecular Ecology*, **14**, 3989-4001.

Raboudi F, Mezghani H, Makni H, Marrakchi M, Rouault DJ, Makni M (2004) Aphid species identification using cuticular hydrocarbons and cytochrome *b* gene sequences. *Journal of Applied Entomology*, **129**, 75-80.

Rank NE (1992) A hierarchical analysis of genetic differentiation in a montane leaf beetle *Chrysomela aeneicollis* (Colleoptera: Chrysomelidae). *Evolution*, **46**, 1097-1111.

Richardson BJ, Baverstock PR, Adams M (1986) *Allozyme electrophoresis. A handbook for animal systematics and population studies*. London, Academic Press.

Roberts JR (1984) Restriction and modification enzymes and their recognition sequences. *Nucleic Acids Research*, **12**, r167-r204.

Rokas A, Nylander JAA, Ronquist F, Stone GN (2002) A maximun likelihood analysis of eight phylogenetic markers in gallwasp (Hymenoptera: Cynipidae); implications for insect phylogenetic studies. *Molecular Phylogenetics and Evolution*, **22**, 206-219.

Ryckewaert P, Alauzet C (2001) Characterization of *Bemisia* (Hom., Aleyrodidae) from the lesser Antilles by electrophoresis. *Journal of Applied Entomology*, **125**, 263-266.

Salvato P, Battisti A, Concato S, Masutti L, Patarnello T, Zane L (2002) Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa – wilkinsoni* complex), inferred by AFLP and mitochondrial markers. *Molecular Ecology*, **11**, 2435-2444.

Saux C, Fisher BL, Spicer GS (2004) Dracula ant phylogeny as inferred by nuclear 28S rDNA sequences and implications for ant systematics (Hymenoptera: Formicidae: Amblyoponinae). *Molecular Phylogenetics and Evolution*, **33**, 457-468.

Schiffer M, Carew ME, Hoffmann AA (2004) Molecular, morphological and behavioural data reveal the presence of a cryptic species in the widely studied *Drosophila serrata* species complex. *Journal of Evolutionary Biology*, **17**, 430-442.

Schulthess F, Bosque-Pérez NA, Chabi-Olaye A, Gounou S, Ndemah R, Goergen G (1997) Exchange of natural enemies of lepidopteran cereal stemborers between African regions. *Insect Science and its Application*, **17**, 97-108.

Sebastiani F, Meisweinkel R, Gomulski ML, Guglielmino RC, Mellor SP, Malacrida RA, Gasperi G (2001) Molecular differentiation of the Old World *Culicoides imicola* species complex (Diptera, Ceratopogonidae), inferred using random amplified polymorphic DNA markers. *Molecular Ecology*, **10**, 1773-1786.

Seshu Reddy KV (1998) Maize and Sorghum: East Africa. p. 25-27 *in*: Polaszek A. (ed.), *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. Wallingford, UK: CABI.

Seshu Reddy KV Walker PT (1990) A review of the yield losses in graminaceous crops caused by *Chilo* spp. *Insect Science and its Application*, **11**, 563-569.

Seshu Reddy KV, Sum KOS, Lubega MC (1989) Empirical models for predicting yield loss in sorghum caused by *Chilo partellus* (Swinhoe). *Discovery and Innovation*, **1**, 90-94.

Shitara H, Hayashi JI, Takahama S, Kaneda H, Yonekawa H (1998) Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics*, **148**, 851-857.

Sithole SZ (1987) Maize insect pests in Zimbabwe. Toward insect resistant maize for the Third World. *Proceedings of an Internationnal Symposium on Methodologies for Developing Host Plant Resistance Maize Insects, Mexico*, pp. 286-288. Mexico City: Centro Int. Majoramiento Maiz y Trigo.

Smith TB, Wayne RK (1996) *Molecular genetic approaches in conservation*. Oxford, Oxford University Press.

Smithies O (1956) Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *Biochemical Journal*, **61**, 629-641.

Silvain JF, Delobel A (1998) Phylogeny of West African *Caryedon* (Coleoptera: bruchidae): congruence between molecular and morphological data. *Molecular Phylogenetics and Evolution*, **9**, 533-541.

Solignac M, Genermont J, Monnerot M, Mounolou JC (1984) Genetics of mitochondria in *Drosophila*: mtDNA inheritance in heteroplasmic strains of *D. mauriatiana*. *Molecular Genetics and Genomics*, 197, 183-188.

Songa JM (1999) *Distribution, importance and management of stemborers (Lepidoptera) in maize production systems of semi-arid eastern Kenya with emphasis on biological control*. PhD thesis. Kenyatta University, Nairobi. 251 pp.

Songa JM, Overholt WA, Okello RO, Mueke JM (2002) Control of lepidopteran stem borers in maize by indigenous parasitoids in semi-arid areas of eastern Kenya. *Biological Agriculture and Horticulture*, **20**, 77-90.

Sota T, Ishikawa R (2004) Phylogeny and life-history evolution in Carabus (subtribe Carabina: Coleoptera, Carabidae) based on sequences of two nuclear genes. *Biological Journal of the Linnean Society*, **81**, 135-149.

Starks KJ (1969) Some cereal crop insects in east Africa. *East African Agriculture and Forestry Research Organization*, Serere Research Station, Uganda.

Strong KL, Mahon RJ (1991) Genetic variation in the Old World screw-worm fly, *Chrysomya bezziana* (Diptera: Calliphoridae). *Bullletin of Entomological Research*, **81**, 491-496.

Sword AG, Joern A, Senior LB (2005) Host plant-associated genetic differentiation in the snakeweed grasshopper, *Hesperotettis viridis* (Orthoptera: Acrididae). *Molecular Ecology*, 14, 2197-2205.

Symondson WOC, Liddell JE (1996) *The ecology of agricultural pests: biochemical approaches*. London, Chapman & Hall.

Szalanski AL, Roehrdanz RL, Taylor DB, Chandler L (1999) Genetic variation in geographical populations of western and Mexican corn rootworm. *Insect Molecular Biology*, **8**, 519-525.

Taneja SL, Nwanze KF (1989) Assessment of yield loss of sorghum and pearl millet due to stem borer damage. In: ICRISAT International Workshop on Sorghum Stem Borers, 17-20 November 1987. ICRISAT, Patancheru, India, pp. 95-104.

Thibodeau SN, Bren G, Shaid D (1993) Microsatellite instability in cancer of the proximal colon. *Science*, **260**, 816-819.

Timm AE, Geertsema H, Wamich L (2006) Analysis of population genetic structure of two closely related tortricid species of economic importance on macadamias and litchis in South Africa. *Agricultural and Forest Entomology*, **8**, 113-119.

Timmermans NTJM, Ellers J, Mariën J, Verhoef CS, Ferwerda BE, Van Straalen MN (2005) Genetic structure in *Orchesella cincta* (Collembola): strong subdivision of European populations inferred from mtDNA and AFLP markers. *Molecular Ecology*, **14**, 2017-2024.

Toda S, Komazaki S, Tomita T, Kono Y (2004) Two amino acid substitutions in acetylcholinesterase associated with pirimicarb and organophosphorous insecticide resistance in the cotton aphid, *Aphis gossypii* Glover (Homoptera: Aphididae). *Insect Molecular Biology*, **13**, 549-553.

Usua EJ (1968) The biology and ecology of *Busseola fusca* and *Sesamia* species in southwestern Nigeria: distribution and population studies. *Journal of Economic Entomology*, **61**, 830-833.

van Achterberg C, Walker AK (1998) Braconidae, p. 137-185 *in*: Polaszek A. (ed.), *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. Wallingford, UK: CABI.

Van Rensburg GDJ, Bate R (1987) Preliminary studies on the relative abundance and distribution of the stalk borer borers *Busseola fusca* and *Chilo partellus Technical*

Communication of Department of Agriculture and Water Supply, Republic of South Africa, **212**, 49-52.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fringerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

Wainhouse D, Jukes MR (1997) Geographic variation within populations of *Panolis flammea* (Lepidoptera : Noctuidae) in Britain. *Bulletin of Entomological Research*, **87**, 95-99.

Wahlberg N, Brower AVZ, Nylin S (2005) Phylogenetic relationships and historical biogeography of tribes and genera in the subfamily Nymphalinae (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **86**, 227-251.

Walker PT (1960) The relation between infestation by the stalk borer *Busseola fusca*, and yield of maize in East Africa. *Annals of Applied Biology*, **48**, 780-786.

Walker PT, Hodson MJ (1976) Developments in maize stem-borer control in East Africa, including the use of insecticide granules. *Annals of Applied Biology*, **84**, 111-114.

Warui CM, Kuria JN (1983) Population incidence and the control of maize stalk borers *Chilo partellus* (Swinh.) and *Chilo orichalcociliellus* Strand and *Sesamia calamistis* Hmps. in Coast Province, Kenya. *Insect Sciences and its Application*, **4**, 11-18.

Wiegmann BM, Mitter C, Regier JC, Friedlander TP, Wagner DM, Nielsen ES (2000) Nuclear genes resolve Mesozoic-aged divergences in the insect order Lepidoptera. *Molecular Phylogenetics and Evolution*, **15**, 242-259.

Wigglesworth VB (1948) The insects as medium for the study of physiology. *Proceedings of the Royal Society of London (B)*, **135**, 430-446.

Williams BL, Brawn JD, Paige KN (2003) Landscape scale genetic effects of habitat fragmentation on a high gene flow species: *Speyeria idalia* (Nymphalidae). *Molecular Ecology*, **12**, 11-20.

Yang ML, Wu HH, Guo YP, Ma EB (2005) Toxicological and biochemical characterizations of malathion sensitivity in two populations of *Oxya chinensis* (Orthoptera: Acridoidea). *Insect Science*, **13**, 41-47.

Yezerski A, Stevens L, Ametrano J (2003) A genetic linkage map for *Tribolium confusum* based on random amplified polymorphic DN As and recombinant inbred lines. *Insect Molecular Biology*, **12**, 517-526.

Yotoko KSC, Prado PI, Russo CAM, Solferini VN (2005) Testing the trend towards specialization in herbivore-host plant associations using a molecular phylogeny of

Tomoplagia (Diptera: Tephritidae) *Molecular Phylogenetics and Evolution*, **35**,701-711.

Zhang AB, Kubota K, Takami Y, Kim JK, Sota T (2005) Species status and phylogeography of two closely related *Coptolabrus* species (Coleoptera: Carabidae) in South Korea inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology*, **14**, 3823-3841.

Zhang DX (2004) Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology and Evolution*, **19**, 507-509.

Zhou X, Faktor O, Applebaum WS, Coll M (2000) Population structure of the pestiferous moth *Helicoverpa armigera* in the Eastern Mediterranean using RAPD analysis. *Heredity*, **85**, 251-256.

Zraket CA, Barth JL, Heckel DG, Abbott AG, (1990) Genetic linkage mapping with restriction fragment length polymorphisms in the tobacco budworm, *Heliothis virescens*. pp. 13-20 in Hagedorn, H.H., Hildebrand JG. and Kidwell M.G. (*Eds*) *Molecular insect science*. New York, Plenum Press.

Zwart KWR (1998) Ichneumonidae, p. 205-258 *in*: Polaszek A. (ed.), *African cereal stem borers: economic importance, taxonomy, natural enemies and control*. Wallingford, UK: CABI.