



**HAL**  
open science

# Evolution phylogéographique et structure génétique des populations de *Sitophilus zeamais* (Coleoptera, Curculionidae), ravageur du maïs stocke en Afrique de l'Ouest

Mama Racky Ndiaye

## ► To cite this version:

Mama Racky Ndiaye. Evolution phylogéographique et structure génétique des populations de *Sitophilus zeamais* (Coleoptera, Curculionidae), ravageur du maïs stocke en Afrique de l'Ouest. Génétique des populations [q-bio.PE]. Université Cheikh Anta DIOP de Dakar, Sénégal; Faculté des Sciences et Techniques; Département de Biologie Animale, 2018. Français. NNT: . tel-02116356

**HAL Id: tel-02116356**

**<https://hal.science/tel-02116356>**

Submitted on 30 Apr 2019

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

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

UNIVERSITE CHEIKH ANTA DIOP DE DAKAR



ÉCOLE DOCTORALE SCIENCES DE LA VIE, DE LA SANTÉ ET DE L'ENVIRONNEMENT



FACULTE DES SCIENCES ET TECHNIQUES  
DÉPARTEMENT DE BIOLOGIE ANIMALE

ANNEE 2018

N° 201830

**EVOLUTION PHYLOGEOGRAPHIQUE & STRUCTURE  
GENETIQUE DES POPULATIONS DE *SITOPHILUS ZEAMAI*  
(*COLEOPTERA, CURCULIONIDAE*), RAVAGEUR DU MAÏS  
STOCKE EN AFRIQUE DE L'OUEST**

THESE DE DOCTORAT UNIQUE EN BIOLOGIE ANIMALE

Spécialité : GENETIQUE DES POPULATIONS

Présentée par

**Mme Mama Racky NDIAYE**

---

Soutenue le 27 juillet 2018 à l'UCAD devant le jury

**PRESIDENT :**

**Pr Karamoko DIARRA**                      **Professeur titulaire**                      **FST/UCAD**

**RAPPORTEURS :**

**Dr Malick FALL**                              **Chargé d'enseignement**                      **FST/UCAD**

**Dr Saliou NDIAYE**                              **Maitre de Conférences tit.**                      **ENSA/UT**

**Dr Jean Yves ROUBEAU**                      **Directeur de Recherche**                      **Univ. MONTP.**

**EXAMINATEURS :**

**Dr Toffène DIOME**                              **Maitre de Conférences ass.**                      **FST/UCAD**

**Dr Cheikh THIAW**                              **Chargé de recherches**                      **CNRA/ISRA**

---

**DIRECTEUR DE THESE :**

**Pr Mbacké SEMBENE**                      **Professeur titulaire**                      **FST/UCAD**

## DEDICACES

---

Ce travail est dédié à la mémoire de trois personnes qui ont laissé des marques indélébiles dans ma vie :

- mon grand-père **EL-HADJ OMAR GUEYE BARRO CISSOKHO** qui m'a assisté et soutenu comme un père. Pour moi, tu constitues le symbole de la bonté, de la combativité, de la persévérance et de la tenacité. Que la Miséricorde Divine soit ton compagnon là où tu te trouves!!!

A ton ami de toujours : **Mame Ogo SECK**

*Que serions-nous devenus, si nul ne nous avait donné notre chance ?!!!*

**MERCI MAME OMAR!**

- ma grand-mère **SAFIETOU MAMA NDONG** : à toi qui incarnes la vertu, le stoïcisme, la générosité, l'humilité,... tant de valeurs qui font la dignité humaine!!! Ta dignité et ta bravoure sont pour moi les piliers fondateurs de ce que je suis.

**MERCI NA SAFIE!**

- ma grand-mère **ADJA SOUKEYNA SARR** : parmi tant de paroles de sagesse, je me souviendrai toujours celle où tu disais : « La vérité scientifique n'est pas loin ; elle est juste cachée ! Il faut procéder méthodiquement pour la trouver ». Ce qui, entre autres paroles, me motivait quand j'étais face à un obstacle dans mon domaine et prouve, une fois de plus, la sagacité de nos aïeux.

**MERCI MAME SOUKEY!**

***Que vous ne soyez plus là pour voir que l'énergie et l'effort que vous m'avait consacré n'ont pas été vains est un regret profond pour moi ; mais j'espère que là où vous vous trouvez, vous êtes fiers de moi !!!***

A mon **petit-frère MAMADY NDIAYE!**

Que le Bon DIEU fasse que tu puisses poursuivre ta jeunesse au paradis!!!

A mon **père SEYDOU NDIAYE!!!** Que la terre de Gambie te soit légère!

A mon grand-père **ABDOU KHADRE DIENE**, Que la lumière divine soit ton compagnon là où tu es !!

-----

A **M. NDARY DIOP**, juge au tribunal régional hors-classe de Dakar. Merci pour votre Altruisme!!!

A **Djily AMAR**, un oncle exceptionnel !

A mes couples préférés :

**M. et Mme Babacar NIANG, M. et Mme Salif SY, M. et Mme Omar FAYE**

Que le Bon Dieu préserve vos unions.

A **Abibatou NDOUR HYZAGI**, *yayou* AMINA!

***A LA TERRE DE MES ANCESTRES : L'ILE DE DIONEWAR (une des îles du Saloum) !!!***

## REMERCIEMENTS

*Grâce et Invocations à ALLAH (الله), le Tout Miséricordieux, le Très Miséricordieux, le Tout Puissant qui m'a donné cette chance de la vie de trouver ma voie.*

*Que Ta Sagesse nous guide! Que Ta lumière nous éclaire vers le chemin des Meilleurs!!!*

Ma gratitude est à l'endroit de chacun des membres de ce présent jury. Qu'il soit rassuré de toute ma reconnaissance pour avoir accepté de juger mon travail à travers leur expertise scientifique dont j'ai toute confiance.

Au **Pr Karamoko DIARRA**, je lui ai profondément reconnaissant d'avoir accepté de présider cette soutenance. Pour sa disponibilité, son engagement et son avis d'expert, j'exprime ma sincère reconnaissance et mon respect.

Au **Dr Saliou NDIAYE**, pour l'honneur qu'il nous fait en siégeant dans le jury de cette thèse, je lui adresse vivement l'expression de ma sincère reconnaissance et mon estime envers sa compétence.

Au **Dr Malick FALL**, pour l'attention qu'il nous porte en participant à ce jury et l'honneur qu'il nous fait de juger notre travail, je lui exprime ma sincère gratitude. Je profite de ces quelques lignes pour lui adresser ma reconnaissance et mon respect.

Au **Dr Jean Yves ROUBEAU**, je le remercie d'avoir accepté d'être un des rapporteurs de cette thèse et de me faire l'honneur d'être un des membres du jury.

Au **Dr Toffène DIOME**, je ne pourrai oublier sa disponibilité et son implication dans ce travail, pour les moments de complicité partagés et son soutien. Je lui adresse mes sincères remerciements pour le bel exemple qu'il nous a offert, de par ses compétences professionnelles et ses qualités humaines.

Au **Dr Cheikh THIAW**, je le remercie pour son soutien sans faille qu'il apporte à notre équipe, GENGESOP. Personnellement, je lui voue une sincère reconnaissance pour les appuis qu'il a su apporter à ma formation depuis ma licence ; ainsi que le soutien logistique et financier qu'il m'a apporté lors de mon stage au CERAAS de Thiès. Il n'a ménagé aucun effort pour que je le passe dans de meilleures conditions.

Au **Pr Mbacké SEMBENE**, je ne saurais jamais lui exprimer à sa juste valeur ma reconnaissance. En effet, il m'a accueilli dans sa « grande famille du savoir i.e l'équipe GENGESPOP » et m'a toujours soutenue surtout dans les moments de quasi-découragement, de doute avec une patience qui me surprenait même parfois. Je n'oublierai jamais sa tolérance et sa confiance qui m'ont laissé une grande liberté dans la réalisation de cette thèse tout en gardant un œil critique et scientifique. Sa compétence, sa rigueur scientifique, sa clairvoyance et sa patience m'ont beaucoup appris.

A travers cet encadrement depuis la licence, j'ai pu gagner, en plus d'un professeur émérite dont je suis toute fière, un père, un oncle, un ami. Que le Bon Dieu vous préserve, tonton Sembène !!!

Ce travail n'aurait pu être mené à bien sans nos partenaires au travers de leur soutien : le WAAPP ("West Africa Agricultural Productivity Program) pour le projet "Insect pests of grain" et le projet ERA (Education and Research in Agriculture) de l'USAID, le collège doctoral inter-régional de Biotechnologies végétales et agroalimentaire (BIOVEG) de l'Agence Universitaire de la Francophonie (AUF).

Ces années de recherches m'ont permis de rencontrer des personnes merveilleuses et exceptionnelles qui forment ma « famille scientifique », l'équipe GENGESPOP dont je ne saurais oublier le soutien amical et l'apport scientifique à mon travail, aux docteurs : Toffène DIOME, Penda NDIAYE, Fatima MBAYE SY, Khadim KEBE, Arame NDIAYE, Ndèye Ndiabé Mbaye, Amy Collé GUEYE, Asta E. Bambou, Ange C. Kafom et aux pré-docteurs Matar CISS, Adiouma SARR, Abdou K. M. DIA, Bineta KENEME (mon binome), Déthié Ngom, Makhtar Diop, Bakary NDIAYE et Babacar S. SAMBE (mes profs de statistiques), *etc...*

A mon mentor M. **Assane NDONG**, chef de projet à NESTLE Côte d'Ivoire. Merci beaucoup !!!

A mes collègues de la spécialité « Nutrition et alimentation humaine » : Doudou SYLLA, Abdou et Ousmane.  
A M. **SEYDI** et notre merveilleuse secrétaire **Adama Sané** !!!

Aux étudiants des sections L2BCG et L3SVT, du master de « Biologie animale », de l'ENDSS, de l'AIHES.  
Le chemin est long mais courage à vous !  
A mes protégés Marie Pierre FAYE & Codé DIOL.

Je remercie toutes les personnes formidables que j'ai rencontrées par le biais du collège doctoral de l'AUF.  
Merci pour votre soutien, vos encouragements et pour cette ambiance amicalement merveilleuse qui a caractérisé nos 4 rencontres dans des pays merveilleux (Tunisie, Camérout, Côte d'Ivoire et Maroc). Je ne saurais terminer sans remercier le comité scientifique du réseau BIOVEG, toutes ces personnes dont la contribution à mon travail est non négligeable : **M. Khalef Boulkroune, M. Serge Hamon, M. Khélifi Douadi, Mme Rachida Maouche, Mme Matel Wane et Mme Fanta BADJI, etc.**

Au **Pr Mme Mame Oureye SY**, je lui adresse un remerciement tout particulier pour son appui à ma candidature lors de l'appel à candidature du collège doctoral mais aussi pour sa rigueur scientifique, ses conseils professionnels et maternels. A mes deux compatriotes Mariama NGOM et Joel Romaric NGUEJOP.  
Je remercie les Dr Ndiaga Cissé et Daniel Foncéka et tout le personnel de la plateforme technique du CERAAS de m'avoir accueilli et permis de travailler dans des conditions conviviales.

A toute la famille **GUEYE** : mes oncles **MOUSSA** et **Amadou** ; mes tantes et marraines **NDEYE ASTOU** et **NDEYE COUMBA**. Merci beaucoup pour votre assistance!!! Que le bon DIEU rallonge votre longévité, vous accorde une santé de fer et que vous soyez présentes aux succès et aux bonheurs que nous attendons de la vie!!!

Je ne saurais jamais oublier l'assistance et l'aide accompagnées d'un altruisme sans égal de mon oncle **DJILY AMAR** et sa femme **NDOUMBE** ! Que la miséricorde divine soit le socle de votre famille!!!

A **M. et Mme SANOKHO**, MERCI INFINIMENT !!!

A ma mère Daba NDONG, mon beau-père Bouba NDIAYE et ma petite sœur Sophie,

A mes oncles exceptionnels Mamadou et Mame Serigne NDONG. A mes tantes Adj, Sophie et Fatou NDONG.

A ma jumelle Rane Wade BA et nos deux trésors : Babacar et Baye Zale

A tonton Ba et M. Goudiaby. Merci pour votre Mansuétude!!!

A mes cousins : Fatou et Olimata NDIAYE ; Mariama Sally, Nama, Mouhamed, Mame Fatou et Alima DIOUF ; Astou, Elimane, Soukeyna, Ma Khady et Mame Serigne Jr NDONG

A mon grand-frère Moustapha DIOUF, merci ! A Fatou Sarr FAYE, Tabaski Sarr, Fatou Guèye Mme DIOUF.

A mes tontons Kalidou CISSE, Mame GOOR et mon frangin Papis CISSE, Imam Seck & Youssou Diané.

A Diatou Ngom et son mari Omar Faye, Fatou Thiaw Mme TRA, Aïssatou Yvette, Ngoné Diop, Amina...

A mes sœurs en Gambie! A toute la famille Ndiaye en Gambie!!! A, Ara, Aïda, Adé et Assiétou Sarr !

**Finalement à mon trésor chéri Amina CISSE**, puisses tu trouver dans ce travail un exemple à passer !!!

***Ma profonde gratitude a tout ceux qui m'ont, un jour ou l'autre, donné envie d'y croire encore et dont le soutien m'a permis de ne jamais dévier de mon objectif final.***

***Merci pour vos encouragements dans les moments de doutes.***

***Même si je ne vous cite pas sur cette page, mon cœur ne vous oublie pas!!!***

« **JOKO NJAAL** » (merci dans ma langue maternelle, sérère)  
« **A DIARAMA** » (merci dans ma langue paternelle, toucouleur)

***Passionnément l'Entomo-Génétique !!!***

La connaissance de la distribution spatiale et temporelle de la diversité génétique au sein et entre les populations d'insectes ravageurs est une étape cruciale pour l'établissement de stratégies de gestion des cultures et des produits stockés pour assurer la sécurité alimentaire et nutritionnelle. En outre, le charançon du maïs ou *Sitophilus zeamais* (Motschulsky, 1855) est connu en raison de son extraordinaire potentiel de destruction des céréales stockées ; mais peu de données sont connues sur la phylogéographie et la structure génétique de ce ravageur en Afrique subsaharienne. De ce fait, l'objectif de cette thèse est d'expliquer les relations phylogéographiques entre les populations et d'évaluer leurs flux géniques. Ceci permettrait de mettre en exergue l'impact des zones biogéographiques et des activités commerciales sur la structure génétique des populations grâce à des analyses multivariées et bayésiennes à l'aide de données de séquençage et de génotypage.

Ainsi, l'analyse de séquences des gènes mitochondriaux, *Cytochrome B* et *Cytochrome oxydase I*, a été faite avec 112 et 109 individus respectivement, dans 7 pays d'Afrique de l'ouest (Burkina Faso, Côte d'Ivoire, Ghana, Mali, Niger, République de Guinée et Sénégal) auxquels 2 pays d'Afrique centrale ont été ajoutés (Caméroun et Centrafrique). Notre étude a révélé la présence de 30 haplotypes de par les séquences concaténées. L'haplotype majoritaire a représenté environ 43% de la population totale et a relié l'Afrique de l'Ouest à l'Afrique centrale. Les analyses moléculaires ont indiqué la présence d'une structure génétique sans signal phylogéographique. Ceci est la résultante d'un flux multidirectionnel asymétrique qui serait très probablement lié aux impacts anthropiques à travers les activités commerciales. Elles ont montré également que la plus grande partie de la variance moléculaire est expliquée par les caractéristiques individuelles au sein des populations alors que peu est liée aux déterminants écologiques des régions biogéographiques. L'étude de l'histoire évolutive des populations a suggéré que l'infestation de la zone sahélo-soudanienne serait plus ancienne que celle de la région guinéo-congolaise. Les inférences multivariées et bayésiennes ont suggéré l'existence de 5 principaux *clusters* qui se sont séparés après un goulot d'étranglement suivi d'une expansion démographique modérée depuis la période du pléistocène.

Pour mieux appréhender ces résultats, 10 loci microsattellites ont été génotypés pour 90 individus dans 6 pays (Burkina Faso, Côte d'Ivoire, Ghana, Niger, Mali et Sénégal). Dès lors, ces marqueurs SSR ont confirmé l'existence d'une différenciation génétique et d'un flux génique asymétrique et multidirectionnel ; d'où l'existence d'une structure génétique sans pattern phylogéographique. Ceci serait lié à la dispersion qui est le résultat d'une migration passive par les échanges commerciaux. En effet, il a été démontré une affinité génétique entre le Burkina Faso, le Niger et la Côte d'Ivoire et une différenciation génétique importante entre le Mali, le Ghana et le Sénégal. L'analyse de la variance moléculaire a réaffirmé que les caractéristiques individuelles ont concentré l'essentiel de la variance totale. L'analyse des populations a aussi révélé les effets d'un *bottleneck* démographique récent et d'une potentielle sélection directionnelle ayant une action érosive de la biodiversité. Le *clustering* par analyses multivariées et bayésiennes des relations génétiques ont mis en évidence de 4 *clusters* génétiques.

En définitif, il existe une congruence entre les résultats de l'analyse de marqueurs de l'ADNmt avec ceux des SSR : la présence d'une structure génétique sans signal phylogéographique résultant d'une différenciation génétique inégale et flux génique multidirectionnel (infestation croisée) ; les populations ont aussi présenté un bon potentiel de viabilité renforcé par une dispersion passive à travers le commerce des grains ; la diversité génétique est plus importante dans le Nord d'où partiraient originellement les voies de dispersion.

Cette étude a fourni de nouveaux aperçus dans la description des caractéristiques et de la différenciation génétiques des populations de *Sitophilus zeamais* dans la sous-région ouest-africaine et centrale. En effet, elle a permis de mettre en exergue la biodiversité de *S. zeamais* dans la sous-région. Elle a aussi permis de comprendre les processus qui gouvernent les paramètres biogénétiques de ces populations or la connaissance de la biologie évolutive permet d'influencer sur ces paramètres pour diminuer voire éliminer les effets néfastes de ces ravageurs. Par conséquent, il constitue donc un outil dans les perspectives de gestion durable des ressources agricoles, particulièrement le maïs.

**MOTS-CLES :** *Sitophilus* ; *zeamais* ; Phylogéographie ; Flux génique ; *Cytochrome b* ; *Cytochrome oxydase I* ; SSR ; Zones biogéographiques ; Afrique de l'Ouest ; Afrique subsaharienne.

Knowledge of spatial and temporal distribution of genetic diversity within and between populations is a crucial step for establishing strategic management of stored crops to assure food safety. The maize weevil or *Sitophilus zeamais* (Motschulsky, 1855) is known because of its extraordinary potential for destroying stored grain; but little is known about its phylogeography and genetic structure in sub-Saharan Africa. The aim of this thesis is to explain the phylogeographical relationships of populations afterwards to evaluate gene flow between them. This would allow to highlight the influence of biogeographical zones and anthropogenic activities on maize weevil populations by multivariate and Bayesian analyses using sequencing and genotyping data.

Thereby, the analyses of sequences of mitochondrial genes, *Cytochrome B* and *Cytochrome oxidase I*, were done with 112 and 109 samples respectively, in 7 countries in West Africa (Burkina Faso, Ivory Coast, Ghana, Mali, Niger, Republic of Guinea and Senegal) and 2 countries in Central Africa (Cameroon and Central Africa). Our study revealed the presence of 30 haplotypes characterized by concatenated sequences. The majority haplotype represented about 43% of the total population. It connected West Africa to Central Africa. The molecular analyses indicated a genetic structure without phylogeographic pattern. This was the results of a multidirectional flow which would be probably related to impacts of commercial activities. The study of the evolutionary history of the global population suggested that the infestation of the North would be longer than the South one. Multivariate and Bayesian inferences suggested the existence of 5 main *clusters* that have separated after a bottleneck followed by a moderate population expansion since the Pleistocene period.

In order to better understand these results, 10 microsatellite loci were genotyped with 90 samples in 6 countries (Burkina Faso, Ivory Coast, Ghana, Niger, Mali and Senegal). These SSR markers indicated the existence of unequal genetic differentiation and gene flow. This asymmetric flow was multidirectional; hence the existence of a genetic structure without phylogeographical pattern. This was related to dispersion which is mainly linked to passive migration due to commercial activities. Besides, the populations' analyses have also revealed effects of a recent bottleneck and a potential directional selection. The clustering by multivariate and Bayesian analyses pointed up 4 genetic *clusters*.

Finally, there was a congruence between the results of mtDNA marker analysis with those of SSR: presence of a genetic structure without phylogeographic pattern resulted from an unequal genetic differentiation and multidirectional gene flow. Besides, populations showed a good viability potential enhanced by passive dispersal through the grain trade. Genetic diversity was more important in the North. So, the infestation would come from the North to the South.

This study provided novel insights into the description of the genetic basis and differentiation. Indeed, it has highlighted the biodiversity of *S. zeamais* in the sub-region. Thus, it has enabled for understanding of processes which drove genetic diversity of populations in West African sub-region, because knowledge of evolutionary biology can influence on these parameters in order to increase and even eliminate harmful effects of these pests. Therefore, it constitutes a tool in perspectives for sustainable management of resources.

**KEY-WORDS:** *Sitophilus; zeamais; Phylogeography; Gene flow; Cytochrome B; Cytochrome oxidase I; SSR; biogeographical zones; West Africa; Sub-Saharan Africa.*

## SIGLES ET ACRONYMES

---

**ADN:** Acide Désoxyribonucléique  
**AIC:** *Akaike Information Criterion*  
**ABC:** *Approximated Bayesian Computation*  
**ACP :** Analyse en Composantes Principales  
**ADNmt:** ADN mitochondrial  
**AFC :** Analyse Factorielle de Correspondance  
**AICc:** *Akaike Information Criterion corrected*  
**AMOVA:** Analysis of MOlecular VAriance  
**ASS:** Afrique SubSaharienne  
**ATP:** Acide Triphosphate  
**BAPS:** Bayesian Analysis Population Structure  
**BEAST:** Bayesian Evolutionary Analysis by Sampling Trees  
**BEAUi:** Bayesian Evolutionary Analysis Utility  
**BIOPASS :** Laboratoire de Biologie des Populations Animales Sahélo-Soudaniennes  
**BP:** Before Present (1950, date de référence par la méthode carbone 14)  
**CEDEAO :** Communauté Economique des États de l'Afrique de l'Ouest  
**CERAAS :** Centre de Recherche d'Amélioration de l'Adaptation à la Sécheresse  
**CFA :** Communauté financière d'Afrique  
**CILSS :** Comité permanent Inter-état de Lutte contre la Sécheresse dans le Sahel  
**CIMMYT:** International Centre for Maize and Wheat Improvement Center  
**COI :** *Cytochrome oxydase I*  
**OCDE :** Club du Sahel et de l'Afrique de l'Ouest  
**Cyt-B:** *Cytochrome B*  
**CACID :** Centre Africain pour le Commerce l'Intégration et le Développement  
**ESASU :** Enquête de Sécurité Alimentaire en Situation d'Urgence en milieu rural  
**FAO :** Organisation des Nations unies pour l'alimentation et l'agriculture  
**FIDA :** Fonds International de Développement Agricole  
**GGC :** Groupe des pays du domaine guinéo-congolais  
**GSA :** Groupe des Pays du Sahel  
**HWE:** *Hardy-Weinberg Equilibrium*  
**IBD:** *Isolement By Distance*  
**IRD :** Institut de recherche pour le développement  
**UEMOA :** Union Economique et Monétaire Ouest Africaine  
**LD:** *Lynkage Desequilibrium*  
**LRT:** *Likelihood Ratio Test*  
**MCMC :** *Markov Chain Monte Carlo*  
**MCMCMC:** *Metropolis Coupling Markov Chain Monte Carlo*  
**MRCA:** *Most Recent Common Ancestor*  
**NCPA:** *Nested Clade Phylogeographic Analysis*  
**NJ:** *Neighbor-Joining method*  
**Pb:** Paire de bases  
**PCR:** *Polymerase Chain Reaction*  
**PIC:** Polymorphism Information Content  
**Rg:** (Raggedness statistic)  
***S. oryzae :*** *Sitophilus oryzae*  
***S. zeamais :*** *Sitophilus zeamais*  
**SAMOVA** (Spatial Analysis of MOlecular Variance  
**SGPE:** *Sitophilus granarius* Primary Endosymbiont  
**SOPE:** *Sitophilus oryzae* Primary Endosymbiont  
**SSD:** Sum of Squared Deviation  
**SSR:** Simple Sequence Repeat  
**SZPE:** *Sitophilus zeamais* Primary Endosymbiont  
**UPGMA:** Un-weighted Pair-group Method with an Arithmetic average  
**Ver.:** Version  
**WECAMAN:** West and Central Africa Collaborative Maize Research Network.



## LISTES DES FIGURES

Figure 1 : Morphologie de <i>Sitophilus zeamais</i> .....	14
Figure 2 : La distribution de l'endosperme dans les cinq types de grains de maïs.....	19
Figure 3 : Les sous-ensembles économiques de l'Afrique de l'Ouest ; OCDE, 2006.....	21
Figure 4 : Principaux bassins de production des cultures vivrières (Blein <i>et al.</i> , 2008) .....	21
Figure 5 : Les différentes zones agro-écologiques de l'Afrique de l'Ouest ; FAO, 2004.....	23
Figure 6 : Zone d'étude et localités échantillonnées.....	37
Figure 7 : Correspondance entre la phylogénie des haplotypes et leur distribution géographique. ....	47
Figure 8 : Cartes de distribution spatiale des haplotypes de <i>Sitophilus zeamais</i> . .....	56
Figure 9 : Réseaux d'haplotypes de <i>S. zeamais</i> . .....	57
Figure 10 : Analyse en composantes principales des populations .....	59
Figure 11: Classification hiérarchique ascendante des populations.....	59
Figure 12: Test de Mantel (Distances géographiques en fonction des distances génétiques).....	60
Figure 13: Cartes de distribution spatiale des groupes génétiques <i>Sitophilus zeamais</i> . .....	61
Figure 14 : Répartition de la diversité génétique chez <i>S. zeamais</i> en Afrique de l'Ouest.....	62
Figure 15: <i>Clustering</i> des populations par le modèle mixture avec les Séquences concaténées .....	62
Figure 16: <i>Clustering</i> des populations par le modèle mixture avec le <i>Cyt-B</i> .....	62
Figure 17 : <i>Clustering</i> des populations par le modèle mixture avec le <i>COI</i> .....	63
Figure 18 : Réseau d'haplotypes en clades emboîtés des séquences concaténées .....	64
Figure 19 : Mismatch distributions par paires d'individus de la population globale.....	65
Figure 20 : Mismatch distributions par paires d'individus de la population guinéo-congolaise .....	65
Figure 21 : Mismatch distributions par paires d'individus de la population sahélienne .....	65
Figure 22 : Diagrammes triangulaires du test « Four-cluster Likelihood-mapping ».....	67
Figure 23: Inférence bayésienne des haplotypes des séquences concaténées .....	68
Figure 24 : Neighbor-joining des haplotypes des séquences concaténées .....	68
Figure 25 : Phylogénie des haplotypes des séquences concaténées par le MP et MV .....	68
Figure 26 : Inférence bayésienne des haplotypes par le <i>Cyt-B</i> et le <i>COI</i> .....	69
Figure 27 : Origine probable de l'infestation et voies de dispersion de <i>Sitophilus zeamais</i> .....	71
Figure 28 : Modèles alléliques dans les populations .....	98
Figure 29 : Différenciation génétique entre populations .....	101
Figure 30 : Classification hiérarchique ascendante des populations .....	101
Figure 31 : Les réseaux standards de migration relative directionnelle des populations .....	102
Figure 32 : Relation de regression avec corrélation de Spearman .....	104
Figure 33 : Analyses des loci sous sélection .....	105
Figure 34 : Modèles alléliques dans la population globale.....	105
Figure 35 : Diagramme de l'Analyse en Composantes Principales (ACP).....	106
Figure 36 : Mise en évidence des relations génétiques entre populations par AFC .....	107
Figure 37 : Contributions à la fonction discriminante 1 (a) et 2 (b) des allèles .....	108
Figure 38 : Diagramme de dispersion de l'analyse du DAPC .....	108
Figure 39 : Assignment des individus selon le DAPC.....	108
Figure 40 : Détermination de K le plus probable .....	109
Figure 41 : Assignment des individus aux différents groupes génétiques .....	109
Figure 42 : Evolution de la structuration génétique des populations.....	110
Figure 43 : Contenu des populations en fonction des 4 <i>clusters</i> .....	110
Figure 44 : Inférence de la structure génétique des <i>clusters</i> .....	111
Figure 46 : Neighbor-Joining à partir de DAS : analyses avec (a) et sans (b) le locus 3G7 .....	113
Figure 45 : Neighbor-Joining à partir de la distance de Nei (1983) et (a) à partir de DAS (b) .....	113
Figure 47 : Spectre des fréquences alléliques des populations du Charançon du maïs .....	114

## TABLE DES MATIERES

Dédicaces .....	II
Remerciements .....	III
Résumé .....	V
Abstract .....	VI
Sigles et acronymes .....	VII
Listes des figures .....	VIII
Table des matières .....	IX
Préambule .....	XV
<b>Introduction générale</b> .....	2
<b>Chapitre I : Présentation de modèles d'étude, cadre agro-écologique et méthodologie en génétique des populations</b> .....	7
Introduction .....	7
I. 1 Modèles d'étude : <i>Sitophilus zeamais</i> (Motschulsky, 1855)/ <i>Zea mays</i> (Linné, 1753) .....	7
I.1.1 Généralités sur les insectes coléoptères phytophages .....	7
I.1.2 Généralités sur les « charançons » .....	8
I.1.3 Présentation du « charançon » : <i>Sitophilus zeamais</i> (Motschulsky, 1855) .....	9
I.1.3.1 Position systématique .....	10
I.1.3.2 Origine et répartition géographique .....	11
I.1.3.3 Importance agro-économique et sanitaire .....	11
I.1.3.4 Caractéristiques morphologiques, phénotypiques et génétiques .....	12
I.1.3.5 Biologie de la reproduction et cycle de développement .....	14
I.1.3.6 Particularités symbiotiques des endobactéries .....	15
I.1.3.7 Parasitoïdes du genre <i>Sitophilus</i> .....	16
I.1.3.8 Présentation synoptique des études sur <i>Sitophilus</i> spp. ....	17
I.2 Présentation de la plante-hôte : maïs ( <i>Zea mays</i> L., 1753) .....	18
I.3 Contexte agro-économique et politiques d'intégration sous-régionale .....	20
I.3.1 Organismes d'intégration dans la situation agro-économique sous régionale .....	20
I.3.2 Les différentes zones agro-écologiques de la sous-région ouest-africaine .....	22
I.3.3 Particularités de la zone Sahélo-soudanienne et de la région guinéo-congolaise .....	23
I.4 Concepts et notions fondamentales en génétique des populations .....	24
I.4.1 Espèce .....	24
I.4.2 Sous-espèce .....	25
I.4.3 Population .....	25
I.4.4 Métapopulation .....	25
I.4.5 Ecotype .....	26
I.4.6 Biotype .....	26
I.5 Approches et méthodologie utilisées en génétique des populations .....	26
I.5.1 Quantification de l'évolution par approche prospective .....	26
I.5.1.1 Loi de Hardy-Weinberg d'une population infinie .....	26
I.5.1.2 Modèle de Wright-Fisher d'une population finie .....	28
I.5.2 Quantification de l'évolution par approche rétrospective : théorie de la coalescence .....	29
I.5.3 Analyses statistiques par les tests d'hypothèses .....	29
I.5.4 L'inférence bayésienne, la phylogénie et la génétique des populations .....	30
I.5.4.1 Inférence bayésienne et phylogénie .....	30

I.5.4.2 Inférence bayésienne et génétique des populations.....	32
<b>Chapitre II : Evolution phylogéographique et structure genetique des populations de <i>Sitophilus zeamais</i>, ravageur du maïs stocke en Afrique de l’Ouest</b> .....	35
Introduction.....	35
II.1 Matériel et méthodes .....	36
II.1.1 Echantillonnage .....	36
II.1.2 Etude Génétique .....	37
II.1.2.1 Intérêt du <i>Cytochrome B</i> et du <i>Cytochrome Oxydase I</i> .....	37
II.1.2.2 Extraction D’ADN .....	39
II.1.2.3 Amplification par la technique de la PCR .....	40
II.1.2.4 Séquençage .....	41
II.1.3 Analyses génétiques et statistiques.....	41
II.1.3.1 Alignements et corrections des séquences.....	41
II.1.3.2 Diversité génétiques des populations.....	41
II.1.3.3 Structure génétique et relations phylogéographiques des populations .....	43
II.1.3.3.1 Différenciation et flux génétiques inter-populations.....	43
II.1.3.3.2 Distances génétique et isolement par la distance .....	44
II.1.3.3.3 Distribution de la variance génétique et répartition spatiale.....	45
II.1.3.3.3.1 Source de la variance moléculaire.....	45
II.1.3.3.3.2 <i>Clustering</i> par une analyse bayésienne de la structure génétique .....	45
II.1.3.3.3.3 Détection de signal phylogéographique .....	45
II.1.3.3.4 Structure des populations et processus historiques .....	48
II.1.3.4 Evolution démographique .....	48
II.1.3.5 Analyses des relations phylogénétiques.....	50
II.1.3.5.1 Réseaux des haplotypes.....	50
II.1.3.5.2 Reconstructions Phylogénétiques.....	51
II.1.3.5.3 Datations moléculaires .....	53
II.2 Résultats .....	55
II.2.1 Diversité génétiques des populations .....	55
II.2.1.1 Polymorphisme et évolution génétiques .....	55
II.2.1.2 Diversité et répartition des haplotypes.....	56
II.2.1.3 Diversité génétique des populations .....	57
II.2.1.4 ACP des indices de diversité génétique des populations .....	58
II.2.2 Structure génétique et relations phylogéographiques des populations .....	59
II.2.2.1 Divergence et différenciation génétiques des populations.....	59
II.2.2.2 Isolements génétiques .....	60
II.2.2.3 Distribution de la variance génétique et répartition spatiale.....	60
II.2.2.3.1 Source de la variance moléculaire (AMOVA).....	60
II.2.2.3.2 Analyse spatiale de la variance moléculaire .....	61
II.2.2.4 Résumé de la répartition de la diversité génétique .....	62
II.2.2.5 <i>Clustering</i> par analyse bayésienne de la structure génétique .....	62
II.2.2.6 Indices de différenciation génétique et détection de signal phylogéographique .....	63
II.2.2.7 Processus historiques et structure des populations .....	63
II.2.3 Evolution démographique des populations.....	64
II.2.4 Relations phylogénétiques des haplotypes .....	66
II.2.4.1 Arbres phylogénétiques et signal phylogénétique .....	66

II.2.4.2	Divergence et structure des haplogroupes identifiés .....	69
II.2.4.3	Analyse de la coalescence et temps de divergence .....	70
II.3	Discussion .....	72
III.3.1	Diversité génétique des populations .....	72
III.3.2	Structure génétique et phylogéographique .....	73
III.3.2	Affinités phylogénétiques, biogéographie et histoire évolutive .....	74
III.3.3	Evolution démographique .....	75
	Conclusion partielle .....	77
	<b>Chapitre III : Structure génétique et flux génique entre populations de <i>Sitophilus zeamais</i> (motschulsky, 1855), ravageur de maïs stocké en Afrique de l’Ouest.....</b>	<b>79</b>
	Introduction.....	79
III.1	Matériel et méthodes.....	80
III.1.1	Echantillonnage .....	80
III.1.2	Génotypage des microsatellites de <i>Sitophilus zeamais</i> .....	80
III.1.2.1	Intérêt et limites des microsatellites .....	80
III.1.2.2	Caractérisation du procédé de génotypage des populations.....	81
III.1.2.2.1	Extraction de l’ADN génomique .....	81
III.1.2.2.2	Génotypage des loci de <i>S. zeamais</i> .....	81
III.1.3	Analyses génétiques et statistiques .....	82
III.1.3.1	Lecture, correction des génotypes et constitutio du jeu de données.....	82
III.1.3.2	Variabilité et diversité génétiques des populations .....	82
III.1.3.2.1	Nombre moyen d’allèles par locus ( $N_A$ ).....	82
III.1.3.2.2	Indice de polymorphisme (I.P.) .....	83
III.1.3.2.3	Fréquences alléliques ( $F_a$ ).....	83
III.1.3.2.4	Taux d’hétérozygotie .....	83
III.1.3.2.5	Contenu Informatif du polymorphisme et index informatif de Shannon.....	83
III.1.3.2.6	Richesse allélique .....	84
III.1.3.3	Equilibre génétique.....	84
III.1.3.3.1	Test de conformité à l’équilibre Hardy-Weinberg.....	84
III.1.3.3.2	Test de l’équilibre de liaison.....	85
III.1.3.4	Différenciation et structure génétiques inter-populations .....	85
III.1.3.4.1	Différenciation génétique et flux génique .....	86
III.1.3.4.1.1	Différenciation génétique des populations.....	86
III.1.3.4.1.2	Flux génique et nombre de migrants par génération.....	87
III.1.3.4.2	Facteurs de différenciation et de structuration génétiques.....	88
III.1.3.4.2.1	Isolement par la distance.....	88
III.1.3.4.2.2	Sources de la variance moléculaire .....	89
III.1.3.4.2.3	Impact des loci dans la différenciation génétique.....	89
III.1.3.4.2.4	Détection de signal phylogéographique.....	90
III.1.3.5	<i>Clustering</i> , organisation spatiale et introgression dans les populations .....	90
III.1.3.5.1	Analyse multivariée : AFC, ACP et DAPC.....	90
III.1.3.5.2	Analyse bayésienne.....	91
III.1.3.5.3	Impact de la mutation/flux génique sur la différenciation génétique .....	93
III.1.3.6	Analyses phylogénétiques .....	93
III.1.3.7	Evolution démographique.....	94
III.2	Résultats .....	95

III.2.1 Screening des loci microsatellites .....	95
III.2.2 Diversité et équilibre génétiques des populations .....	95
III.2.3 Variabilité allélique et diversité génétique des populations .....	97
III.2.4 Différenciation et structure génétiques des populations .....	98
III.2.4.1 Différenciation génétique et flux génique .....	98
III.2.4.2 Différenciation génétique entre populations du Nord et celle du Sud.....	102
III.2.4.3 Facteurs de différenciation et de structuration génétiques .....	103
III.2.4.3.1 Sources de la variance moléculaire.....	103
III.2.4.3.2 Isolement par la distance génétique .....	103
III.2.4.3.3 Impact des loci dans la différenciation génétique des populations.....	104
III.2.4.3.4 Détection de signal phylogéographique.....	105
III.2.3 <i>Clustering</i> , organisation spatiale et introgression des populations .....	106
III.2.3.1 Analyse multivariée .....	106
III.2.3.1.1 Analyse principale des correspondances (ACP).....	106
III.2.3.1.2 Analyse factorielle de correspondance (AFC).....	106
III.2.3.1.3 Analyse discriminante des composantes principales (DAPC).....	107
III.2.3.2 Analyse bayésienne .....	109
III.2.3.2.1 Structure basée sur le modèle de Hardy-Weinberg .....	109
III.2.3.2.2 Détection de migrants de première génération .....	111
III.2.3.2.3 Influence de la mutation et du flux génique sur la différenciation génétique .....	112
III.2.4 Reconstructions phylogénétiques .....	112
III.2.5 Evolution démographique : signature de goulot d'étranglement .....	113
II.3 Discussion .....	115
III.3.1 Variabilité allélique et diversité génétique des populations.....	115
III.3.2 Diversité et équilibre génétiques des populations .....	115
III.3.3 Répartition hiérarchique et spatiale de la diversité génétique.....	116
III.3.4 Populations de <i>Sitophilus zeamais</i> sous sélection directionnelle .....	119
III.3.5 Evolution démographique .....	120
Conclusion partielle .....	122
<b>Conclusion générale .....</b>	<b>123</b>
Recommandations.....	124
Perspectives.....	125
Références bibliographiques .....	126
Annexes.....	152
Publications et communications.....	156

A mes Grand-Parents : Safiétou, Omar & Soukeyna !!!

Jeunes africains, armés vous en science jusqu'aux dents, et si vous manquez d'unir l'Afrique, vous ne serez pas au rendez-vous du troisième millénaire.

Pr Cheikh Anta DIOP

Il n'est d'un homme raisonnable de blamer par caprice l'étude des insectes, ni de s'en dégouter par la considération des peines qu'elle donne.

La nature ne renferme rien de bas.

Tout y est sublime, tout y est digne d'admiration.

Aristote

La trilogie :

- recherche scientifique,
- l'application des découvertes et
- la formation du personnel sous l'impulsion des investissements constitue le pilier du développement durable.

*[www.solidariteetprogres.org](http://www.solidariteetprogres.org)*

Si tu ne sais pas où tu vas, regarde d'où tu viens.

Proverbe africain

Avoir une bonne production agricole, c'est bien !  
Savoir bien la conserver, c'est encore mieux !!!

La recherche scientifique constitue un élément-moteur du développement durable. Dans le vaste domaine de l'agriculture, l'objectif fondamental de cette recherche est l'atteinte d'une agriculture durable et par conséquent d'une sécurité alimentaire et nutritionnelle en parallèle avec une préservation de l'environnement. Ainsi, la contribution de la recherche scientifique s'appuie sur une trilogie : l'identification des problèmes, la prévision des évolutions futures et la proposition de solutions dont les priorités tournent autour de la recherche de l'équilibre entre les 3 principales finalités du développement durable à savoir l'aspect social, économique et environnemental. Par ailleurs, la protection des récoltes et la conservation des denrées stockées sont des facteurs essentiels pour assurer la sécurité alimentaire et nutritionnelle. Selon FAO (2011a), les pertes post-récoltes pourraient combler les besoins minimum annuels d'au moins 48 millions de personnes alors que plus de 700 millions de personnes souffrent de la sous-alimentation dans le monde (FAO *et al.*, 2015). Ceci montre que le volet post-récolte de l'agriculture constitue une question cruciale et mérite, plus que jamais, d'être considéré comme une question de survie.

L'enjeu est donc d'identifier les facteurs qui déterminent le succès invasif et dépréciateur des organismes qui dégradent les produits agricoles dans les champs et dans les infrastructures de stockage particulièrement leur biologie évolutive afin de mieux prédire et de contrôler voire supprimer leurs actions dépréciatives. A cet effet, la génétique des populations constitue un bon outil d'investigation car elle explore les groupes d'individus à une échelle très fine, la diversité génétique qui est l'un des 3 niveaux de la biodiversité. Dès lors, le but de cette étude est, entre autres, d'utiliser les principes de la génétique des populations, les outils et techniques de la biologie moléculaire, de la bio-informatique et des statistiques, ainsi que l'inférence bayésienne pour déterminer les caractéristiques génétiques des populations de *Sitophilus zeamais* infestant le maïs. En effet, une des étapes cruciales de la lutte intégrée, alternative à la voie chimique (dangereuse et coûteuse), est la connaissance du statut bioévolutif des populations incriminées (Sembène, 2000). Dès lors, l'obtention de ces informations, plus particulièrement l'évolution, l'intensité et la structure de la variabilité génétique et les facteurs qui lui sont liés contribuerait à identifier une faiblesse dans la biologie évolutive de ces insectes ravageurs. En effet, la variabilité génétique constitue le maillon fondamental de la biodiversité car le destin des populations, l'apparition, tout comme l'extinction des espèces, dépendent de cette variabilité génétique (Lacy, 1987 ; Montaigne, 2011 ; Greenbaum *et al.*, 2014). Ceci permettrait de définir des scénarii de gestion de ces populations phytophages afin d'intervenir ultérieurement et efficacement par des méthodes respectueuses de la santé des populations et de la préservation de l'environnement. Cette thèse trouve donc son utilité dans les principes du développement durable, particulièrement dans le domaine agricole et sûrement dans le volet post-récolte.



## INTRODUCTION GENERALE

## INTRODUCTION GENERALE

La sécurité alimentaire et nutritionnelle exige une disponibilité des ressources alimentaires, en qualité et en quantité, à travers une bonne production et surtout une bonne conservation des produits agricoles. Par ailleurs, la notion de sécurité alimentaire, jadis considérée comme une question d'adéquation entre l'offre et la demande alimentaires et que pour l'atteindre les politiques de sécurité alimentaire se résument à augmenter la production agricole et/ou à ralentir l'accroissement démographique, a essentiellement évolué. En effet, à la suite de la conférence mondiale de l'alimentation de 1996 à Rome, elle a été redéfinie et elle est dite assurée quand toute personne a, à tout moment, un accès physique et économique à une nourriture suffisante, saine et nutritive lui permettant de satisfaire ses besoins énergétiques et ses préférences alimentaires pour mener une vie saine et active (FAO, 1996 ; FAO *et al.*, 2013). Dès lors, la sécurité alimentaire est axée sur 4 piliers : l'accès par la disposition d'un pouvoir d'achat suffisant pour le faire ; la disponibilité des ressources alimentaires à travers la conservation des récoltes ; la qualité de l'alimentation et la régularité. De ce fait, l'agriculture constitue la source vitale de l'alimentation qui est l'un des besoins primaires des animaux en général et de l'homme en particulier. Elle constitue, en effet, un maillon central dans la chaîne agroalimentaire en lui assurant l'approvisionnement en matières premières (céréales, féculés, fruits, légumineuses, *etc...*). En plus, elle constitue le « moteur de croissance » en Afrique ; et les céréales telles que le sorgho, le millet, le blé, le maïs et le riz sont des aliments de base majeurs pour la majorité de la population ; avec le rôle central du maïs comme aliment de base en Afrique subsaharienne (ASS) qui est comparable à celui du riz ou du blé en Asie (Macaulley et Ramadjita, 2015). Néanmoins, les changements climatiques accompagnés de la dégradation de la nature et de la croissance démographique galopante ont des effets négatifs sur la productivité agricole. En effet, selon Schlenker et Lobell (2010), en raison du réchauffement climatique en ASS, la production de céréales majeures pourrait décliner jusqu'à 20% d'ici le milieu du siècle. Cependant, par l'adoption d'une gamme de variétés tolérantes au facteur abiotique qu'est la sécheresse, le maïs deviendrait une culture commerciale et assurerait la sécurité alimentaire mieux que toute autre culture (Smith *et al.* 1997). Selon Rosegrant *et al.* (2008), il est prévu pour devenir la plus grande récolte en termes de production d'ici 2050.

Pour répondre aux besoins alimentaires d'une population, en augmentation progressive, qui devrait atteindre un peu plus de 9 milliards d'individus à l'échelle mondiale à l'aube 2050 (Dufumier, 2010 ; FAO, 2011b) ; la production agricole devrait augmenter de 60 à 70 % à l'échelle mondiale d'ici 2050 ; et de près de 100% dans les pays en développement (FAO, 2011b ; FIDA, 2011 ; 2012). Par conséquent, il faudra d'autant plus faire une meilleure prise en charge des questions relatives au volet pré-récolte mais aussi au volet post-récolte et donc aux populations qui détériorent et déprécient les cultures et denrées stockées. Ce qui permettra d'éviter que le monde s'expose aux risques de

situations de famines et/ou d'intoxications alimentaires. Cela permettrait aussi de pouvoir sortir les zones endémiques de l'insécurité alimentaire, essentiellement concentrée dans les zones tropicales et subtropicales (Ngamo et Hance, 2007 ; FAO *et al.*, 2013). Les pertes post-récoltes en grains en ASS se chiffrent à 4 milliards de dollars/an selon la FAO (2011a) ; elles affectent 14% à 50% du maïs (*Zea mays*, Linnaeus 1753) produit dans les pays en développement (Ojo et Omoloye, 2012). Dans ces zones, les stocks sont assaillis et détériorés par les insectes, et il est estimé à 35% de phytophages parmi les espèces connues de Coléoptères selon Montesinos (1998). Parmi ces insectes dépréciateurs, il y a les coléoptères, plus particulièrement les charançons dont 3 espèces sont inféodées aux céréales : *Sitophilus granarius* (Linnaeus, 1758), *Sitophilus oryzae* (Linnaeus, 1763) et *Sitophilus zeamais* (Motschulsky, 1855) caractérisées par leur fort potentiel de dépréciation des céréales stockées (Lefèvre, 2004 ; Anselme, 2006) ; et donc par leur importance agro-économique. En effet, ces charançons peuvent déprécier 25% voire 40% de la récolte maïsicole en 6 mois de stockage (Prates *et al.*, 1998 ; Papachristos et Stamopoulos, 2002 ; Tapondjou *et al.*, 2002). De ces 3 ravageurs du grain entreposé, *Sitophilus zeamais* est parmi les ravageurs majeurs du maïs dans les régions tropicales et subtropicales (Carvalho *et al.*, 2012 ; Paes *et al.*, 2012). Etant un insecte primaire, il ouvre la voie aux insectes secondaires comme *Tribolium confusum* (Du Val, 1868) mais aussi aux champignons qui secrètent, en plus, des substances nuisibles et cancérigènes telles que les aflatoxines. Cette situation compromet ainsi la sécurité alimentaire et le statut sanitaire des groupes les plus vulnérables en Afrique (Hell *et al.*, 2000 ; Sembène, 2000). En effet, le maïs constitue une denrée de base pour beaucoup de pays africains et leur permet de traverser, dans certaines contrées, les périodes de soudure (Guèye *et al.*, 2011). Ce qui fait de *Sitophilus zeamais* un problème agro-économique et sanitaire.

Pour faire face à ces ravageurs, plusieurs méthodes de lutte ont été mises en œuvre dont la plus utilisée est la lutte chimique. Cependant, la principale résultante de cette forme de lutte est la désorganisation des communautés biologiques qui se manifeste par un affaiblissement progressif de la productivité et de la rentabilité des agro-écosystèmes. A ceci s'ajoute une apparition et/ou une prolifération d'espèces, de populations ou d'écotypes d'insectes nuisibles résistants aux pesticides, une disparition des communautés d'auxiliaires. Les conséquences de toutes ces perturbations constituent une menace à la sécurité alimentaire à travers une infestation et une dépréciation massives des produits de conservations. Cette voie chimique est caractérisée notamment par son anarchisme dans les pays en voie de développement où les principaux acteurs n'ont pas une connaissance approfondie des produits utilisés et des conséquences éventuellement néfastes sur leur santé et l'environnement (Saxena *et al.*, 1990 ; Mbaye, 2014). La lutte intégrée constitue donc une alternative à la voie chimique car proposant des méthodes respectueuses de la santé des consommateurs et de la préservation de l'environnement.

Par ailleurs, une des préalables et étapes cruciales à cette lutte intégrée est la connaissance du

statut bioévolutif des populations incriminées (Sembène, 2000). Les études moléculaires constituent un outil plus déterminant dans la caractérisation des populations. Néanmoins, la grande majorité de ces études concernent des populations américaines et européennes, alors peu sont consacrées à l'entomofaune de l'Afrique subsaharienne ; à quelques exceptions telles *Caryedon serratus* et *Tribolium castaneum* (Sembène, 2000 ; Thiaw *et al.*, 2012 ; Ndiaye, 2014 ; Diome *et al.*, 2013), *Sitophilus* spp (Thiaw *et al.*, 2012 ; Ndiaye *et al.*, 2014 ; Ndong *et al.*, 2015), *Ceratitidis rosa* et *Ceratitidis fasciventris* (Baliraine *et al.*, 2004) ; *Eldana saccharina* (King *et al.*, 2002) ; *Busseola fusca*, (Sézonlin *et al.*, 2006 ; Félix, 2008) ; *Callosobruchus maculatus* (Kébé *et al.*, 2016 ; 2017), *etc...*

En outre, pour une meilleure compréhension de la dynamique d'évolution des populations naturelles africaines, une étude d'envergure sous-régionale (Afrique de l'Ouest) voire régionale (Afrique subsaharienne) aurait beaucoup plus d'impacts avec l'élaboration de stratégies de lutte pouvant être globales à l'échelle des objectifs mais différentielles et adaptatives dans leurs contenus à l'échelle locale. En effet, des insectes phytophages de la même espèce, à large distribution géographique et donc de populations, de plantes hôtes et de conditions écologiques différentes pourraient avoir des caractéristiques génétiques et de développement différentes induisant une adaptation de leurs populations et une variation du cycle biologique et conduire à la formation de races d'hôtes (biotypes et écotypes) voire de sous-espèces ou d'espèces cryptiques (Bekon, 1984 ; Seri-Kouassi *et al.*, 1987 ; Sézonlin, 2006). Chez la bruche de l'arachide, *Caryedon serratus* (Olivier, 1790), par exemple, il y a une structuration en biotype en fonction de la plante-hôte (Sembène *et al.*, 2012a).

En plus, la sous-région ouest-africaine est caractérisée par son dynamisme commercial et le développement des activités humaines a facilité l'apparition de nouvelles zones infestées de façon volontaire ou non. En effet, l'importance des échanges de marchandises qu'ont connue les produits agricoles ont conduit à l'importation de grains (semences) par les populations rurales d'un pays à un autre (Kébé, 2013 ; Ndiaye, 2014).

C'est dans cette perspective que se trouve l'intérêt de l'étude proposée en « Génétique des Populations » qui utilise les principes et techniques de la biologie moléculaire, de la bio-informatique, et des statistiques multivariées et bayésiennes. En effet, l'analyse moléculaire est basée sur la théorie neutraliste de l'évolution (Kimura, 1968 ; 1983a ; King et Jukes, 1969) selon laquelle la diversité génétique est le résultat de processus stochastiques (dérive génétique) car les mutations sont majoritairement ou presque neutres et sont fixées ou perdues aléatoirement sous l'action de la dérive génétique. Elle postule que le temps évolutif séparant différents taxa est reflété par leur divergence au sein des séquences d'ADN (Sembène, 2000). Donc, à travers l'analyse d'échantillons d'ADN provenant d'individus de différentes zones du système étudié, il y a une meilleure compréhension des processus (forces évolutives) qui régissent l'évolution et le fonctionnement des populations étudiées (dynamique et structure génétique) mais aussi l'intensité de leurs liens (flux génique). Dès lors, cette discipline permet de faire le lien entre cette diversité génétique et les données écologiques par une

analyse temporelle et spatiale.

Partant de ces observations, les hypothèses de recherche de cette étude sont :

- (i) les zones biogéographiques ou de provenance impactent sur l'évolution phylogéographique et la structure génétique des populations de *Sitophilus zeamais* ;
- (ii) et/ou l'action anthropique (cas des échanges commerciaux) influent sur la structure génétique des populations de *S. zeamais*.

L'objectif général de cette étude est donc de connaître les caractéristiques de l'évolution et de la structure génétiques des populations de *S. zeamais* afin de participer à l'élaboration des stratégies de lutte pour atténuer voire supprimer la dépréciation massive des denrées stockées pour préserver la valeur nutritive et marchande des céréales. Cette étude est axée sur 2 objectifs spécifiques distincts mais complémentaires destinés à mieux comprendre les facteurs qui déterminent les liens génétiques entre les populations (phylogénie), leur distribution géographique (phylogéographie) et la structuration des variants génétiques (génétique des populations, *sensu stricto*) de *Sitophilus zeamais* dans la sous-région ouest-africaine. Chaque objectif spécifique est organisé sous forme d'activités. Ces objectifs sont de :

- déterminer la diversité et la structure génétiques et définir les relations phylogéographiques entre populations par les séquences d'ADN ;
- estimer le flux génique et caractériser la structure génétique à l'aide des loci microsatellites (SSR).

Pour se faire, les travaux de cette thèse sont structurés en 3 chapitres : le premier est destiné à la synthèse bibliographique ; et les 2 derniers consacrés respectivement aux deux objectifs spécifiques précités dont chacun présentés selon une méthodologie bien définie suivie de la présentation des résultats obtenus suivi de leur discussion.

- Le chapitre I fait une présentation du complexe *Sitophilus*/maïs à travers les caractéristiques et l'importance agro-écologique de *S. zeamais*, mais aussi l'importance alimentaire et économique du maïs. Il présentera également le cadre agro-écologique de l'étude, les principales approches et méthodes utilisées en génétique des populations ;

- Le chapitre II est destiné à déterminer la structure génétique, les affinités phylogéniques et les relations phylogéographiques entre populations ouest-africaines par les séquences d'ADN ;

- Le chapitre III est consacré à la caractérisation de la structure génétique et à l'estimation du flux de gènes entre les différentes populations étudiées par les loci microsatellites.

Enfin, la conclusion générale est suivie par des recommandations et des perspectives.

**CHAPITRE I**  
**SYNTHESE BIBLIOGRAPHIQUE**

# CHAPITRE I : PRESENTATION DE MODELES D'ETUDE, CADRE AGRO-ÉCOLOGIQUE ET METHODOLOGIE EN GENETIQUE DES POPULATIONS

## INTRODUCTION

Ce chapitre est consacré au complexe « ravageur/plante-hôte ». Il a permis de mettre en exergue l'importance agro-économique des coléoptères en général et de *Sitophilus zeamais* en particulier mais aussi la valeur socio-économique, alimentaire et nutritionnelle du maïs et donc de l'agriculture dans le développement socio-économique en Afrique de l'Ouest dont le contexte agro-écologique est marqué par le dynamisme d'une intégration politico-économique. Enfin, vu que la génétique des populations constitue un outil d'investigation, nous allons faire un bref rappel des outils, techniques et postulat utilisés par cette discipline afin d'apporter des réponses aux questions scientifiques.

### I. 1 MODELES D'ETUDE : *SITOPHILUS ZEAMAI* (MOTSCHULSKY, 1855)/*ZEAMAYS* (LINNE, 1753)

#### I.1.1 GENERALITES SUR LES INSECTES COLEOPTERES PHYTOPHAGES

Les insectes constituent le groupe terrestre le plus diversifié (Daly *et al.*, 1998). Leur particularité est due à leur diversité interspécifique élevée, leur impact écologique et à leur 400 millions d'années d'histoire évolutive (Grimaldi et Engel, 2005 ; Yeates *et al.*, 2012). Cette particularité est aussi liée à leur importance économique due surtout à la compétition interspécifique pour la ressource végétale entre les insectes et l'homme d'une part et d'autre part à l'abondance de la classe des insectes représentant plus de 80% de la biodiversité totale décrite (Erwin, 1997). Parmi ces insectes, il y a les phytophages qui sont un groupe se nourrissant de tissus vivants provenant de plantes terrestres vascularisées appelées trachéophytes (Ptéridophytes, Prespermaphytes et Spermaphytes) (Strong *et al.*, 1984) ; et occasionnent alors des pertes pré et post-récoltes énormes. Cette phytophagie a constitué une barrière évolutive que seul les phytophages ont eu à franchir. Cette adaptation est ensuite suivie par une radiation adaptative (Strong *et al.*, 1984) qui est corrélée à la diversification des Angiospermes (Farrell, 1998 ; Marvaldi *et al.*, 2002). En effet, cette diversification de part et d'autre des 2 groupes est liée au concept de co-évolution d'Ehrlich et Raven (1964) ou « course aux armements » qui caractérise l'évolution des interactions plantes-insectes où les plantes vont développer au fur et à mesure des mécanismes de défenses (composés secondaires toxiques) et où les insectes vont élaborer des dispositifs (mécanismes de détoxification) pour contrer ces mécanismes. Ainsi, l'élaboration de chaque mécanisme au sein des 2 groupes est suivie de leur diversification. En outre, la spécialisation alimentaire est un trait caractéristique des insectes phytophages où plus de 75% des espèces sont strictement monophages et donc se nourrissent sur une seule espèce de plante donnée. Il y a aussi les oligophages qui s'alimentent sur un nombre limité d'espèces de plantes appartenant généralement à un ou plusieurs genres d'une même famille (Bernays et Chapman, 1994). Alors que les polyphages se caractérisent par un régime alimentaire basé sur des plantes appartenant à des familles botaniques différentes. Cette spécialisation alimentaire (selon la plante à consommer)

associée à la spécialisation trophique (selon les parties de la plante à consommer) ont joué un rôle décisif dans le succès évolutif des insectes phytophages (Marvaldi *et al.*, 2002). Parmi les phytophages, il y a principalement les Lépidoptères et les Coléoptères (Bernays et Chapman, 1994). L'ordre des coléoptères (*Coleoptera* ; Linnaeus, 1758 ; *Koleos* = étui, fourreau, *Ptera* = ailes) constitue le plus vaste de la classe des insectes : 40% des espèces d'insectes connues (Gullan et Cranston, 2005) ; 30% des espèces animales de la planète soit ~ 300 000 espèces décrites dans le monde soit le 1/3 de l'ordre des insectes (Bourbonnais, 2013). Selon Montesinos (1998), 35% des espèces connues de coléoptères sont phytophages. Il y a plus d'espèces de coléoptères connues que d'espèces de plantes. La présence d'élytres (ailes antérieures sclérifiées et donc rigides) est caractéristique de l'ordre. Ces élytres ne participent pas au vol mais protègent les ailes postérieures membraneuses aptes au vol. Ces élytres peuvent servir de défense avec d'autres organes, soit à l'aide de couleurs vives qui effraient leurs prédateurs (aposématisme), soit en adoptant les tons de l'habitat naturel de l'insecte (mimèse) ou bien en imitant un autre animal nocif (mimétisme). Les élytres sont parfois appelés *tegmina* (ou *tegmen* au singulier). Présents dans presque tous les milieux (sauf marins), la majorité des coléoptères sont herbivores ; mais on trouve aussi de nombreux prédateurs, des détritivores, des nécrophages et même quelques espèces parasites (Delvare et Aberlenc, 1989 ; Bourbonnais, 2013). Certaines espèces mangent les feuilles à l'âge adulte, comme le charançon vert ou le charançon noir, ce qui développe les maladies fongiques. Les espèces associées aux denrées entreposées peuvent causer, en plus des dommages directs, la dépréciation massive des cultures et denrées stockées. En général, les larves causent plus de dommages que les adultes en s'alimentant. Elles sont réparties en 2 groupes selon leur capacité à s'attaquer aux grains (Bekon et Fleurat-Lessard, 1989). En effet, il y a les ravageurs primaires qui s'attaquent aux grains intacts dont *Sitophilus zeamais* ou charançon du maïs et les ravageurs secondaires qui ne sont capables de s'attaquer qu'aux grains déjà rongés par les ravageurs primaires ; c'est le cas de *Tribolium confusum* (Du Val, 1868).

### **I.1.2 GENERALITES SUR LES « CHARANÇONS »**

Le terme « charançon » (coléoptères à museau) est un nom vernaculaire ambigu donné à plusieurs espèces d'insectes. Les charançons sont classés dans la plus grande des superfamilles de l'ordre des coléoptères, les *Curculionoidea* (Lefèvre, 2004 ; Anselme, 2006) composée de 10 familles (Bouchard *et al.*, 2011) ; ces insectes sont répartis principalement dans les familles *Curculionidae* et *Dryophthoridae*. Les 2 caractéristiques principales des charançons sont la présence d'une trompe proéminente (rostre) au bout de laquelle se situent les pièces buccales et les antennes coudées qui sont situées souvent vers le milieu du rostre. Les *Curculionoidea* constitue un groupe monophylétique soutenu par les données morphologiques et moléculaires dont l'apomorphie (caractère dérivé) la plus importante et visible est le rostre (Farrell, 1998 ; Marvaldi et Morrone, 2000). Une condition préalable pour le succès évolutif des charançons était l'utilisation de leur rostre dans l'oviposition. En plus,



l'adaptation clé des *Curculionidae*, *sensu stricto*, qui représentent la plupart des espèces de charançons, était probablement l'utilisation de tissus de plantes qui leur a permis d'exploiter les angiospermes au moment où ils sont apparus et leur expansion avec ce groupe de plantes (Anderson, 1995). En effet, la diversité de l'ensemble des phytophages s'explique par une co-évolution précoce des clades les plus riches en espèces des *phytophaga* (*Curculionoidea* et *Chrysomeloidea*) avec les angiospermes (Ehrlich et Raven, 1964 ; Farrell, 1998 ; Marvaldi *et al.*, 2002).

Leur reproduction est directe et sexuée, mais chez certaines espèces, elle peut être parthénogénétique. Le nombre d'œufs pondus par les femelles est d'importance très variable (quelques-uns à plusieurs centaines). La plupart des larves sont endophages, mais certaines sont ectophages et vivent comme les chenilles, en se déplaçant sur les parties aériennes des plantes. D'autres sont terricoles et consomment le pourtour des racines sans y creuser de galeries, au contraire de la majorité des espèces radicicoles (vivant sur les racines). La métamorphose est complète (insectes holométaboles). Le cycle vital des coléoptères comporte donc quatre étapes : l'œuf, la larve, la nymphe et l'adulte. La larve, éruciforme et apode, va subir plusieurs transformations avant de se nymphoser (Anonyme, 2009). La vie d'un charançon a une durée qui varie selon l'espèce.

### **I.1.3 PRESENTATION DU « CHARANÇON » : *SITOPHILUS ZEAMAI* (MOTSCHULSKY, 1855)**

La taxonomie du genre *Sitophilus* a été confuse jusqu'à récemment, de sorte que la valeur d'une grande partie de la littérature antérieure sur ces insectes a été réduite, en raison de la difficulté de connaître l'espèce à laquelle elle se réfère (CABI, 2017 ; 2018). Récemment 14 espèces ont été dénombrées dans le genre *Sitophilus* (Delobel et Grener, 1993 ; Corrêa *et al.*, 2012). Parmi ces espèces, il y en a 3 espèces qui forment un groupe monophylétique, à savoir le charançon du blé, *S. granarius*, le charançon du riz, *S. oryzae* et le charançon du maïs, *S. zeamais* (Grenier *et al.*, 1986 ; Rees, 1996 ; Charles, 1997 ; Lefèvre, 2004 ; Plarre, 2010 ; Corrêa *et al.*, 2012). Cette monorphie est attestée, entre autres, par le succès de transférabilité des amorces de marqueurs de microsatellites entre les 3 espèces (Corrêa *et al.*, 2012). Ils jouent un rôle d'importance agro-économique comme insectes ravageurs de grains de céréales stockés. En effet, ils sont responsables des pertes de denrées stockées qui peut atteindre 40% dans les pays tropicaux (Grenier *et al.*, 1986). A ces 3 espèces s'ajoute le charançon du tamarin, *Sitophilus linearis* (Herbst, 1797) (Silva *et al.*, 2015). Sur le plan trophique, la larve de ces insectes phytophages est qualifiée de cléthrophage car vivant à l'intérieur des grains (Rees, 2004 ; Beckett *et al.*, 2007), par opposition à l'adulte qualifié de granivore puisqu'il les ronge extérieurement. En effet, toutes les espèces du genre *Sitophilus* se développent, de l'œuf à la puppe, à l'intérieur du noyau de leurs plantes hôtes (Plarre, 2010). Suivant l'aspect évolutif, *Sitophilus* spp sont qualifiées de sitiobies : insectes incapables d'effectuer leur évolution complète en dehors des denrées.

Ce genre appartient à la famille des *Curculionidae* (Latreille, 1802) qui est la plus grande famille de l'ordre des coléoptères et la famille zoologique qui contient le plus grand nombre d'espèces

(Anonyme, 2009). Il faut noter que *S. zeamais* et *S. oryzae* se présentent comme des espèces jumelles et présentent donc des similarités dans leur morphologie, leur biologie, leur génétique, leur répartition, *etc...* *S. zeamais* se retrouve essentiellement dans le blé, le riz, le maïs, le seigle, le sorgho, le tamarin, les produits dérivés de céréales comme les pattes (Tipping *et al.*, 1987 ; Cruz *et al.*, 1988 ; Trematerra, 2009 et 2013). Leur infestation se caractérise par une augmentation des niveaux d'humidité et le chauffage à la surface et aussi par la présence de trous ronds sur les grains dus à la sortie des adultes (Anonyme, 2013a ; 2013b). L'infestation initiale des grains de maïs par *S. zeamais* débute dans les champs où l'humidité est encore de 50-55% (Hill et Waller, 1990) juste avant la récolte mais l'importance des pertes se passe dans les infrastructures de stockages où la population s'accumule rapidement (Likhayo et Hodges, 2000 ; Adedire, 2001 ; Adedire et Lajide, 2003 ; Iloba et Ekrakene, 2006). D'après De Lima (1979), 6 à 8 semaines plus tard après la récolte de maïs, le charançon a déjà terminé une génération et a pondu ses œufs pour une deuxième génération. Alors que pour *S. oryzae*, une infestation aux champs des grains de maïs est rarement observée à cause de la compétition interspécifique plus favorable à *S. zeamais* qu'à *S. oryzae* sur le maïs (Vásquez-castro *et al.*, 2009).

### I.1.3.1 POSITION SYSTEMATIQUE

D'après Bouchard *et al.* (2011) *Sitophilus zeamais* (Motschulsky, 1855) est un ravageur primaire classé dans les catégories taxinomiques principales comme suit :

- **Règne** : *Animalia* (animaux), Linnaeus, 1758 ;
- **Embrassement** : *Arthropoda* (arthropodes), Latreille, 1829 ;
- **Sous-embranchement** : *Hexapoda* (hexapodes) Blainville, 1816 ;
- **Classe** : *Insecta* (insectes), Linnaeus, 1758 ;
- **Sous-classe** : *Pterygota* (ptérygotes), Lang, 1888 ;
- **Super-Ordre** : *Holometabola*, Sharp, 1898 (*Endopterygota*) ;
- **Ordre** : *Coleoptera* (coléoptères), Linnaeus, 1758 ;
- **Sous-ordre** : *Polyphaga*, Emery, 1886 ;
- **Super-famille** : *Curculionoidea*, Latreille, 1802 ;
- **Famille** : *Curculionidae*, Latreille, 1802 ;
- **Sous-famille** : *Curculioninae*, Latreille, 1802 ;
- **Genre** : *Sitophilus*, Schönherr, 1838.

Notons que *S. zeamais* est classée selon certains auteurs dans la famille des *Dryophthoridae* (Schönherr, 1825) et la sous-famille des *Rhynchophorinae* (Schönherr, 1833) d'après la classification d'Alonso-Zarazaga et Lyal (1999).

### I.1.3.2 ORIGINE ET RÉPARTITION GÉOGRAPHIQUE

L'origine du charançon du maïs reste inconnue de manière certaine (Akol *et al.*, 2011) comme pour un certain nombre d'insectes notamment les espèces à distribution cosmopolite particulièrement celles des denrées stockées (Sauvion *et al.*, 2013). Elles sont ainsi nommées « espèces cryptogéniques » (Carlton, 1996) car on ne sait si elles sont indigènes ou introduites par rapport à une aire géographique donnée. En effet, leur distribution originale et naturelle a disparu après des siècles de transferts. Cependant, le charançon du maïs est supposé être originaire d'Asie ou d'Eurasie (Delobel et Grenier 1993 ; Plarre, 2010). Selon Plarre (2010), les espèces du genre *Sitophilus*, particulièrement les 3 espèces d'importance économique, *S. granarius*, *S. oryzae* et *S. zeamais*, sont considérées comme endémiques dans les zones forestières d'Asie. Ces auteurs suggèrent que les premières infestations des grains entreposés ont eu lieu près des montagnes boisées du sud-ouest asiatique ; et que pendant ce stockage, une adaptation par une co-évolution aurait eu lieu, ce qui a avantage ces 3 espèces par rapport aux autres charançons dans les grains stockés. Cependant, l'étude d'Obata *et al.* (2011), grâce à des recherches archéologiques, a mis en évidence des impressions de charançons du maïs sur la poterie jomon (Japon), âgées de 10.500 BP par la méthode du Carbone 14. En effet de nombreuses études ont montré que les charançons sont fréquemment trouvés dans les sites archéologiques et fournissent des informations importantes sur l'histoire des environnements urbains de l'humanité, les origines des grains et de même que les routes de dispersion et d'échanges ; en plus, de l'histoire des infrastructures de stockage (Levinson et Levinson, 1994 ; Kenway et Carrott 2006 ; Smith et Kenward, 2011). Notons que, selon Metcalf et Flint (1962), *Sitophilus oryzae* serait originaire de l'Inde et répandue dans le monde entier, à travers le transport des grains infestés, par les navires. Par ailleurs, Zimmerman (1993) met en évidence que la répartition actuelle des *Dryophthoridae* laisse présager que le foyer de diversification de la famille se situe en Afrique et dans la région indo-pacifique ; et que l'Amérique tropicale serait un foyer secondaire.

Le charançon du maïs est une espèce cosmopolite et plus particulièrement à répartition tropicale et subtropicale. Il se rencontre le plus souvent en association avec le charançon du riz et spécialement dans les zones maïsicoles (Longstaff, 1981). Le charançon de maïs montre de grandes possibilités de dispersion naturelle accentuée par l'intermédiaire des activités anthropiques (Corrêa *et al.*, 2012). En effet, la répartition mondiale de ces insectes a eu lieu récemment, et ils ont continué leur propagation jusqu'aux régions tempérées. Ceci constitue le résultat des échanges de céréales et de matières premières par le commerce dans le monde entier au cours du 20<sup>e</sup> siècle (Obata *et al.*, 2011 ; Anonyme, 2013a ; 2013b).

### I.1.3.3 IMPORTANCE AGRO-ECONOMIQUE ET SANITAIRE

L'importance agro-économique de *Sitophilus zeamais* se trouve dans son potentiel invasif et dépréciateur des céréales stockées. En effet, la détérioration des céréales affecte les valeurs nutritives

(qualité et quantité), commerciales et agronomiques avec une perte de poids de plus de 30% provoquée par l'infestation (Bell *et al.*, 1998 ; Hell *et al.*, 2008). En effet, en zone rurale où les techniques de conservation des denrées agricoles sont peu élaborées, *S. zeamais* peut occasionner des pertes post-récoltes pouvant atteindre 90% pendant 5 mois de stockage (Nukenine *et al.*, 2002 ; Boura, 2006 ; Noosidum et Sangprajan, 2014). Ce ravageur constitue donc un des obstacles majeurs à l'atteinte de la sécurité alimentaire dans les pays en développement car il provoque des pertes post-récoltes énormes. Étant un insecte primaire, il ouvre la voie aux insectes secondaires mais aussi aux champignons. En effet, la détérioration de la qualité du grain est aussi exacerbée par la prolifération de champignons et de bactéries d'où la production de substances nuisibles telles les aflatoxines. Les aflatoxines sont un groupe de mycotoxines produites principalement par *Aspergillus flavus* (Link, 1809) et *Aspergillus parasiticus* (Speare, 1912) ; rarement par *Aspergillus nomius* (Kurtzman, Horn et Hesselt, 1987). Ce qui contamine souvent l'alimentation humaine et animale, provoquant une maladie et/ou la mort des consommateurs (Hell *et al.*, 2008). Ces aflatoxines compromettent la sécurité alimentaire et le statut sanitaire des groupes les plus vulnérables de personnes en Afrique car ils sont cancérogènes (Hell *et al.*, 2000 ; Sembène, 2000 ; Strosnider *et al.*, 2006 ; Hell *et al.*, 2008, 2010 ; Magnussen *et al.*, 2013 ; Villers, 2014). En effet, une enquête menée par Williams (2011) révèle qu'environ 40% de produits sur les marchés locaux africains dépassent les niveaux admissibles d'aflatoxines dans les aliments. Certaines des expositions humaines les plus élevées et les plus persistantes à l'aflatoxine se produisent en Afrique de l'Ouest où presque 99% des enfants étaient positifs pour un test à l'aflatoxine (Gong *et al.*, 2002 ; 2004) ; et d'après Egal *et al.* (2005), la consommation de maïs est une source importante d'exposition à aflatoxine pour ces enfants. Par conséquent, *S. zeamais* constitue une espèce cosmopolite, d'importance agro-économique et sanitaire.

#### **I.1.3.4 CARACTÉRISTIQUES MORPHOLOGIQUES, PHÉNOTYPIQUES ET GÉNÉTIQUES**

*Sitophilus zeamais* a des caractéristiques morphologiques et biologiques similaires à *Sitophilus oryzae* (Gallo *et al.*, 2002). Chez *S. zeamais*, la larve blanche à jaune-clair est charnue et apode avec une tête plus sombre ; et elle mesure entre 2 à 3 mm. Les chrysalides montrent une couleur blanche laiteuse. Les adultes sont de petits insectes avec 3-4 mm de long (Figure 1A), de couleur brun foncé, avec quatre taches jaune à rougeâtres sur les élytres (Alleoni et Ferreira, 2006). Étant des espèces jumelles, leur distinction morphologique est difficile (Hidayat *et al.*, 1996), cependant des critères de distinction subsistent tels que :

- sur le plan morphologique, selon Yasutomi *et al.* (2000), les 2 espèces ont été considérées comme distinctes sur la base de la taille du corps, le charançon du maïs est légèrement plus grand que le charançon du riz ; en effet, il mesure en moyenne 2,8 mm contre 2,3 mm ;
- des préférences alimentaires, *S. zeamais* préfère les grains de grande taille comme le maïs alors

que *S. oryzae* a une préférence pour les grains de petite taille comme le riz, le blé (Seck, 2009 ; Obata *et al.*, 2011) ;

- l'incompatibilité pour l'accouplement/reproduction, même si *S. oryzae* et *S. zeamais* présentent le même caryotype  $2N=22$  (Yang *et al.*, 1989 ; Silva *et al.*, 2015) elles sont génétiquement bien distinguées comme 2 espèces différentes.
- la morphologie du pronotum et des organes génitaux ; en effet, les études de Boudreaux (1969) puis Hidayat *et al.* (1996) mettent en évidence la présence chez *S. oryzae* et l'absence chez *S. zeamais* d'une zone exemptée de ponctions sur la ligne médiane. En plus, chez *S. oryzae*, les ponctions sont ovales alors qu'elles sont rondes chez *S. zeamais* (Figure 1B). Le décompte du nombre de ponctions le long de la ligne médiane du pronotum est inférieur à 20 pour *S. oryzae* et supérieur à 20 pour *S. zeamais*. Les tergites abdominaux sont généralement noirâtres chez *S. zeamais* et rougeâtres pour *S. oryzae* (Zimmerman, 1968).
- l'activité de vol selon Vásquez-castro *et al.* (2009) : *Sitophilus zeamais* a une plus grande activité de vol que *S. oryzae* ; et donc une capacité de dispersion naturelle plus grande. Il y a aussi un effet significatif du type de céréales, en général, les insectes élevés sur le maïs ont une plus grande activité de vol que ceux élevés sur le blé. En effet, l'activité de vol de *Sitophilus zeamais* est de 35% sur le maïs contre 18% sur le blé.
- l'existence d'un dimorphisme sexuel basé sur le caractère du rostre qui est équipé à l'extrémité de pièces buccales de type broyeur. En effet, chez le mâle, le rostre est plus court, plus épais et rugueux sur la surface alors que chez la femelle, il est effilé et lisse (Halstead, 1963a).

Par ailleurs, pour les caractéristiques de l'édéage-genitalia, certains auteurs ont conclu que seul l'édéage du mâle et le sclérite en forme d'Y (formé par le spiculum ventrale et les demi-sternites fusionnés) de la femelle ont donné des séparations consistantes (Kuschel, 1961 ; Halstead, 1963b ; Kuschel, 1978 ; Gallo *et al.*, 1988 ; Hidayat *et al.*, 1996). En effet :

- chez le mâle : il y a la présence, chez *S. zeamais*, de crêtes longitudinales sur le dos de l'édéage du mâle (phallus) alors qu'ils sont absents chez *S. oryzae* où la surface est lisse ;
- chez la femelle : Il y a un sclérite en forme d'Y du 8<sup>e</sup> sternite invaginé (spiculum ventral) de la femelle qui est censé être arrondi et épais chez *S. oryzae* alors qu'il est pointu et mince chez *S. zeamais*.

D'autres auteurs proposent des critères de distinction moléculaire. En effet, d'après Yingjie *et al.* (1999), il y a 3 bandes protéiques évidentes caractéristiques pour *S. oryzae* et une bande pour *S. zeamais* avec une valeur de pH égale à 9,1 et un gel de concentration à 7,5% sur électrophorèse. Par ailleurs, d'après des analyses métaphasiques, *S. zeamais* possède  $2n = 22$  chromosomes avec 10 paires d'autosomes et 1 paire de gonosomes ( $2n = 20 + Xy$ ) (Yang *et al.*, 1989). Cependant, chez différentes populations, ce nombre peut varier de 22 à 26 paires de chromosomes dû à la présence de 1 à 4 chromosomes surnuméraires (Moraes *et al.*, 2003 ; Silva *et al.*, 2015). La présence de ces

chromosomes surnuméraire chez les espèces de la famille des *Curculionidae* est néanmoins rare (Silva *et al.*, 2015). En outre, il faut noter qu'il y a 11 paires de chromosomes chez *Sitophilus oryzae* sous le meioformula ( $2n = 20 + Xy$ ) (Smith et Virkki, 1978 ; Yang *et al.*, 1989 ; Moraes *et al.*, 2003) et 12 paires de chromosomes ( $5 + XX$ ) (Inkman, 1933 *in* Smith et Virkki, 1978) chez *Sitophilus granarius*.



A) Vue dorsal

B) Aspects différentiels du pronothum entre *S. oryzae* (gauche) et *S. zeamais* (droite)

**Figure 1** : Morphologie de *Sitophilus zeamais*

#### Sources :

(A) : <https://www.cabi.org/isc/datasheet/10926>

(B) : Hong *et al.*, 2018

#### I.1.3.5 BIOLOGIE DE LA REPRODUCTION ET CYCLE DE DÉVELOPPEMENT

Le cycle de développement complet des charançons, qui sont des espèces holométaboles et polyvoltines, est de 26 à 35 jours dans les zones tropicales et sub-tropicales (Haryadi et Fleurat-Lessard, 1994 ; Seck, 2009 ; Gallo *et al.*, 2002) ; et enfin, il est de 31 jours environ à 27,5°C pour la souche symbiotique (Charles, 1997 ; Lefèvre, 2004). Ce cycle peut atteindre 110 jours dans des conditions très défavorables (Cornelius *et al.*, 2008). Le temps de développement optimal est presque identique à celui de *S. oryzae* et est de 6 jours environ plus rapide que chez *S. granarius* dans les conditions semblables (Anonyme, 2016). Cependant, il faut noter que des insectes, appartenant à la même espèce mais provenant de régions différentes, peuvent présenter un cycle de développement différent (Bekon, 1984). En effet, selon Seri-Kouassi *et al.* (1987), la diversité des conditions écologiques entraînent l'adaptation des populations tandis que leurs cycles subissent des variations. Les données écologiques influenceraient donc les paramètres biologiques des populations.

Chez *S. zeamais*, la maturité sexuelle est acquise dès le jour même où l'insecte sort du grain. Les femelles se déplacent sur la nourriture pour libérer l'hormone sexuelle, la phéromone ; et les mâles sont alors attirés par cette phéromone pour l'accouplement (Mason, 2003). La ponte a lieu à partir du 3<sup>ème</sup> jour et se fait, très souvent, au voisinage du sillon central du grain près du germe où, grâce à son rostre, la femelle pratique un trou et y dépose directement un œuf (Kehe, 1975 ; Thrône,

1994). Après la ponte, le trou est recouvert d'une sécrétion mucilagineuse de l'ovoducte qui durcit à l'air (Nardon, 1963 ; Plarre, 2010). Il y a des cas où plus d'un œuf est déposé dans un seul grain ; cependant à cause du cannibalisme, il est rare que plus d'une larve se développe à l'échéance (Longstaff, 1981). Le grain de céréale apporte donc à l'insecte à la fois sa nourriture et sa protection. Le comportement de ponte des femelles joue un rôle important car il détermine le nombre et la qualité de la progéniture (Honek, 1993 ; Stejskal et Kucerova, 1996). En effet, il varie selon l'espèce et/ou la souche, la densité de la population, les conditions environnementales, la nourriture, l'âge et la taille des individus (Stejskal et Kucerova, 1996 ; Danho *et al.*, 2002). La femelle s'attaque aux grains intacts d'où leur avilissement ; en effet les larves et les adultes se nourrissent et se développent à l'intérieur du noyau (Lopez *et al.*, 2008 ; Ojo et Omoloye 2012). De ce fait, après l'émergence de l'adulte, le grain est complètement creux (Plarre, 2010).

Le développement larvaire (3 à 4 stades larvaires et un stade nymphal) jusqu'à l'âge adulte se déroule à l'intérieur du grain. L'incubation des œufs va de 3 à 6 jours (Gallo *et al.*, 2002). En général, les larves du genre *Sitophilus* se nourrissent préférentiellement d'endosperme ou de germe, réduisant alors les valeurs nutritives des protéines et vitamines (Dal Bello *et al.*, 2001). La limite inférieure concernant la température pour le développement de l'œuf à l'adulte des charançons est de 15,6°C et la limite supérieure est de 32,5°C à 75% d'humidité relative (Thrône, 1994). Après la mue imaginale, l'adulte reste un à deux jours dans le grain avant l'émergence selon la température (Longstaff, 1981 ; Charles, 1997). Pendant ce temps, sa cuticule se durcit et mûrit et enfin l'adulte émerge. Les adultes peuvent vivre 3 à 6 mois ou plus (Alleoni et Ferreira, 2006 ; Anonyme, 2013b). Par ailleurs, les charançons sont caractérisés par un chevauchement de générations et un comportement grégaire ; ce qui favorise la consanguinité (Corrêa *et al.*, 2012).

### I.1.3.6 PARTICULARITES SYMBIOTIQUES DES ENDOBACTERIES

Chez les insectes, la survie de 10 à 20% des espèces dépendent de la symbiose avec des bactéries (Buchner, 1965 ; Douglas, 1989 ; Wernegreen, 2002). Chez le genre *Sitophilus*, cette symbiose a été découverte en 1927 par Pierantoni. A l'exception d'une espèce, *Sitophilus linearis* (Herbst, 1795), les charançons du genre *Sitophilus* vivent en symbiose intracellulaire avec une bactérie,  $\gamma$ -protéobactérie, (Charles, 1997 ; Lefèvre, 2004 ; Anselme, 2006), transmise à la descendance par voie ovocytaire qui, d'après Heddi *et al.* (1998), se nomme SPE pour « *Sitophilus* Primary Endosymbiont ». Plus spécifiquement, elle est nommée SZPE (*Sitophilus zeamais* Primary Endosymbiont) chez *S. zeamais* ; SOPE (*Sitophilus oryzae* Primary Endosymbiont) chez *S. oryzae* ; et SGPE (*Sitophilus granarius* Primary Endosymbiont) chez *S. granarius* (Heddi *et al.*, 1999 ; 2001). Cette bactérie est qualifiée d'endocytobiotte primaire obligatoire (Heddi *et al.*, 1999) et logée dans un organe appelé bactériome (Buchner, 1965). Cette transmission est caractérisée par une réduction de la taille du génome au cours du temps (Pérez-Brocal *et al.*, 2006), un enrichissement global en bases

nucléiques A-T (Heddi *et al.*, 1998) et une rapidité du taux de substitution (Moran, 1996). Dans cette symbiose, la bactérie produit des nutriments tels que les acides aminés et vitamines pour l'hôte, complétant dès lors son régime alimentaire de céréales (Vallier *et al.*, 2009). Alors que l'insecte assure à la bactérie habitat, source de nourriture et protection. Cette bactérie est Gram négatif proche de la famille des entérobactéries. Les trois charançons *S. granarius*, *S. oryzae* et *S. zeamais* sont étroitement liés sur la base de l'analyse de l'ADN de leurs endosymbiotes (Obata *et al.*, 2011).

L'importance de leur présence est liée à la physiologie en général et particulièrement au système reproductif, immunologique et nutritif (Grenier *et al.*, 1986, 1994 ; Charles, 1997 ; Heddi *et al.*, 2005 ; Heddi, 2009). En effet, chez *S. oryzae*, l'élimination des endocytobiotés, par exemple, entraîne une augmentation de la durée de développement, une baisse de la fertilité et la perte de l'aptitude au vol chez l'adulte (Grenier *et al.*, 1986 ; 1994). Chez *Sitophilus zeamais* (Anselme *et al.*, 2006), l'endosymbiote primaire peut activer les protéines responsables de la détection de peptidoglycanes dans certaines phases du développement de l'insecte, ce qui peut entraîner des conséquences sur son immunité. Il y a aussi des endosymbiotes qualifiés de secondaires comme le *Wolbachia* impliqué dans le système reproductif des hôtes et qui est une bactérie endosymbiote obligatoire appelée  $\alpha$ -*Proteobacteria* (Stouthamer *et al.*, 1999). Ces endocytobiotés peuvent être éliminés par des conditions thermiques à la chaleur de 35°C et sous une humidité relative de 90% (Nardon, 1973) ; ces souches aposymbiotiques constituent des modèles d'étude au laboratoire car étant vivantes et fertiles. La comparaison d'individus aposymbiotiques et symbiotiques a permis de mettre en évidence sur de nombreux modèles les apports métaboliques des endocytobiotés (Lefèvre, 2004). Il devient donc évident que le fitness et le potentiel destructeur de ces dépréciateurs de denrées stockées est lié à cette symbiose (Nardon et Charles, 2001 ; Heddi *et al.*, 1999). A cela s'ajoute le rôle probable de ces endocytobiotés dans l'élargissement de la gamme de plantes alimentaires (Tsuchida *et al.*, 2004), leur résistance aux températures extrêmes (Montllor *et al.*, 2002) et la protection contre les parasites (Oliver *et al.*, 2003). Par conséquent, la compréhension de la relation endocytobiotés-hôte, dans toute sa diversité, constituerait un défi à relever pour mettre en place des stratégies de lutte efficaces respectueuses de la nature et des consommateurs.

### I.1.3.7 PARASITOÏDES DU GENRE *SITOPHILUS*

Un insecte est défini comme un parasitoïde lorsqu'il présente les caractéristique suivantes selon Douth (1959) :

- il se développe aux dépens d'un seul hôte (insectes essentiellement, Boivin, 1996) et entraîne la mort de ce dernier ;
- en général, l'hôte appartient à la même classe taxonomique ;
- le parasitoïde et l'hôte présentent le plus souvent des tailles similaires ;
- le parasitisme concerne seulement une phase de leur existence ; le plus souvent, en partie ou en



totalité, la phase larvaire ; alors que la phase imaginaire est généralement libre ;

- le parasitoïde en développement ne peut pas changer d'hôte ;
- et enfin, l'impact du parasitoïde sur la dynamique des populations de l'hôte est plus proche de celle d'un prédateur que de celle d'un véritable parasite.

Dans le cadre de la gestion des populations naturelles à travers la lutte intégrée, les insectes parasitoïdes de l'ordre des *Hymenoptera* et de la famille des *Pteromalidae* sont utilisés comme agents biologiques contre les insectes ravageurs des denrées stockées. Des recherches ont montré que des parasitoïdes tels qu'*Anisopteromalus calandrae* (Howard, 1881) et *Lariophagus distinguendus* (Forster, 1841) pourront être utilisés comme agents de lutte biologique contre *S. oryzae* et *S. zeamais* (Smith, 1993 ; Wen et Brower, 1995 ; Helbig, 1998 ; Pérez-Mendoza *et al.*, 1999 ; Lucas et Riudavets, 2002 ; Chaisaeng *et al.*, 2010). Ces ennemies naturelles des charançons s'attaquent à leurs larves et nymphes. Sur le riz non poli, Lucas et Riudavets (2002) ont mis en évidence l'action prépondérante de *L. distinguendus* par rapport à *A. calandrae* par une réduction des populations de charançons de 98%. Alors que Smith (1993) a observé une préférence de *L. distinguendus* pour les grosses larves-hôtes avec 87% de parasitisme. Selon Chaisaeng *et al.* (2010) et Arbogast et Mullen (1990), le parasitoïde *A. calandrae* peut être un agent de lutte biologique efficace, s'il est introduit en quantité suffisante au début de la période de stockage de manière à supprimer l'accroissement initial des populations de charançon de maïs.

### I.1.3.8 PRESENTATION SYNOPTIQUE DES ETUDES SUR *SITOPHILUS* SPP

En Afrique, en général, la plupart des études très nombreuses sur ce genre concerne les méthodes de lutte avec les produits bio-insecticides (huiles essentielles, formules poudreuses ou liquides) issus de plantes dont la plupart sont utilisées par les agriculteurs. La propagation de ces méthodes est liée à leur nombreux avantages par rapport aux pesticides de synthèse, tels que la haute sélectivité, la faible ou non-toxicité pour les organismes non-ciblés et sur de l'environnement, la dégradation rapide, la faiblesse des résidu, la disponibilité locale, peu de résistance croisée en raison de leurs agents naturels complexes et de leurs modes d'action inédit contre les insectes (Isman 2006 ; 2008 ; KuDom *et al.*, 2011 ; Lv *et al.*, 2012 ; Ladhari *et al.*, 2013). La base de ces études est une connaissance de la biologie de *Sitophilus*, qui est une étape préalable à l'élaboration des méthodes de contrôle efficaces (Zaghloul *et al.*, 2012). Le but de cette étape est la détermination des mécanismes de réponse aux effets des facteurs environnementaux et biologiques. Alors qu'au niveau moléculaire, les études sont peu nombreuses ; il y a par exemple Ndiaye *et al.* (2014), Ndong *et al.* (2015), Sarr *et al.* (2016) en Afrique de l'Ouest et Centrale. En revanche, en Amérique du Nord et en Europe, de nombreuses études se sont focalisés sur les organismes symbiotiques de *Sitophilus*, les endocytobiotés (Heddi *et al.*, 1998, 1999, 2001 ; Rio *et al.*, 2003 ; Lefèvre, 2004 ; Anselme, 2006). Il y a aussi les études basées sur les marqueurs moléculaires comme Cao (2007). C'est aussi le cas de

Corrêa *et al.* (2012) qui ont réalisé la première étude sur la détermination, la caractérisation et la transférabilité de loci microsatellites de *S. zeamais*, *S. oryzae* et *S. granarius* regroupant des populations du Brésil, des Etats-Unis et de l'Afrique centrale. La même dynamique a été poursuivie par Thangaraj *et al.* (2016) avec *S. oryzae* avec des populations issues de la Chine, de l'Australie et des Etats-Unis.

## **I.2 PRESENTATION DE LA PLANTE-HOTE : MAÏS (*ZEAMAYS* L., 1753)**

Originaire de l'Amérique Centrale (Le Conte, 1950 ; FAO, 1993), le maïs, *Zea mays* ssp, s'est adapté avec succès à de nombreuses régions du monde (tropicales et tempérées) de par sa grande variabilité morphologique et physiologique. Il serait introduit en Afrique par les explorateurs portugais. C'est une plante tropicale, monoïque (la même plante porte, à différents endroits, les fleurs mâles et femelles), allogame (la fécondation est majoritairement croisée, hybridation naturelle, les fleurs femelles sont fécondées par le pollen d'une autre plante) et une herbacée annuelle voire biannuelle de la famille des graminées (Poacées). Le cycle du maïs se décompose en trois phases de développement bien distinctes, définies par la formation d'un ou de plusieurs organes essentiels de la plante : la phase végétative, la phase de reproduction, la phase du développement du grain.

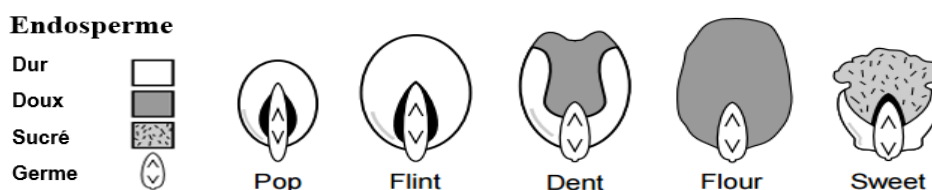
La fixation du carbone chez le maïs est effectuée par le système photosynthétique en C4 qui confère au maïs une plus grande efficacité pour fixer le CO<sub>2</sub> atmosphérique comparé aux autres céréales principales. Plus de cinq mille variétés de maïs ont été dénombrées à travers le monde. Ces variétés sont regroupées en sept types selon les particularités de l'albumen des caryopses et la morphologie de l'épi femelle. Selon la classification de Cronquist (1981), l'espèce *Zea mays* (L., 1753) est classée dans :

- **le règne :** *Plantae*
- **le sous-règne :** *Tracheobionta*
- **la Classe :** *Liliopsida*
- **la Sous-classe :** *Commelinidae*
- **l'Ordre :** *Cyperales* ou *Poales*
- **la Famille :** *Poaceae*
- **la Sous-famille :** *Panicoideae*
- **le Genre :** *Zea*

Le maïs est classifié en 6 voire 7 variétés principales (Singh *et al.*, 2009) selon des caractéristiques de l'endosperme (Figure 2), la couleur du grain, la maturité et l'utilisation finale (Paliwal *et al.*, 2000), selon le type de noyau (Dickerson, 2003) mais aussi selon les particularités de l'albumen des caryopses et la morphologie de l'épi femelle pour le commerce et la consommation humaine :

- maïs indenté ou dent corn (*Zea mays indenata*), (grains très petits et pointus, à albumen presque entièrement constitué d'amidon vitreux) ;
- maïs à albumen corné ou flint corn : (*Zea mays indurata*) ;
- maïs amidacé ou flour/soft corn : (*Zea mays amylacea*) ;
- maïs sucré ou sweet corn : (*Zea saccharata or Zea rugosa*) ;
- maïs visqueux ou Waxy corn (*Zea mays cerotina*) ;
- maïs fulminant ou Pop-corn (*Zea mays everta*) ;
- maïs tunique ou Pod corn (*Zea mays tunicata*).

Chacune de ces formes a déjà reçu le statut d'espèce ou de sous-espèce, mais les analyses génétiques ont démontré que ce statut n'était pas justifié car un seul gène, plus l'influence de quelques gènes régulateurs-modificateurs, est responsable de l'expression des caractères qui différencient ces types de maïs. A l'exception de *Zea perennis* qui est tétraploïde,  $2n = 40$ , toutes les autres espèces et sous-espèces du genre *Zea* ont le même nombre de chromosomes ( $2n = 20$ ).



**Figure 2 :** La distribution de l'endosperme dans les cinq types de grains de maïs ; Dickerson, 2003

Le maïs est la 1<sup>ère</sup> en termes de production et la 3<sup>ème</sup> céréale mondiale en termes de surfaces cultivées après le blé et le riz (Anzala, 2006 ; Larsen, 2012) de par sa large utilisation dans l'alimentation animale et humaine ainsi que pour des usages industriels. Il constitue une denrée de base pour beaucoup de pays africains et se place comme un produit de substitution pour le millet et le sorgho et permet de traverser les périodes de soudure dans certains contrées sahéliennes (Ndiaye, 2005 ; Guèye *et al.*, 2011). On estime que 200-300 millions de personnes en Afrique sub-saharienne (ASS) dépendent du maïs comme source de sécurité alimentaire et de bien-être économique (Macauley et Ramadjita, 2015). Parmi les 20 pays d'ASS qui représentent 96 % de la production totale de maïs, 7 pays (Nigeria, Ghana, Mali, Burkina Faso, Benin, Togo et Côte d'Ivoire) sont d'Afrique de l'Ouest (Ibid). La production de maïs dans l'Afrique subsaharienne a triplé depuis le début des années 1960 à la fin des années 1990 (Hell *et al.*, 2008). Le Nigéria se présente comme 1<sup>er</sup> producteur de maïs et représente l'économie agricole la plus puissante de la sous-région ouest-africaine avec plus de la moitié de la population sous-régionale (Blein *et al.*, 2008). En outre, le maïs deviendrait une culture commerciale et assurerait la sécurité alimentaire mieux que toute autre culture (Smith *et al.* 1997) par l'adoption d'une gamme de variétés tolérantes au facteur abiotique qu'est la sécheresse. En effet, le maïs est prévu pour devenir la plus grande récolte en termes de production

d'ici 2050 (Rosegrant *et al.*, 2008).

### **I.3 CONTEXTE AGRO-ECONOMIQUE ET POLITIQUES D'INTEGRATION SOUS-REGIONALE**

L'Afrique de l'Ouest couvre les 16 pays de la CEDEAO (Communauté Economique des États de l'Afrique de l'Ouest) en plus de la Mauritanie qui a réintégré cette espace politico-économique depuis 2017. Cependant, la plupart des études socio-économiques et environnementales réalisées dans la sous-région incluent le Tchad et le Caméroun, voire la Centrafrique, pour des raisons de continuité économique, de peuplement et de complémentarité à l'échelle régionale (OCDE, 2007). Cette région couvre une superficie de 7.800.000 km<sup>2</sup>, soit un peu moins de 2 fois celle de l'union européenne ; avec une population de 290 millions d'habitants en 2010 ; soit 28% par rapport à celle du continent africain avec un taux d'urbanisation de 41%. Elle représente donc la 4<sup>ème</sup> région la plus peuplée au monde, derrière la Chine, l'Inde et les Etats-Unis (OCDE, 2013). Elle présente une croissance agricole impressionnante avec une production agricole brute qui augmente annuellement de 3,7% en moyenne entre 1980 et 2010 ; plus que la moyenne mondiale (2,2%) et une croissance de la production des céréales plus rapide encore (OCDE, 2013). Cependant, la sécurité alimentaire reste inaccessible face, principalement, aux pertes post-récoltes essentiellement dues aux insectes ravageurs. Notre zone d'étude recouvre 7 pays de la sous-région ouest-africaine à savoir le Burkina Faso, la Côte d'Ivoire, le Mali, la Mauritanie, le Niger, la République de Guinée et le Sénégal. A ces pays nous avons ajouté 2 pays de l'Afrique centrale le Caméroun et la Centrafrique. Ces derniers ont été inclus afin de mettre en exergue les relations phylogénétiques entre les populations de *S.zeamais* dans les 2 sous-régions. Ces différents pays ont été choisis principalement selon l'intensité de la culture du maïs et la présence du genre *Sitophilus*.

#### **I.3.1 ORGANISMES D'INTEGRATION DANS LA SITUATION AGRO-ECONOMIQUE SOUS REGIONALE.**

Selon le FMI (2013), la sous-région ouest-africaine de par la CEDEAO constitue la 20<sup>ème</sup> puissance économique du monde et se caractérise par une hétérogénéité économique de ses différents pays. En effet, elle englobe quatre pays non PMA (Pays Moins Avancés) (Nigeria, Côte-d'Ivoire, Ghana et Cap Vert) (CACID, 2012). Selon le OCDE (2006), la sous-région peut être subdivisée en 3 sous-ensemble économiques avec pour chacun une contribution de plus 25% au PIB (Figure 3) : les pays du golfe de Guinée autour du Nigeria qui constituent le cœur de l'économie régionale ; la façade atlantique qui a pour option soit une extraversion ou une insertion dans l'espace régional et les grands pays enclavés qui sont caractérisés par une marge de manœuvre étroite et donc ont un avenir nécessairement régional. Globalement, la part du secteur agricole au PIB brut régional est d'environ 35% pour plus de 60% de la population active de la sous-région (Blein *et al.*, 2008 ; CEDEAO, 2008). En fait, l'agriculture ouest-africaine est dans une phase de transition. Certaines exploitations agricoles sont désormais bien insérées dans le marché et fonctionnent dans une logique pleinement marchande, alors que pour de nombreux ménages, l'agriculture reste encore et avant tout

un mode de vie (voire de survie). Même dans l’hypothèse d’un développement rapide des marchés, cette dualité de l’agriculture restera durablement une caractéristique de la région (OCDE, 1998). En outre, les régions à climat humide et semi-humide jouent un rôle déterminant dans la dynamique du secteur agricole régional. C’est notamment le cas des franges transfrontalières qui traversent le nord des pays côtiers et le sud des pays sahéliens. Ces zones sont dynamiques à la fois sur le plan des productions et sur le plan des échanges. C’est la principale zone de production alimentaire (Figure 4), concourant largement à la sécurité de l’approvisionnement alimentaire des populations de la CEDEAO (Blein *et al.*, 2008).

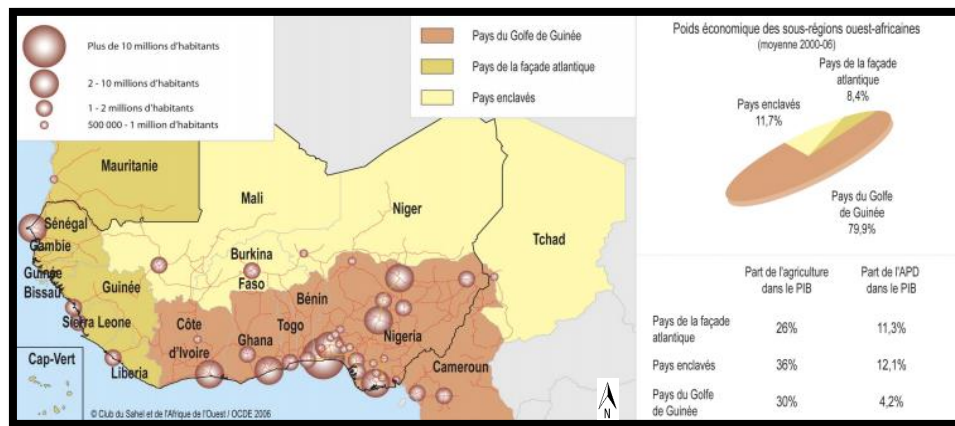


Figure 3 : Les sous-ensembles économiques de l’Afrique de l’Ouest ; OCDE, 2006.

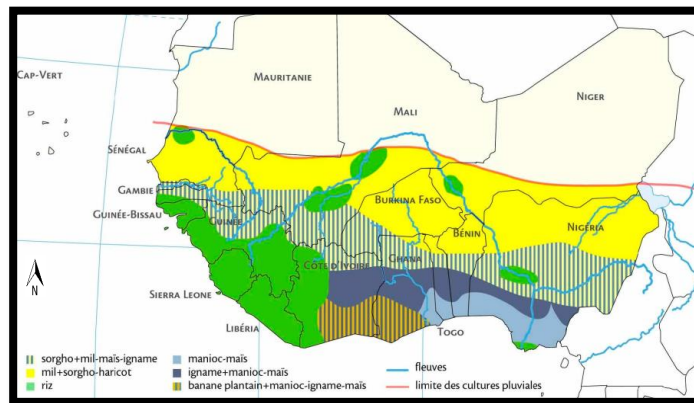


Figure 4 : Principaux bassins de production des cultures vivrières (Blein *et al.*, 2008)

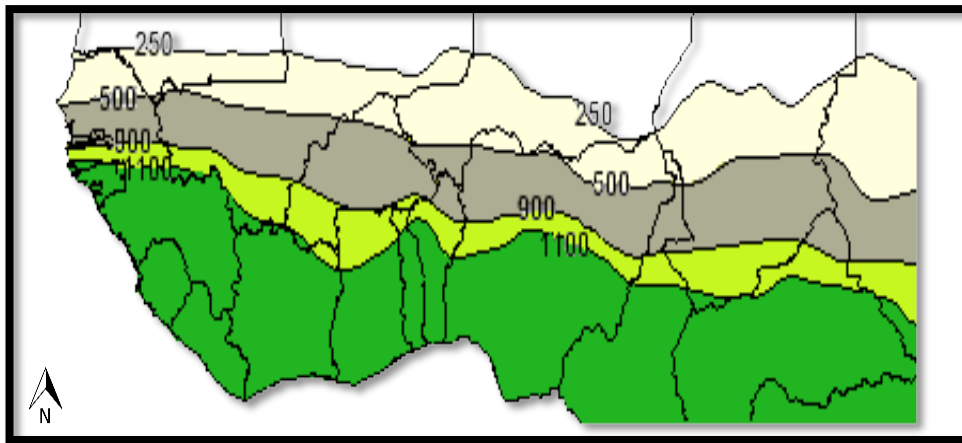
La sous-région constitue aussi une communauté de peuples, caractérisée par le dynamisme d’organismes comme notamment la CEDEAO qui est une organisation politique, intergouvernementale et de coopération sous-régionale regroupant tous les 16 pays de la sous-région à l’exception de la Mauritanie. Il y a aussi l’UEMOA (Union Economique et Monétaire Ouest Africaine) qui œuvre pour l’intégration économique avec 8 pays (Bénin, Burkina Faso, Côte d’Ivoire, Guinée-Bissau, Mali, Niger, Sénégal et Togo) ou le CILSS (Comité permanent Inter-état de Lutte contre la Sécheresse dans le Sahel) qui est recentré sur des activités comme la sécurité alimentaire et la gestion des ressources naturelles avec 13 pays (Bénin, Burkina Faso, Cap Vert, Côte d’Ivoire, Gambie, Guinée, Guinée-Bissau, Mali, Mauritanie, Niger, Sénégal, Tchad et Togo).

### I.3.2 LES DIFFERENTES ZONES AGRO-ÉCOLOGIQUES DE LA SOUS-REGION OUEST-AFRICAINE

En Afrique de l'Ouest, en général, les zones agro-écologiques se distinguent sur la base de la durée de la période croissante, c'est-à-dire, la période où l'eau est disponible pour la production végétale dans les sols bien-drainés. Cette période est en fonction des précipitations, de l'évaporation et de l'eau disponible dans le sol (Hell *et al.*, 2008). D'autre part, on peut distinguer isolément dans chaque pays des zones agro-écologiques qui regroupent des communautés rurales ayant des similitudes en termes de structure du sol et de climat (ESASU, 2008). En effet, ces différentes zones peuvent être identifiées en fonction du climat, de l'écologie et des pratiques agricoles englobant des dimensions environnementales, sociales, économiques et ethniques (Ndong, 2015). Il y a globalement, du Nord au Sud, 3 biomes terrestres : le désert, la savane et la forêt tropicale. D'après FAO (2003, 2004), il y a 4 zones agro-écologiques/climatiques du Nord au Sud caractérisées suivant une pluviométrie croissante (mm) (FAO, 2004 ; OCDE et CEDEAO, 2008) et leurs caractéristiques agricoles (Figure 5) :

- **Zone sahélienne** (zone aride, < 500 mm) : les précipitations annuelles moyennes varient de 250 à 500 mm. C'est la zone située à la limite de la végétation pérenne ; là où les précipitations sont inférieures à 350 mm, il n'y a que des pâturages et, parfois, des cultures céréalières à cycle court résistant à la sécheresse. Dans cette zone, toutes les activités agricoles sont hautement aléatoires.
- **Zone soudano-sahélienne** (zone semi-aride, 500-1000 mm) : les précipitations annuelles se situent entre 500 et 900 mm. On pratique surtout des cultures ayant un cycle de végétation bref de 90 jours, c'est-à-dire principalement du sorgho et du mil.
- **Zone soudanienne** (zone sub-humide, 1000-1500 mm) : les précipitations annuelles moyennes varient de 900 à 1100 mm. La plupart des céréales cultivées ont un cycle de végétation de 120 jours ou plus. C'est la zone où l'on produit l'essentiel des céréales, notamment du maïs, des racines et tubercules, et des cultures de rapport.
- **Zone guinéenne** (zone humide, > 1500) : les précipitations annuelles moyennes dépassent 1100 mm. Dans cette zone, il est plus facile de cultiver des racines.

Cette sous-région couvre un gradient pluviométrique qui oscille entre moins de 150 mm en zone saharienne au Nord à près de 3.000 mm dans le sud des pays côtiers du Golfe de Guinée. Ceci induit une très large diversité de systèmes de production agricole en fonction des opportunités de développement des cultures potentiellement rentables (CEDEAO, 2008). La zone subhumide caractérisée par une diversité climatique où la diversité des cultures est plus accentuée, constitue la principale zone de culture du maïs qui y domine (FAO, 2014).



**Figure 5** : Les différentes zones agro-écologiques de l’Afrique de l’Ouest ; FAO, 2004

### I.3.3 PARTICULARITES DE LA ZONE SAHELO-SOUDANIEENNE ET DE LA REGION GUINEO-CONGOLAISE

L’Afrique de l’Ouest peut être subdivisée globalement, en 4 régions biogéographiques dont le critère principal est la présence d’un gradient croissant de la pluviométrie, du Nord au Sud, qui module les paramètres écologiques (température, humidité relative, *etc...*). Il s’agit du désert du Sahara, de la zone transitoire du Sahel, des savanes de la zone soudanienne et des forêts de la zone guinéenne (White, 1986 ; Mestre et Chiffaud, 2005). Dans cette étude, nous avons globalement scindé les localités échantillonnées en deux groupes correspondant à deux régions biogéographiques qui correspondent aux deux biomes définis d’après Olson *et al.* (2001) :

- les forêts tropicales et subtropicales qui correspondent à la région biogéographique Guinéo-Congolaise de White (1983) ;
- et les prairies tropicales et subtropicales, savanes et arbustives qui équivalent au domaine englobant la zone de transition du Sahel et le centre d’endémisme de la région soudanienne de White (1983) que nous nommerons sous le terme « zone sahélo-soudanienne » dans l’analyse des résultats.

De nombreuses études ont déterminé des régions biogéographiques et/ou des biomes basés sur des paramètres climatiques et/ou bioécologiques dans la sous-région de l’Afrique de l’Ouest et du Centre. Cependant, il y a des différences dans la nomination de ces zones et les localités qu’elles ont exactement englobées (White, 1983; Olson *et al.*, 2001; Arbonnier, 2004; Kottek *et al.*, 2006; Linder *et al.*, 2012). Ainsi, le Sahel est une région semi-aride située dans le sud du Sahara et se situe principalement entre 10 et 20 "N (OMS, 2001 ; Epule *et al.*, 2017). Il est caractérisé par une variabilité climatique importante et a 2 caractéristiques climatiques : une saison des pluies par an et un mois d’Août très pluvieux (Heinrigs, 2010). La pluie (100 mm et 700 mm) constitue de loin la variable climatique la plus déterminante pour la vie des populations ; certains auteurs considèrent qu’elle permet à elle seule de déterminer l’évolution de l’environnement dans cette région du monde (AGRHYMET, 2010). Il fait partie des zones endémiques de l’insécurité alimentaire et est caractérisé par l’exacerbation des pertes post-récoltes due à la longue période de stockage car la production agricole est généralement saisonnière (Ngamo et Hance, 2007). Cette bande de terres traverse six

pays de l’Afrique de l’Ouest : la Mauritanie, le Sénégal, le Mali, le Burkina Faso, le Niger et le Nigéria (OCDE, 2006). C’est une zone de contact située à la limite de deux milieux plus ou moins différents regorgeant une forte diversité entomologique. Le Sahel constitue donc un écotone (Uvarov, 1957 *in* Popov, 1996). D’après cet auteur, ces écotones répondent le mieux aux besoins écologiques des insectes dans leur recherche de milieu servant de survie, de refuge, d’alimentation au cours du développement larvaire et en partie pendant la vie imaginaire ; et de milieu de reproduction constitué par les plages de sols nus, ensoleillés, convenant à la ponte et au développement embryonnaire. Alors que le centre d’endémisme de la région soudanienne forme une ceinture de la côte ouest en Guinée à la côte de la mer rouge au Soudan et en Erythrée. Sa pluviométrie donne environ 1 000 mm par an au sud où il y a des bois ; à environ 600 mm par an dans les prairies du Nord. Enfin, la région Guinéo-Congolaise, qui est une région humide (Couteron, 1996 ; FAO, 2002), a été définie par Anhuf (2000) après la dernière maximum glaciaire (8000 B.P) et correspond aux refuges forestiers putatifs proposés par Maley (1996) lors des changements climatiques du Pléistocène (18000 B.P). Enfin, d’après White (1983), la forêt pluvieuse de l’Afrique de l’Ouest et Centrale forme la région Guinéo-Congolaise qui était un centre phytogéographique d’endémisme d’espèces et a été subdivisée en trois sous-centres : la Haute-Guinée, la Basse-Guinée et le Congolia. Il abrite le onzième des 25 sites qualifiés de « points chauds » de biodiversité ou *biodiversity hotspot* avec 0,8 et 1,0 % pour des plantes et vertébrés endémiques respectivement dans le monde (Myers, 1988 ; 1990 ; Mittermeier *et al.*, 2000 ; Myers *et al.*, 2000). Ces sites correspondent à des régions biogéographiques forestières de la sous-région ouest-africaine et sont caractérisées par une grande richesse de biodiversité mais menacées par les activités anthropiques (Myers *et al.*, 2000). Cette richesse de la biodiversité de ces sites serait liée au fait qu’ils correspondraient aux zones refuges pendant les périodes défavorables de glaciations (Maley, 1996 ; 2001). Ils recouvrent également les forêts guinéennes de l’Afrique de l’Ouest et sont divisés en 2 sous-régions par le Dahomey Gap, une région de savanes et de forêts sèches du Togo et du Bénin dont l’existence pourrait être liée à 2 facteurs selon Ballouche *et al.* (2000) : le climat et l’action anthropique.

## **I.4 CONCEPTS ET NOTIONS FONDAMENTALES EN GENETIQUE DES POPULATIONS**

### **I.4.1 ESPECE**

Selon Lee (2003), la notion d’espèce a une définition plurielle qui varie d’un concept à l’autre suivant les domaines scientifiques ; et Reydon (2004) décompte des douzaines de concepts d’espèce définis dans la littérature scientifique. Dans le cadre de cette thèse, le terme « espèce » fait référence au concept biologique de l’espèce de Mayr (1942) qui fait référence au critère d’isolement reproducteur par le biais des barrières géniques (Dobzhansky, 1937). Alors, l’espèce biologique se définit comme des groupes de populations naturelles qui sont génétiquement isolés d’autres groupes similaires et dont les individus sont effectivement ou potentiellement capables de se reproduire entre



eux pour engendrer une descendance viable et féconde dans les conditions naturelles (Coyne et Orr, 2004 ; Futuyma, 2005). Ce concept semble la plus pertinente dans ce cas où il est question de contrôle de populations nuisibles. La quantification du degré de connectivité (i.e. flux génique) entre les populations délimitées de façon précise dans le système étudié en fonction de l'espace et du temps est primordial (Webster *et al.*, 2002). En effet, c'est une étape cruciale dans la définition des stratégies d'intervention dans la gestion des populations naturelles. Cela permet de déterminer si l'ensemble des populations étudiées forment un groupe plus ou moins homogène et dynamique (métapopulation) ou les différentes populations se sont adaptées à leurs milieux respectifs engendrant une différenciation et structuration génétiques distinctes. Dès lors que les individus d'une espèce sont souvent regroupés naturellement en populations locales soumises à différentes forces évolutives et écologiques, et de différentes intensités ; ce qui engendre des compositions génétiques distinctes pour chacune des populations.

#### **I.4.2 SOUS-ESPECE**

Une sous-espèce constitue une des étapes du processus de la spéciation définie selon Mayr (1942). En effet, une sous-espèce est formée par un ensemble d'individus isolés des autres individus de la même espèce pour des raisons anatomiques (incompatibilité des organes reproducteurs mâle et femelle), écologiques (distance géographique, barrières environnementales...) ou éthologiques (saison de reproduction). Il faut noter, cependant, que l'interfécondité demeure entre les sous-espèces malgré leurs divergences qui ne sont pas encore capables d'induire des caractères propres à une d'espèce.

#### **I.4.3 POPULATION**

La notion de « population » la plus utilisée en biologie et plus spécifiquement en « Génétique des Populations » est défini par Hartl (1994) qui la considère comme étant une unité évolutive de reproduction. Dès lors, une population est un groupe d'individus de la même espèce vivant dans une zone géographique suffisamment restreinte de telle sorte que chaque membre ait potentiellement une chance équivalente de se reproduire avec chaque membre du sexe opposé (Hartl et Clark, 2007). Cette notion d'unité de reproduction est essentielle dans le cadre de gestions des populations d'insectes ravageurs.

#### **I.4.4 METAPOPOPULATION**

Introduite par Richard Levins en 1969 avec un modèle déterministe d'occupation/absence pour étudier les insectes nuisibles dans les surfaces agraires, une métapopulation est une structure dynamique composée de plusieurs populations séparées spatialement ou temporellement et connectées par la migration.

### **I.4.5 ECOTYPE**

Introduite par Turesson (1922a, b) la notion d'écotype désigne un individu ou une population d'individus d'une espèce, d'une sous-espèce ou d'une race qui présente des caractéristiques morphologiques, physiologiques adaptées à un type de milieu en particulier. Ces caractéristiques propres à l'écotype sont héréditaires et peuvent induire des variations morphologiques plus ou moins importantes car étant déterminées par les conditions du milieu qui est dit écotypique. Une population écologiquement différenciée constitue donc un écotype. Ces écotypes sont appelés variétés ou sous-espèces voire espèce dans le cas où le milieu a induit un isolement reproductif complet par rapport à la population ancestrale. C'est donc dû à une adaptation locale régie par une sélection directionnelle. Les adaptations de ces écotypes sont basées sur les interactions de leurs propres ensembles spéciaux de gènes avec leur propre environnement.

### **I.4.6 BIOTYPE**

Un biotype constitue un ensemble d'individus ayant le même génotype, le même patrimoine héréditaire homozygote. Chez les insectes ravageurs, un biotype fait référence, en général, au régime alimentaire (type de céréales ou de légumineuses, *etc...*) qui modulent les caractéristiques génétiques généralement cryptiques des populations. Un exemple typique est mis en exergue par l'étude Sembène *et al*, (2012a) sur la bruche de l'arachide où des populations inféodées à des plantes différentes se distinguent les unes les autres sur le plan génétique. Ainsi, la différenciation génétique de telles populations est en fonction de la plante-hôte et donc du régime alimentaire.

## **1.5 APPROCHES ET METHODOLOGIE UTILISEES EN GENETIQUE DES POPULATIONS**

La configuration d'une population est régie par un ensemble de forces évolutives et écologiques qui font varier sa constitution génotypique et phénotypique et moduler son évolution. La compréhension de ces processus est donc capitale pour une bonne gestion des populations naturelles. A cela, les modèles mathématiques constituent un socle dans la représentation hypothétique de la population à étudier ; et ce sont les écarts aux hypothèses concernant la population théorique qui rendent compte de la réalité vécue par la population. Ces hypothèses reposent sur le régime reproductif, la taille de la population et les forces évolutives pouvant les gouverner, *etc...* De ce fait, les fréquences génotypiques et alléliques peuvent être liées par des relations qu'il est possible de formaliser suivant un modèle mathématique dans le but de quantifier l'évolution d'une population donnée (Huvet, 2000).

### **I.5.1 QUANTIFICATION DE L'EVOLUTION PAR APPROCHE PROSPECTIVE**

#### **I.5.1.1 LOI DE HARDY-WEINBERG D'UNE POPULATION INFINIE**

L'équilibre de Hardy-Weinberg (Lange, 2002) est le principe théorique et fondamental utilisé en génétique des populations. Il a été élaboré indépendamment et simultanément en 1908 par le

mathématicien anglais, Hardy, et le médecin allemand, Weinberg. Cette loi stipule que les fréquences alléliques à un locus donné d'un organisme diploïde et sexué restent invariantes d'une génération à l'autre, et par conséquent, les fréquences génotypiques calculées à partir de ces fréquences restent invariantes d'une génération à l'autre. Ce modèle d'évolution est soumis aux hypothèses/conditions suivantes :

- 1) la population est panmictique : les couples se forment au hasard (panmixie) et leurs gamètes se rencontrent au hasard (pangamie) ;
- 2) la population est de taille "infinie" afin de minimiser les variations d'échantillonnage, et donc de la dérive génétique ;
- 3) la population n'est soumise ni à la sélection (même probabilité de survie des gamètes jusqu'à la fécondation, puis des zygotes jusqu'à l'âge de la reproduction, indépendamment de leur génotype), ni mutation (fréquences alléliques ne sont pas modifiées car il n'y a pas de création de nouveaux allèles par modification), ni migration (pas d'échanges de gènes entre populations) ;
- 4) les générations successives sont discrètes (pas de croisement entre générations différentes).

Cependant, cette loi n'est jamais parfaitement observée dans la nature et, en réalité, ce sont les écarts au modèle qui sont informatifs. La détection d'éventuels écarts au modèle théorique est faite par les tests de conformité qui permettent de déterminer les facteurs influençant l'évolution et la constitution génétique de la population.

### CONSEQUENCES DE LA LOI

Cependant, cette loi n'est jamais parfaitement observée dans la nature et, en réalité, ce sont les écarts au modèle qui sont informatifs. La détection d'éventuels écarts au modèle théorique est faite par les tests de conformité qui permettent de déterminer les facteurs influençant l'évolution et la constitution génétique de la population. Le plus utilisé de ces tests est le Test du  $\chi^2$  de Pearson. Toutefois, lorsque la fréquence d'un des génotypes présents est faible, le test exact de Fisher est plus recommandé. Ces éventuels écarts dans une population traduit donc qu'une des hypothèses, énoncées dans le modèle, n'est pas vérifiée. Dès lors, les écarts observés dans une population naturelle peuvent être dus :

- au système de reproduction comme la consanguinité par homogamie ou par hétérogamie ;
- à la sélection naturelle : la capacité de survie et la performance reproductive des individus sont liées à leurs génotypes ; et donc la constitution génotypique de la génération suivante se fait de façon inégale (Hartl et Clark, 2007). Ceci peut aboutir à une adaptation locale selon Salemi et Vandamme (2003) par la fixation de certains allèles (certains génotypes) ;
- à la dérive génétique : c'est la variation des fréquences alléliques due au hasard des événements de reproduction dans une population finie (Wright, 1931). Elle provoque donc l'hétérogénéisation de deux populations indépendantes par le simple fait du hasard. La population étudiée étant de petite taille, en plus d'une fluctuation hasardeuse des fréquences alléliques au cours du temps ; ce qui aboutit

à la fixation et la perte d'allèles. Elle s'oppose donc à la sélection naturelle en fixant des allèles contre-sélectionnés. Elle aurait tendance à réduire la diversité génétique au sein des populations et à augmenter la divergence entre les populations ;

- aux mutations qui augmentent le nombre d'allèles (haplotypes) dans une population donnée et s'accumulent avec le temps (Chao, 1990) et donc la divergence entre populations ;
- à la migration qui, en fonction de son importance, vise à homogénéiser les populations (Slatkin, 1987) et participe dès lors à la réduction de la divergence entre les populations ;
- à la subdivision de la population globale en différentes sous-populations caractérisées par une absence de panmixie et des fréquences alléliques différentes, ce phénomène est nommé effet Walhund (Walhund, 1928) ;
- à l'admixture qui est le résultat d'un second contact de populations préalablement séparées et génétiquement différentes ;
- échantillon biaisé (ascertainment) ;
- erreurs de génotypage.

#### **I.5.1.2 MODELE DE WRIGHT-FISHER D'UNE POPULATION FINIE**

Ce modèle est introduit par Fisher (1930) et Wright (1931) et constitue le modèle le plus largement utilisé en génétique des populations. Il permet de suivre et d'évaluer l'évolution d'un ou plusieurs allèles à l'intérieur d'une population diploïde ou haploïde ; c'est une illustration du phénomène de perte de diversité allélique au sein d'une population isolée. Il est conçu avec les hypothèses suivantes :

- 1) la population, à reproduction sexuée avec des possibilités d'autofécondation, est considérée comme panmictique : les couples se forment au hasard (panmixie) indépendamment du phénotype et/ou du génotype et les gamètes s'unissent indépendamment de leur génotype (pangamie) ;
- 2) elle est à l'équilibre démographique avec une taille constante ( $N$  allèles ou  $2N$  allèles pour respectivement les individus haploïdes et diploïdes) ;
- 3) ses générations sont non-chevauchantes ;
- 4) il n'y a pas de migration (pas d'échanges de gènes entre populations), ni sélection (tous les individus ont la même fitness i. e. même chance de survie et de reproduction) ;
- 5) avec des mutations neutres apparaissant à taux  $\mu$  constant et des variations des fréquences alléliques au cours du temps par effet stochastique (dérive génétique).

Le processus stochastique d'échantillonnage des individus participant à la reproduction d'un nombre limité d'individus, change la fréquence des allèles d'une génération à l'autre. Ce processus engendre, au cours des générations, la fixation d'allèles et entraîne une perte de diversité génétique. Il peut être utilisé pour prédire l'évolution du polymorphisme moléculaire sous le modèle neutre de l'évolution à l'équilibre mutation/dérive et peut donc être utilisé pour détecter d'éventuels signes de

sélection ou de changements démographiques. A l'équilibre mutation-dérive (Kimura, 1968), les allèles nouvellement créés par la mutation compensent ceux qui disparaissent sous l'effet de la dérive génétique ; ainsi, la population est isolée (pas d'échange de gènes par migration) et à l'équilibre démographique (taille constante au cours du temps) où il y'a pas de sélection. D'après Denton (2010), l'état d'équilibre mutation-dérive est largement respecté, puisque la distance génétique entre les différents phylums reste très faible. Sous l'hypothèse d'une population à l'équilibre mutation-dérive, il existe des méthodes pour détecter des écarts à cet équilibre dont une variation d'effectifs démographiques de la population. Ces effets peuvent être causés par des variations climatiques, par exemple.

### **I.5.2 QUANTIFICATION DE L'EVOLUTION PAR APPROCHE RETROSPECTIVE : THEORIE DE LA COALESCENCE**

La théorie de la coalescence (Kingman, 1982 et 2000 ; Rosenberg et Nordborg, 2002 ; Nordborg, 2004) repose sur la reconstruction rétrospective de l'histoire d'un échantillon de gènes sélectivement neutres d'une population sous l'impact de la dérive génétique. En plus, la population à étudier est caractéristique du modèle de Wright-Fisher où de génération en génération, il y a fixation et extinction de gènes et de lignées par dérive génétique car les gènes présents sont originaires du pool de  $2N$  gènes de la génération précédente et ainsi de suite. De ce fait au bout d'un certain nombre de générations en remontant dans le passé, il y a identification d'un ancêtre commun (MRCA : « Most Recent Common Ancestor ») pour les différentes générations explorées. Ainsi, la généalogie des gènes échantillonnés est établie. De ce fait, la coalescence utilise le polymorphisme moléculaire neutre pour connaître l'histoire évolutive des populations.

De nombreuses méthodes s'appuient sur cette théorie pour explorer et décrire les différents évènements qui ont eu à marquer l'histoire évolutive et démographique (taille de la population, taux de migration entre populations...) des populations naturelles par l'estimation d'un ou plusieurs paramètres du modèle (Beaumont et Rannala, 2004 ; Excoffier et Heckel, 2006) à partir de la variation génétique observée (Girod, 2010). En effet, il y a un lien probabiliste direct entre les fréquences des allèles et leur âge (Watterson et Guess, 1977 ; Donnelly et Tavaré, 1986) car les haplotypes/allèles les plus fréquents seraient en moyenne les plus anciens. Dès lors, elle aide à comprendre les forces évolutives qui sont les plus compatibles avec la généalogie observée (Montaigne, 2011).

### **I.5.3 ANALYSES STATISTIQUES PAR LES TESTS D'HYPOTHESES**

Les méthodes de détermination de la structure génétique et/ou du flux génique de populations à partir des données moléculaires se fait à partir de l'estimation de paramètres qui repose sur un certain nombre d'hypothèses dont les plus importantes sont :

- 1) la taille effective des populations reste stable au cours du temps ;
- 2) les gènes échantillonnés sont sous neutralité sélective ;
- 3) les taux de migration entre populations sont égaux ;

- 4) les loci ne sont pas liés (déséquilibre de liaison) ;  
 5) et il n'y a pas de recombinaisons au sein des gènes (les allèles ont des histoires évolutives indépendantes).

Ces hypothèses peuvent être irréalistes et représentent alors des limites significatives en fonction des résultats et des conclusions subséquentes de ces analyses. Cependant, les récents progrès, notamment à travers l'utilisation de la théorie de coalescence, permettent d'évaluer indépendamment certaines de ces hypothèses.

#### I.5.4 L'INFERENCE BAYESIENNE, LA PHYLOGENIE ET LA GENETIQUE DES POPULATIONS

L'inférence bayésienne est le processus d'adaptation d'un modèle de probabilité à un ensemble de données et résume le résultat par une distribution de probabilité sur les paramètres du modèle et sur des quantités inaperçues telles que des prévisions pour de nouvelles observations (Gelman *et al.*, 2014). L'idée centrale dans une analyse bayésienne est que les paramètres du modèle statistique doivent être traités comme des variables aléatoires ; les inférences sont alors basées sur la probabilité du paramètre dépendant des observations. Cette probabilité conditionnelle peut être calculée en combinant la vraisemblance avec un *a priori* utilisant le théorème de Bayes (Huelsenbeck, 2000). L'analyse bayésienne est donc une méthode probabiliste qui consiste à baser l'estimation des paramètres sur une distribution nommée loi *a posteriori* qui mélange l'information portée par les données (*via* la vraisemblance) et une information *a priori*.

La probabilité postérieure d'une hypothèse pouvant être calculée par le théorème de Bayes, elle apparaît être fonction de la vraisemblance et de la probabilité *a priori* de cette hypothèse selon l'équation suivante :

$$P(H|D) = \frac{P(D|H) * P(H)}{P(D)}$$

Où  $P(D|H)$  est la fonction de vraisemblance ou la probabilité d'observer les données D sous une hypothèse H ;  $P(H)$  : la probabilité *a priori* de l'hypothèse H et  $P(D)$  : la probabilité des données D. La probabilité postérieure d'une hypothèse H sachant les données D peut donc être interprétée comme la probabilité que cette hypothèse soit vraie sachant les données D.

##### I.5.4.1 INFERENCE BAYESIENNE ET PHYLOGENIE

Selon Delsuc et Douzery (2004), l'alliance entre la phylogénie et l'approche bayésienne a débuté en 1996 par trois études indépendantes (Li, 1996 ; Mau, 1996 ; Rannala et Yang, 1996) qui ont formalisé le calcul des probabilités postérieures (calculées *a posteriori*) des arbres phylogénétiques à partir de probabilités définies *a priori*. En phylogénie, l'inférence bayésienne est typiquement appliquée pour créer la distribution postérieure pour un paramètre basé sur la probabilité

*a priori* de ce paramètre et la vraisemblance des données de séquence d'ADN (Lu et Hanrahan, 2014). Ce qui distingue l'avantage de cette méthode de son prédécesseur, le maximum de vraisemblance. En effet, la notion de probabilité postérieure considérée ici est la probabilité de l'hypothèse H sachant les données X :  $P(H|X)$ , ce qui diffère du maximum de vraisemblance où l'on cherche la probabilité d'observer les données X sous une hypothèse H :  $P(X|H)$  (Delsuc et Douzery, 2004). En plus, le maximum de vraisemblance se base entièrement sur des inférences de la fonction de la vraisemblance, n'incorporant aucune information préalable et choisissant des estimations de point des paramètres qui maximisent la probabilité des données déterminées du paramètre (c'est-à-dire, maximisant la vraisemblance d'une fonction de paramètres pour un ensemble défini de données). Alors que les statistiques bayésiennes permettent aux scientifiques d'incorporer facilement une connaissance à priori dans l'analyse des jeux de données (Beaumont et Rannala, 2004). Ainsi, l'inférence bayésienne de la phylogénie combine la probabilité *a priori*  $P(A)$  d'un arbre A avec la vraisemblance  $P(X/A)$  des données X sachant cet arbre A pour produire une distribution de probabilité postérieure  $P(A/X)$  sur les arbres en utilisant la formule de Bayes :

$$P(A|X) = \frac{P(X|A) * P(A)}{P(X)}$$

- $P(A|X)$  : C'est la distribution à estimer ; elle est appelée la distribution à *posteriori* sur les arbres phylogénétiques (*posterior distribution*). C'est une distribution conjointe, c'est-à-dire qu'il s'agit de la distribution de tous les paramètres du modèle d'évolution des séquences de l'arbre phylogénétique A connaissant les données X. La probabilité postérieure d'un arbre A pouvant être interprétée comme la probabilité que cet arbre soit vrai sachant les données X ;
- $P(X|A)$  est appelée la vraisemblance des données X sachant l'arbre A, aussi écrit  $L(X)$  ;
- $P(A)$  est appelée la distribution *a priori* (ou *prior distribution*) des paramètres d'un arbre phylogénétique A ;
- $P(X)$  : C'est la vraisemblance marginale des données.

La distribution postérieure des paramètres est estimée à partir des données et des distributions *a priori* sur les paramètres. L'inférence bayésienne dépend, dépend par conséquent, fortement de la probabilité postérieure et dans la plupart des cas, la sommation et les intégrales sont difficiles à être calculées par les méthodes analytiques (Huelsenbeck, 2000 ; Delsuc et Douzery, 2004 ; Lu et Hanrahan, 2014) ; et en raison de l'impossibilité d'explorer tout l'espace des paramètres (Girod, 2010). D'où l'adoption de la méthode algorithmique utilisant une technique d'échantillonnage de l'espace, Monte Carlo, par une chaîne de Markov qui est définie comme une séquence de variables aléatoires et que le passage à l'état E+1 dans la séquence ne dépend que de l'état E ; ce processus est appelé MCMC (Markov Chain Monte Carlo) (Beaumont et Rannala, 2004 ; Delsuc et Douzery, 2004 ; Girod, 2010 ; Lu et Hanrahan, 2014). Le principe de cette méthode est donc de construire une chaîne

de Markov qui va converger vers la distribution stationnaire des paramètres. Cette distribution stationnaire est la distribution postérieure des paramètres qui présente la vraisemblance la plus forte (Girod, 2010). Les différents états (E+1, E+2, E+3, E+4...) par lesquels passent les MCMC sont souvent désignés sous le terme de « générations » et les échantillons tirés de la chaîne de Markov représentent un échantillon de la distribution des probabilités postérieures (Delsuc et Douzery, 2004).

Cependant, du fait de l'impossibilité d'un échantillonnage exhaustif dans l'espace des paramètres par une chaîne de Markov (Girod, 2010) d'où le défaut de convergence des MCMC qui se manifestent par les optimums locaux (Delsuc et Douzery, 2004) ; et que l'approximation de la distribution sera d'autant plus exacte que le nombre de pas effectués par la chaîne de Markov sera élevé (Lewis, 2001), une variante de l'algorithme Metropolis Hastings est utilisée par un couplage de Metropolis des MCMC (MCMCMC ou MC<sup>3</sup> pour « Metropolis Coupling Markov Chain Monte Carlo ») qui contribue à la réduction des risques de défaut de convergence des MCMC (Huelsenbeck et Ronquist, 2001). Cet algorithme permet d'utiliser  $n$  MCMC simultanément dont les  $n - 1$  sont dites « chauffées » de manière graduelle. Ces chaînes « chaudes » explorant l'espace des paramètres plus largement que la chaîne « froide », des permutations entre les états des différentes chaînes « chaudes » et ceux de la chaîne froide sont tentées à chaque génération, ce qui peut permettre de sortir la chaîne « froide » d'un optimum local éventuel. Ainsi, ces chaînes « chaudes » sont utilisées pour guider la chaîne dite « froide » à partir de laquelle les inférences sont faites (Delsuc et Douzery, 2004). Par ailleurs, les probabilités postérieures bayésiennes représentent la probabilité qu'un clade donné soit vrai étant donné le modèle d'évolution, les probabilités *a priori*, et les données considérées (Huelsenbeck *et al.*, 2001).

#### **1.5.4.2 INFERENCE BAYESIENNE ET GENETIQUE DES POPULATIONS**

L'analyse bayésienne est basée sur trois étapes principales qui sont : formuler le modèle, adapter le modèle aux données et améliorer le modèle en vérifiant son ajustement et en le comparant à d'autres modèles (Gelman *et al.*, 2004 ; 2014). Les méthodes bayésiennes sont utilisées dans des domaines comme la génétique des populations pour inférer sur les processus d'évolutions ayant impacté sur les caractéristiques génétiques des populations ; notamment les approches bayésiennes de type ABC dans la compréhension des forces qui influencent la variation génétique au sein et entre les populations (Beaumont *et al.*, 2002 ; Beaumont et Rannala 2004 ; Csillery *et al.*, 2010). Le calcul bayésien approximatif ou ABC (*Approximated Bayesian Computation*) (Pritchard *et al.*, 1999 ; Beaumont *et al.*, 2002) est une méthode populaire pour obtenir une approximation des estimations *a posteriori* sans évaluer la vraisemblance. Selon Beaumont et Rannala (2004), les approches de type ABC contournent des calculs exacts de vraisemblance en employant des statistiques résumées et des simulations ; elles constituent une approche sans la vraisemblance pour les inférences bayésiennes basées sur une méthode d'algorithme de rejet qui applique une tolérance de dissemblance entre les



statistiques résumées de données observées et simulées. La base de l'algorithme de rejet, une technique simple pour générer des échantillons à partir d'une distribution de probabilité, consiste à simuler un grand nombre d'ensembles de données dans un scénario évolutif hypothétique. Les paramètres du scénario ne sont pas choisis de manière déterministe, mais échantillonnés à partir d'une distribution de probabilité. Les données générées par la simulation sont alors réduites à des statistiques résumées, et les paramètres échantillonnés sont acceptés ou rejetés sur la base de la distance entre la simulation et les statistiques sommaires observées. Les statistiques résumées sont des valeurs calculées à partir des données pour représenter la quantité de l'information maximale sous la forme la plus simple possible ; il y a l'exemple des logiciels M-ratio et BOTTLENECK (Girod, 2010) utilisés dans l'inférence de l'évolution démographique des populations. En plus, ces méthodes bayésiennes ont été également employées pour assigner des individus à leur population d'origine ou pour détecter si la sélection agit sur les gènes (Beaumont et Rannala, 2004).

## CHAPITRE II

# EVOLUTION PHYLOGEOGRAPHIQUE ET STRUCTURE GENETIQUE DES POPULATIONS DE *Sitophilus zeamais*, RAVAGEUR DU MAÏS STOCKE EN AFRIQUE DE L'OUEST

## **CHAPITRE II : EVOLUTION PHYLOGEOGRAPHIQUE ET STRUCTURE GENETIQUE DES POPULATIONS DE *SITOPHILUS ZEAMAI*S, RAVAGEUR DU MAÏS STOCKE EN AFRIQUE DE L'OUEST**

### **INTRODUCTION**

La connaissance de la distribution spatiale et temporelle de la variabilité génétique au sein et entre les populations constitue une des étapes cruciales dans l'établissement de stratégies de gestion des populations naturelles. En effet, la valeur adaptative ou fitness d'un caractère, d'un individu, d'une population voire d'une espèce serait tributaire de la variabilité génétique. Elle est perçue comme une garantie qui permettrait l'adaptation et la pérennité des populations dans leurs environnements changeants (Lacy, 1987 ; Luck *et al.*, 2003 ; Duminil, 2006 ; Montaigne, 2011 ; Greenbaum *et al.*, 2014). La caractérisation de cette composante temporelle par l'établissement de l'histoire évolutive des relations génétiques entre populations/lignées se fait en décrivant leurs liens de parenté au cours de l'évolution grâce à la phylogénie. Ceci permet de connaître leurs patrons historiques et d'établir des hypothèses sur les causes évolutives possibles ayant engendré la dynamique de la diversité observée (composante spatiale) car la structuration actuelle des populations résulte d'événements du passé (Grivet, 2002 ; Cyr, 2012). Ceci permettrait de mieux appréhender les mécanismes de cette évolution. En outre, l'étude des principes et processus qui gouvernent la distribution géographique de ces lignées généalogiques (lignées génétiques) est l'objet de la phylogéographie (Avise, 2000 ; 2004). Cette dernière intègre des informations géographiques et génétiques afin d'inférer l'histoire démographique et les processus évolutifs des espèces. Elle s'intéresse aux phénomènes génétiques et démographiques ayant conduit à la distribution et à la structuration actuelle des populations. A partir de là, elle tente de reconstruire des scénarios évolutifs prenant en compte les événements climatiques, géologiques et environnementaux ayant jalonné l'histoire des régions concernées. Selon Beaumont (2010), grâce aux progrès récents dans les techniques de séquençage de l'ADN, l'analyse phylogéographique de la variation moléculaire est devenue une approche de plus en plus importante pour trouver des indices d'interaction entre les facteurs écologiques, la dispersion et l'évolution moléculaire.

L'hypothèse de ce chapitre est qu'il y a un impact des zones biogéographiques et/ou des échanges commerciaux sur l'évolution et la structuration de la diversité génétique des populations de *S. zeamais* (Motschulsky, 1855). Pour vérifier cette hypothèse, nous avons effectué différentes activités grâce aux gènes mitochondriaux, le *cytochrome B* (*Cyt-B*) et le *cytochrome oxydase I* (*COI*) pour identifier les différents haplotypes circulant dans la sous-région ouest-africaine; caractériser la structure génétique et déterminer les relations phylogéographiques entre les populations. Nous avons aussi identifié l'origine probable de l'infestation et daté les périodes déterminantes dans l'histoire évolutive de *S. zeamais* (séparation de lignées, phases d'expansion, goulot d'étranglement...).

## II.1 MATERIEL ET METHODES

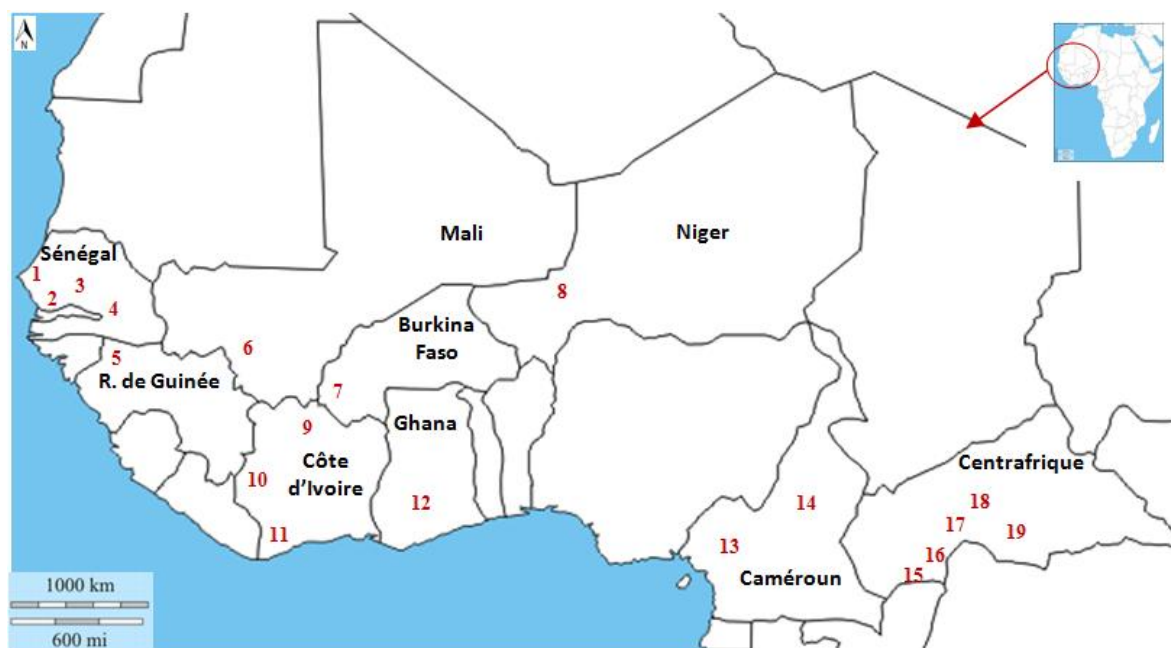
### II.1.1 ECHANTILLONNAGE

Les sites échantillonnés (Figure 6) dans cette étude se situent dans différentes localités de 6 pays de la sous-région ouest-africaine : Sénégal, République de Guinée, Mali, Burkina Faso, Côte d'Ivoire, Ghana et Niger. A ces 7 pays d'Afrique de l'Ouest, nous avons ajouté deux pays de l'Afrique centrale, le Camérout et la Centrafrique afin de mieux déterminer les impacts anthropiques des échanges commerciaux et écologiques dans la structuration génétique des populations. En effet, ils n'appartiennent pas aux organismes d'intégration de la sous-région ouest-africaine et en plus ils permettent d'approfondir les différences écologiques entre les pays de la zone sahélo-soudanienne et ceux de la région guinéo-congolaise. Dans notre étude, la population correspond à un ensemble d'individus issus d'un pays. Ceci suit les critères de définition d'Hartl et Clark, (2007) d'une « population évolutive » (voir page 25). Alors que le groupe de populations rassemble un ensemble de populations appartenant à une région biogéographique (zone Sahélo-soudanienne ou région guinéo-congolaise). Ces deux régions correspondent aux deux biomes définis d'après Olson *et al.* (2001) :

- les forêts tropicales et subtropicales qui correspondraient à la région biogéographique Guinéo-Congolaise de White (1983) ;
- et les prairies tropicales et subtropicales, savanes et arbustives qui équivaudraient au domaine englobant la zone de transition du Sahel et le centre d'endémisme de la région soudanienne de White (1983).

Dans notre stratégie d'échantillonnage, environ 250 g à 1 kg de maïs sont prélevés dans les lieux de stockage de producteurs grâce à l'appui de différents partenaires du projet. Ces échantillons sont ensuite acheminés au laboratoire puis conservés dans des bocaux à couvercle grillagés. Un suivi régulier de cet élevage de masse permettra de collecter les individus de *S. zeamais* qui émergent des échantillons. Ces spécimens sont ensuite conservés dans l'alcool 95° en vue de la réalisation de l'étude moléculaire. Chaque échantillon est identifié et codé selon un code alphanumérique défini par la 1<sup>ère</sup> lettre du nom binomial de l'espèce (S pour *Sitophilus* et z pour *zeamais*), la 1<sup>ère</sup> lettre est parfois suivi par une 2<sup>ème</sup> lettre du pays ; de même que pour la localité échantillonnée et enfin suivi du numéro d'ordre (exemple : SzGhA1 avec S = *Sitophilus*, z = *zeamais*, Gh = Ghana ; A = Accra et 1 = individu numéro 1). Les différents paramètres de l'échantillonnage sont résumés dans le tableau I.

**NB** : Dans le cas de l'échantillonnage au Camérout, les échantillons proviennent non d'une localité particulière mais d'une zone. En effet, dans ce pays chaque zone, regroupant plusieurs localités, cultive une variété de maïs qui y est la plus adaptée.



**Figure 6** : Zone d'étude et localités échantillonnées

**Légende** : 1 : Bambey, 2 : Keur-Ayip, 3 : Mbassis et 4 : Missirah (Sénégal) ; 5 : Labé (Guinée Conacry) ; 6 : Bamako (Mali) ; 7 : Bobodioulasso (Burkina Faso) ; 8 : Youry (Niger) ; 9 : Korhogo, 10 : Man et 11 : San Pédro (Côte d'Ivoire) ; 12 : Accra (Ghana) ; 13 : Zone centre et 14 : Zone ouest (Cameroun) ; 15 : Mbaïki ; 16 : PK12 ; 17 : Sibut ; 18 : Mala et 19 : Bambari (Centrafrique).

**Tableau I** : Tableau récapitulatif de l'échantillonnage

Pays	Localités	Codes d'échantillons	Nombre d'individus séquencés		Coordonnées géographiques	
			<i>Cyt-B</i>	<i>COI</i>	Latitude	Longitude
Burkina Faso	Bobodioulasso	SzBB	10	10	11°9'53,719" N	4°18'18,555" O
Cameroun	Zone centre	SzCmZC	04	04	7°22'10,999" N	20°40'13,9" E
	Zone ouest	SzCmZO	03	03		
Centrafrique	Bambari	SzCaBA	04	04	5°45'51,641" N	20°40'13,9 E
	Mala	SzCaMA	04	04	6°7'12,875" N	19°30'11,48" E
	Mbaïki	SzCaMB	04	04	3°51'59,406" N	17°58' 58,129" E
	PK12	SzCaPK	04	04	4°21,6732" N	18°33,2976" E
	Sibut	SzCaSI	04	04	5°44'49,827" N	19°5'49,145" E
Côte d'Ivoire	Korhogo	SzCIK	05	05	9°28'0,961" N	5°36'51,32" O
	Man	SzCIM	05	05	7°24'23,139" N	7°33'26,003" O
	San Pédro	SzCISP	05	05	4°45'28,327" N	6°38'32,758" O
Ghana	Accra	SzGhA	10	10	5°36'13,38 N	0°11'13,071" O
Mali	Bamako	SzMIB	10	10	12°38'21,234" N	8°0'10,401" O
Niger	Youri	SzNgY	10	10	13°20'08,0" N	2°14'44,0" E
République de Guinée	Labé	SzGL	10	10	11°19' 23,535" N	12°17'20,873" O
Sénégal	Bambey	SzSB	05	00	14°41'47,441" N	16°27'29,721 O
	Keur-Ayip	SzSKa	05	08	13°35'45,065" N	15°36'16,56" O
	Mbassis	SzSMb	05	07	14°26'54,842" N	16°40'47,204" O
	Missirah	SzSMi	05	05	13°31'25,845" N	13°30'42,775" O
<b>Total</b>			<b>117</b>	<b>112</b>		

## II.1.2 ETUDE GENETIQUE

### II.1.2.1 INTERET DU *CYTOCHROME B* ET DU *CYTOCHROME OXYDASE I*

La mitochondrie est considérée comme la « centrale énergétique » de la cellule car produisant la quasi-totalité de l'énergie sous forme d'ATP à travers la phosphorylation oxydative par la chaîne

respiratoire et le complexe de l'ATP synthase (Saraste, 1999). Collectivement, elles peuvent occuper 25% du volume du cytoplasme (Ballard et Whitlock, 2004) ; soit 10.000 à 100.000 mitochondries (Newman *et al.*, 1996). Selon la théorie endosymbiotique, elle est issue de la symbiose entre une bactérie de la classe des  $\alpha$ -protéobactéries avec la première cellule précurseur (Margulis, 1981 ; Gray *et al.*, 2001). La très grande majorité des espèces animales possède un *ADNmt* bicaténaire et circulaire, d'une taille de 15 à 20 kb (Boore, 1999 ; Lynch, 2006 ; Avise, 2009). Ce type d'ADN compte 37 gènes codant pour 13 polypeptides de la chaîne respiratoire mitochondriale et de la phosphorylation oxydative (Nad1 à Nad6 et Nad4L ; *COI* à *COIII*, *Cyt-B*, *Atp6* et *Atp8*), 2 sous-unités ribosomales (12S et 16S ARNr) et 22 ARNt (Boore, 1999). L'*ADNmt* constitue un marqueur de référence dans les études de phylogénie, de génétique des populations, de phylogéographie (Avise *et al.*, 1987 ; Simon *et al.*, 1994 ; Sémonlin, 2006).

Il se distingue donc par :

- sa transmission est essentiellement maternelle (Birky, 2001), processus nommé SMI (*strict maternal inheritance*,) via les ovocytes. Ce mode d'hérédité a pour but d'éviter toute incompatibilité entre les produits d'expression des ADN des compartiments nucléaires et mitochondriaux. Alors, dans un organisme, toutes les cellules renferment des copies d'*ADNmt* qui ont la même succession nucléotidique ; c'est l'homoplasmie. En effet, l'hétéroplasmie qui correspond à la coexistence dans les cellules de différents *ADNmt* (ADN normal/ADN muté) peut entraîner une compétition inter-mitochondriale et une modification des interactions nucléo-mitochondriales ; se traduisant par l'apparition de phénotypes indésirables comme une insuffisance de production d'énergie sous forme d'ATP (Birky, 2001 ; Rand, 2001 ; Ballard et Whitlock, 2004). Néanmoins, il existe des cas exceptionnels à cette règle (Zhao *et al.*, 2004 ; Breton *et al.*, 2006 ; Breton et Beaupré, 2007).
- son taux d'évolution qui est de 1 à 10 fois plus rapide que le taux d'évolution global pour l'ADN nucléaire chez les animaux (Avise *et al.*, 1987) ; et le taux de mutations synonymes est un ordre de grandeur plus important que celui de l'ADN nucléaire (Birky *et al.*, 1989). Il possède une taille efficace quatre fois plus petite que celle des marqueurs nucléaires (Avise, 2000 ; 2004). Il est donc bien plus sensible que le génome nucléaire aux effets de la dérive génétique et donc aux changements démographiques des populations. Par conséquent, les populations se différencient en moins de générations sous l'effet de la dérive génétique.
- une absence, en général, de recombinaison méiotique (pas de métissage) et donc ses variations ne sont dues qu'à des mutations cumulées ; ce qui permet d'utiliser les séquences polymorphes comme des haplotypes uniques (Avise, 2000).

Cependant il présente certaines limites, parmi lesquels il y a l'hérédité maternelle qui, selon Ballard et Whitlock (2004), implique que l'information soit limitée lorsque la dispersion est liée au sexe. Ces auteurs mettent aussi en évidence l'absence de recombinaison qui signifie que toute la molécule a la même histoire. Donc la phylogénie construite à partir de l'*ADNmt* ne reflétera pas forcément la

totalité de l'histoire de l'espèce en question. Malgré ces limites, ils possèdent une suprématie dans les études en biologie évolutive (Avisé 2000 ; 2004). Par ailleurs, la neutralité est capitale ; en effet, la plupart des méthodes d'inférence dans la génétique de population reposent sur l'hypothèse que les marqueurs moléculaires sont sélectivement neutres (Avisé, 1994). En outre, un marqueur moléculaire est considéré comme neutre lorsque ses différentes formes alléliques n'affectent pas le phénotype des individus qui les portent.

Nous avons par conséquent choisi 2 gènes du génome mitochondrial que sont :

❖ **le CYTOCHROME B (*Cyt-B*)** qui est considéré comme l'un des gènes les plus utilisés pour la phylogénie, et est probablement le gène mitochondrial le mieux connu en ce qui concerne la structure et la fonction de la protéine produite (Esposti *et al.*, 1993). Chez les insectes sa longueur varie généralement de 1140 à 1150 pb (Takehana *et al.*, 2003 ; Wang *et al.*, 2012 ; Shi *et al.*, 2013). Il se distingue par une forte variabilité intra-spécifique (Avisé, 1994) surtout chez les insectes (Simmons et Weller, 2001) ; et est plus variable que le *COI* ; et donc plus résolutif dans les analyses intra-spécifiques que le *COI* (Ndong *et al.*, 2015).

❖ **le CYTOCHROME OXYDASE I (*COI*)** : sa longueur est d'environ 1540 pb (Spicer, 1995 ; Lunt *et al.*, 1996 ; Wang *et al.*, 2012 ; Shi *et al.*, 2013) chez les insectes ; et est très utilisé en phylogénie et en phylogéographie. Ce marqueur joue un rôle très important dans le processus du barcoding (Decaëns *et al.*, 2013 ; Ndong *et al.*, 2015). En effet, c'est le fragment standard du génome choisi (du moins pour le règne animal) comme marqueur génétique pour la discrimination des espèces et pour révéler la diversité cryptique (Hebert *et al.*, 2004 ; Decaëns *et al.*, 2013). En effet, son taux d'évolution moléculaire est assez rapide pour permettre l'accumulation substantielle de mutations (souvent neutres) et donc la discrimination non seulement des espèces très voisines (espèces jumelles) et de sous-espèces, mais aussi la détection de signal phylogéographique au sein d'une même espèce ou populations écologiquement différenciées (écotypes) (Cox and Hebert, 2001 ; Wares and Cunningham, 2001a ; 2001b ; Decaëns *et al.*, 2013). De ce fait, il est très déterminant dans les études phylogéographiques (Avisé, 2000).

### II.1.2.2 EXTRACTION D'ADN

L'ADN génomique de *S. zeamais* a été extrait grâce à la méthode standard Qiagen (kit Qiagen Dneasy Tissue) en 4 étapes. Pour la phase de la digestion : chaque individu a été broyé dans 180 µl de tampon ATL de digestion contenant des détergents qui entraînent une dissociation des tissus et une individualisation des cellules. Puis s'en suit la lyse cellulaire où 20 µl de protéinase K ont été ajoutés pour dégrader toutes les protéines, après une incubation à 55°C pendant 3h à toute une nuit. Les débris tissulaires ont été éliminés après centrifugation rapide et le surnageant récupéré. A ce dernier, 200 µl du tampon AL de lyse cellulaire ont été ajoutés avant passage immédiat au vortex et incubation pendant 10 minutes à 70°C. Pour les étapes suivantes à savoir l'isolement et la purification

de l'ADN : 200 µl d'éthanol absolu 96-100% ont été ajoutés à la solution récupérée ; ensuite, le mélange a été vortexé et transvasé sur une colonne fourni par le kit puis centrifugé à 13.000 rpm (rotation par minute) pendant 1 minute pour retenir l'ADN au niveau de la membrane de silice de la colonne. En effet, l'ADN chargé négativement se fixe, par interactions ioniques, sur la membrane de silice chargée positivement. Par contre les protéines, les lipides et les polysaccharides sont éliminés. L'ADN fixé sur la colonne est ensuite complètement purifié par une élimination de toutes traces de contaminants. Ce lavage est réalisé par ajouts successifs de 2 tampons de lavage, AW1 et AW2 dont 500 µl chacun, qui passent à travers la membrane par centrifugation à 13.000 rpm pendant respectivement 1 et 3 minutes. La colonne est ensuite placée sur un tube de 1,5 ml établi selon la prédisposition d'une traçabilité de 2 éluions avec le tampon AE auparavant incubé à 70°C pour augmenter le rendement de 15 à 20%. Après une première élution à 30 µl et une seconde à 15 µl, l'ADN est conservé à -20°C pour les études moléculaires. La qualité des extraits ADN a été appréciée après avoir effectué une migration électrophorétique sur gel d'agarose. Pour ce faire, 5 µl d'extrait d'ADN ont été mélangés avec 2 µl de bleu de bromophénol (bleu de charge) et le mélange a été déposé sur un gel d'agarose 1,5% et migré à 100 Voltes pendant 30 minutes.

### II.1.2.3 AMPLIFICATION PAR LA TECHNIQUE DE LA PCR

La Polymerase Chain Reaction (PCR) ou Réaction de Polymérase en Chaîne permet d'amplifier spécifiquement une séquence d'ADN cible. Cette activité se fait grâce à l'ADN polymérase qui a été isolée d'une archéobactérie thermophile (*Thermus aquaticus*, Brock et Freeze, 1969) qui vit dans des sources chaudes et qui peut supporter de hautes températures jusqu'à 100°C. Les amorces utilisées ont été définies par Simon *et al.* (1994) (Tableau II). La PCR a été réalisée dans un thermocycleur de type Eppendorf dans les conditions présentées au Tableau II en multipliant les volumes par le nombre d'échantillons à amplifier. L'amplification de chaque échantillon d'ADN a été réalisée dans un volume réactionnel total de 25 µl contenant 19,3 µl d'eau MilliQ ; 0,5 µl de dNTP ; 2,5 µl du tampon 10X Qiagen ; 1 µl de MgCl<sub>2</sub> ; 0,25 µl de chaque amorce (Ron et Nancy) et 0,2 µl de Taq polymerase ; en plus d'1 µl d'ADN matrice pour le *COI*. Alors que pour le *Cyt-B* avec un échantillon, il y a 25 µl contenant 18,3 µl d'eau MilliQ ; 0,2 µl de Taq polymerase ; 0,5 µl de dNTP ; 2,5 µl du tampon 10X Qiagen ; 1 µl de MgCl<sub>2</sub> et 0,25 µl de chaque amorce (CB1 et CB2) ; en plus d'1 µl d'ADN matrice. Elle a été réalisée suivant les trois étapes présentées dans le Tableau II. Ces étapes constituent un cycle et assurent une multiplication exponentielle du segment d'ADN cible en fonction du nombre de répétitions (cycles). La PCR a été bouclée par un hold à 10°C pour la conservation du produit par le thermocycleur. Pour voir si les amorces se sont accrochées, une migration électrophorétique sur gel d'agarose 1,5% a également été effectuée.



**Tableau II** : Identification des différentes amorces utilisées et programmation de la PCR

Gènes	Nom des amorces	Séquences des amorces	Programme de la PCR
<i>COI</i>	C1-J-1751 (F alias Ron II)	5-GGATCACCTGATATAGCATTCCC-3	<b>(1) Dénaturation initiale</b> : 94°C 3 min ; 35 cycles de dénaturation : 94°C, 1 min ; <b>(2) Hybridation</b> : 47°C, 1 min ; <b>(3) Elongation</b> : 72°C, 1 min ; Elongation finale : 72°C 10 min.
	C1-N-2191 (R alias Nancy)	5-CCCGGTAAAATTTAAAATATAAACTTC-3	
	CB-J-10933 (F)	5-TATGTACTACCATGAGGACAAATATC-3	
<i>Cyt-B</i>	5-	5-TATGTACTACCATGAGGACAAATATC-3	<b>(1) Dénaturation initiale</b> : 93°C, 1min 30s ; 36 cycles de dénaturation : 93°C, 35s ; <b>(2) Hybridation</b> : 47°C, 1min ; <b>(3) Elongation</b> : 72°C, 2 min ; Elongation finale : 72°C, 8 min.
	CB-N-11367 (R)	ATTACACCTCCTAATTTATTAGGAAT-3	

### II.1.2.4 SEQUENÇAGE

Les réactions de séquençage ont été effectuées dans un thermocycleur de type MJ Research PTC-225 Peltier avec les kits ABIPRISM BigDye™ Terminator Cycle. Chaque échantillon a été séquençé en utilisant l'amorce sens. Les fragments fluorescents ont été purifiés avec le protocole BigDye Xterminator purification. Les échantillons ont été suspendus dans de l'eau distillée et soumis à une électrophorèse dans ABI 3730xl sequencer (Applied Biosystems).

## II.1.3 ANALYSES GENETIQUES ET STATISTIQUES

### II.1.3.1 ALIGNEMENTS ET CORRECTIONS DES SEQUENCES

L'alignement de séquences trouve son importance dans la détermination de la similarité ou non des sites (Swofford *et al.*, 1996) et représente une étape fondamentale dans le processus de la reconstruction phylogénétique. En effet, tout le reste du processus phylogénétique repose sur cette étape ; de même que la pertinence des inférences. Les séquences brutes obtenues ont donc été alignées et corrigées avec le logiciel BioEdit ver. 7.2.5 (Hall, 2001) qui utilise l'algorithme Clustal-W (Thompson *et al.*, 1997) qui fait partie des méthodes dites d'alignement global. Ce logiciel est d'une remarquable efficacité tant que le jeu de données ne présente pas de complexité particulière, séquences très éloignées phylogénétiquement c'est-à-dire longs indels, régions hypervariables, portions non-homologues (Mahé, 2009) ; et donc très pertinente dans l'étude de la diversité intraspécifique.

### II.1.3.2 DIVERSITE GENETIQUES DES POPULATIONS

La genèse et l'évolution de cette diversité génétique est conditionnée par les forces évolutives (mutations, sélection naturelle, dérive génétique et migrations) ou forces écologiques (système de reproduction, sex-ratio, démographie, capacité de dispersion, distance géographique, ressources naturelles) ou par les facteurs pluvio-climatiques et anthropiques. Dès lors, nos analyses ont été faites avec 3 matrices : les séquences du *Cyt-B* et du *COI* séparément en plus d'un troisième par une approche de concaténation des deux gènes. Différents paramètres ont été déterminés afin d'inférer sur la diversité et la structure génétiques des populations étudiées.

#### ➤ Indices standards de la diversité génétique tels que :

- le nombre de sites polymorphes (S) ;

- le nombre de sites informatifs en parcimonie ( $S_{IP}$ ) : un site est informatif uniquement s'il y a au moins 2 types de nucléotides présents dans ce site et si chacun est représenté dans au moins 2 séquences comparées ;
- le nombre de sites singletons ( $S_S$ ) ;
- le nombre de différences nucléotidiques ( $K$ ) ;
- le taux transitions/transversions ( $R$ ) ;
- et les fréquences nucléotidiques sont déterminées avec le logiciel MEGA 7 ver. 7.0.18 (Kumar *et al.*, 2016) ;
- $\pi$  : la diversité nucléotidique définissant la probabilité que deux sites nucléotidiques homologues choisis au hasard soient différents (Nei, 1987).
- $H$  : les indices des haplotypes (Nei, 1987) tels que :
  - $H_T$  : le nombre total d'haplotypes ;
  - $H_P$  : le nombre d'haplotypes privés/endémiques ; ce sont des haplotypes spécifiques à une population ;
  - $H_d$  : la diversité haplotypique correspondant à la probabilité que deux allèles ou haplotypes tirés au hasard dans un échantillon soient différents (Nei, 1987). C'est donc une mesure des fréquences et du nombre d'haplotypes parmi les individus avec une valeur comprise entre 0 et 1 (Grant et Bowen, 1998).

Ces derniers paramètres ont été déterminés par le logiciel DNAsp ver. 5.10.01 (Librado and Rozas, 2009 ; Librado *et al.*, 2010).

En outre, le nombre de sites polymorphiques différenciant les séquences est important, plus la diversité nucléotidique sera élevée avec une valeur de 0 (pas de divergence) à plus de 10% (divergence profonde) (Grant et Bowen, 1998 ; Orémus *et al.*, 2011). Il est une mesure de la diversité génétique qui, en plus, intègre un caractère évolutif car 2 séquences présentant une seule paire de base qui les différencie auront un ancêtre commun plus proche que deux séquences avec cinq paires de bases différentes contrairement au  $H_d$ .

Par ailleurs, un haplotype désigne un ensemble d'individus qui ont une succession nucléotidique identique pour une séquence d'ADN donnée. Les haplogroupes ou lignées sont ceux qui partagent des mutations qui leurs sont spécifiques. Autrement dit, c'est un *cluster* monophylétique d'haplotypes.

Des tests de T-Student et de Chi 2 ont été utilisés avec le Statview ver. 5.0 (SAS Institute Inc., 1998) pour comparer les moyennes et les proportions des différents indices standards de la diversité génétique ( $K$ ,  $H_d$  et  $\pi$ ) entre populations de la zone sahélo-soudanienne et de la région guinéo-congolaise.

➤ **Détection de sélection naturelle au sein des populations** : Comprendre l'importance relative des processus sélectifs et démographiques soulève un intérêt considérable, mais démêler les effets des 2 processus pose un sérieux défi. En effet, la sélection et la démographie peuvent entraîner des schémas similaires de diversité nucléotidique car les signatures de sélection peuvent également être produites par des forces démographiques telles que l'expansion rapide ou le goulot d'étranglement d'une population (Hudson *et al.*, 1987 ; Tajima, 1989a, b ; McDonald et Kreitman, 1991). Utiliser des modèles qui tiennent compte des processus démographiques est essentiel pour identifier les régions du génome sous sélection.

Le test du  $K_A/K_S$  ou  $d_N/d_S$  (Li *et al.*, 1985 ; Nei and Gojobori, 1986) a été utilisé. Ce test compare le rapport du taux de substitutions non-synonymes ( $d_N$ ) et du taux de substitutions synonymes ( $d_S$ ) par site polymorphe entre populations ou espèces proches pour la séquence étudiée (Kimura, 1980 ; Li *et al.*, 1985 ; Yang and Nielsen, 2000). A l'équilibre (en l'absence de sélection),  $d_N$  est (environ) égal à  $d_S$ . Par contre, Si  $d_N$  est supérieur à  $d_S$ , la sélection est divergente. Au contraire un  $d_N$  inférieur à  $d_S$  renseigne d'une sélection purificatrice. Des tests de sélection ont aussi été effectués à l'aide du logiciel HyPhy ver. 2.2.4 (Kosakovsky-Pond *et al.*, 2005) où le modèle de codon MG94xREV (Kosakovsky-Pond et Muse, 2005) est utilisé pour estimer, pour chaque marqueur, le ratio global des synonymes et non synonymes ( $d_N/d_S$ ).

### II.1.3.3 STRUCTURE GENETIQUE ET RELATIONS PHYLOGEOGRAPHIQUES DES POPULATIONS

Les modalités de reproduction et de dispersion, souvent difficiles à évaluer par des moyens directs, peuvent être obtenues à partir des empreintes qu'elles laissent sur la distribution de la diversité génétique à différentes échelles spatiales (Avice, 1994). Par exemple, une diminution de l'efficacité de la dispersion avec la distance géographique, associée à la dérive génétique locale, peut créer un modèle d'isolement par la distance.

#### II.1.3.3.1 DIFFERENCIATION ET FLUX GENETIQUES INTER-POPULATIONS

Le  $F_{ST}$  (indice de différenciation génétique inter populations) permet de mesurer, indirectement, la différenciation et/ou le flux génique des populations à partir de la diversité (polymorphisme) génétique et peut être interprété comme la part de variance des fréquences haplotypiques (alléliques) due aux différences entre populations. Pour les séquences d'ADN, le calcul du  $F_{ST}$  est basé sur les distances génétiques entre haplotypes c'est-à-dire à partir des fréquences haplotypiques or fréquences des sites polymorphiques, en traitant chaque site comme un locus distinct (Hudson *et al.*, 1992b). Il renseigne sur la différenciation et l'effet de subdivision des populations. Il représente le ratio entre la variabilité allélique (haplotypique) inter-populations sur la variabilité allélique (diversité haplotypique) totale (Weir et Cockerham, 1984) selon le modèle en île c'est-à-dire une population globale non soumise à la sélection et subdivisée en sous-populations de tailles identiques échangeant un nombre de gènes identique (Wright, 1951 ; 1969) ; et à l'équilibre migration/dérive (Weir et

Cockerham, 1984 ; Nei, 1987). D'après Hartl et Clark (1989) qui tiennent compte du fait que l'influence de la mutation soit généralement considéré comme négligeable, à l'exception des marqueurs microsatellites qui présentent de forts taux de mutations (Grivet, 2002 ; Duminil, 2006), en considérant un grand nombre de populations avec un taux de migration  $m$  beaucoup plus important que le taux de mutation  $\mu$  :

$$F_{ST} = 1 / (4N_e m + 1) \text{ d'où } Nm = (1 - F_{ST}) / 4 * F_{ST}$$

Où  $N_e$  est la taille efficace de chaque population et  $m$  le taux de migration et donc  $Nm$  correspond au nombre de migrants par génération. Plus la valeur de  $Nm$  est supérieure à 1, plus l'échange de gènes est important. Les valeurs du  $F_{ST}$  sont comprises entre 0 et 1.  $F_{ST}=0$ , quand il n'y a pas de différence entre les fréquences haplotypiques des sous-populations qui sont donc panmictiques ; donc absence de structuration des populations ; et  $F_{ST}=1$  quand tous les échantillons sont totalement isolés les uns des autres.

#### II.1.3.3.2 DISTANCES GENETIQUE ET ISOLEMENT PAR LA DISTANCE

La distance génétique ( $D_s$ ) entre paires de populations définies selon les pays est calculée sous MEGA, en utilisant le modèle Kimura-2-P. Nous avons aussi cherché à tester l'existence d'une relation entre la différenciation des populations et leur distribution géographique ou IBD (Isolement By Distance, Wright, 1943) en réalisant le test de Mantel (Mantel, 1967) à l'aide du logiciel R ver. 3.4.3 (R Core Team, 2017) avec les packages *car* (Fox and Weisberg, 2011), *nlme* (Pinheiro *et al.*, 2017) et *ggplot2* (Wickham, 2016).

La structure des populations est aussi mise en exergue par l'analyse des matrices de différenciation génétique ( $F_{ST}$ ) et de divergence ( $D_s$ ) par l'approche de la classification ascendante hiérarchique (CAH) avec le logiciel statistique R ver. 3.2.3 (R Core Team, 2015) en utilisant la fonction *hclust* () du package *stats* avec la méthode « ward.D2 » et le package *pvclust* (Suzuki et Shimodaira, 2014) pour, en plus, évaluer les *p-value* de tous les *clusters*. Notons qu'ici, le but de cette analyse est de quantifier les affinités entre les populations surtout mais elle ne se focalise pas particulièrement sur la détermination du nombre de *clusters*. L'analyse avec le *pvclust* fournit 2 tests de *p-values*. En effet, en plus de la BP *P-value* (**B**ootstrap **P**robability, Probabilité de Bootstrap), il y a l'UA *p-value* (*Approximately Unbiased*, pour approximativement non biaisé) effectuée par un test très spécifique en phylogénétique (Shimodaira, 2002). L'UA *p-value* est faite par une technique de ré-échantillonnage multi-échelle de bootstrap. C'est est une meilleure approximation de valeurs de *p-value* non-biaisées que la BP *p-value* calculée par un ré-échantillonnage standard de bootstrap.

### II.1.3.3.3 DISTRIBUTION DE LA VARIANCE GENETIQUE ET REPARTITION SPATIALE

#### II.1.3.3.3.1 SOURCE DE LA VARIANCE MOLECULAIRE

La structuration génétique des populations est appréhendée par une analyse hiérarchique, AMOVA (Analysis MOléculaire of Variance) (Excoffier *et al.*, 1992) qui va chercher à identifier la source de la variance moléculaire. Cette procédure estime la proportion de variation génétique totale attribuable à différents niveaux hiérarchiques basés sur des regroupements *à priori* (tels que les zones agro-écologiques ou les pays échantillonnés qui peuvent représenter la population) et sur le degré de divergence génétique entre les regroupements en fonction d'un ou plusieurs facteurs donnés. Les tests de significativité sont effectués après 1023 permutations.

#### II.1.3.3.3.2 CLUSTERING PAR UNE ANALYSE BAYESIENNE DE LA STRUCTURE GENETIQUE

L'analyse bayésienne de la structure de la population ou BAPS (Bayesian Analysis Population Structure) ver. 6.0 (Corander and Tang, 2007) qui est un logiciel d'assignation par le bayésien, a permis de visualiser la structuration génétique spatiale des individus. Dans le cas des séquences d'ADN, le logiciel BAPS traite des fréquences nucléotidiques et du nombre de variants génétiquement divergents dans la population en tant que variables aléatoires. Cependant, les analyses et les comparaisons de modèle peuvent également être effectuées en utilisant un nombre fixé de variants génétiquement divergents ou des structures de population pré-spécifiées. Cette analyse bayésienne assigne des individus à des *clusters* génétiques soit en les considérant comme immigrants (analyse par mixture) ou comme des descendants d'immigrants (analyse par admixture) (Excoffier et Heckel, 2006). BAPS détermine les principaux ensembles populationnels en regroupant les populations qui échangent des migrants à un taux élevé.

#### II.1.3.3.3.3 DETECTION DE SIGNAL PHYLOGEOGRAPHIQUE

##### II.1.3.3.3.3.1 INDICES DE DIFFERENCIATION

Parmi les indices de différenciation génétique entre populations, analogues au  $F_{ST}$  de Wright (1969), il y a le  $G_{ST}$  (Nei, 1973, 1977 et 1987) avec les corrections de Nei et Chesser (1983) et le  $N_{ST}$  (Lynch et Crease, 1990 ; Hudson *et al.*, 1992b ; Pons et Petit 1996). Le  $G_{ST}$  est une mesure de différenciation entre allèles ou haplotypes en utilisant seulement la fréquence des haplotypes. Alors que le  $N_{ST}$  est basé en plus de la proportion des sites polymorphes qui distinguent chaque paire d'haplotypes ou allèles autrement dit sur les distances génétiques entre les allèles (Nei et Li, 1979 ; Lynch et Crease, 1990 ; Hudson *et al.*, 1992b ; Pons et Petit, 1996). Ainsi, le  $N_{ST}$  intègre une dimension phylogéographique (Duminil, 2006). Il devrait, *à priori*, être supérieur au  $G_{ST}$  car les allèles les plus apparentés devraient se retrouver préférentiellement dans les mêmes populations. Une comparaison de ces indices ( $G_{ST}$  et  $N_{ST}$ ) permettent alors de décrire la répartition de la diversité génétique entre et au sein des populations ; d'inférer sur la dynamique historique de population et de

déterminer la structuration génétique et phylogéographique. Pour tester la structure génétique, l'hypothèse stipule que les populations ne sont pas différenciées génétiquement ( $G_{ST}=0$  et  $N_{ST} = 0$ ), en considérant différentes entités de populations. Alors que dans la recherche d'un signal phylogéographique, l'hypothèse nulle est une absence de structure phylogéographique avec  $G_{ST}=N_{ST}$  (permut). D'après Burban *et al.* (1999), la réalisation de ces tests est faite à travers une série de permutations de l'identité des haplotypes suivie de la comparaison des valeurs de  $N_{ST}$  obtenues après permutations avec la valeur observée. Il faut souligner aussi qu'après permutation, les valeurs de  $N_{ST}$  convergent vers le  $G_{ST}$ . Si la valeur de  $N_{ST}$  observée est supérieure à la borne supérieure de l'intervalle de confiance de  $N_{ST}$  (permut), l'hypothèse nulle est rejetée. Un test significatif ( $N_{ST} > G_{ST}$ ) signifie que des haplotypes distincts se retrouvant au sein des populations sont en moyenne plus apparentés que les haplotypes distincts issus de différentes populations (Pons et Petit, 1996). La phylogénie des haplotypes et leur distribution géographique ne sont donc pas indépendantes, et on parle de structure phylogéographique (Grivet, 2002). Il existe donc une correspondance (Pons et Petit, 1996 ; Grivet, 2002) entre la phylogénie des haplotypes et leur distribution géographique (Figure 7) :

✓ s'il y a correspondance des haplotypes phylogénétiquement, les plus proches se retrouvent plus souvent en mélange dans les mêmes populations. La différenciation mesurée en prenant en compte les similarités entre haplotypes ( $N_{ST}$ ) est supérieure à la différenciation basée uniquement sur la fréquence des haplotypes ( $G_{ST}$ ) ;

✓ si les haplotypes sont reliés de manière équivalente, alors  $N_{ST} = G_{ST}$  ;

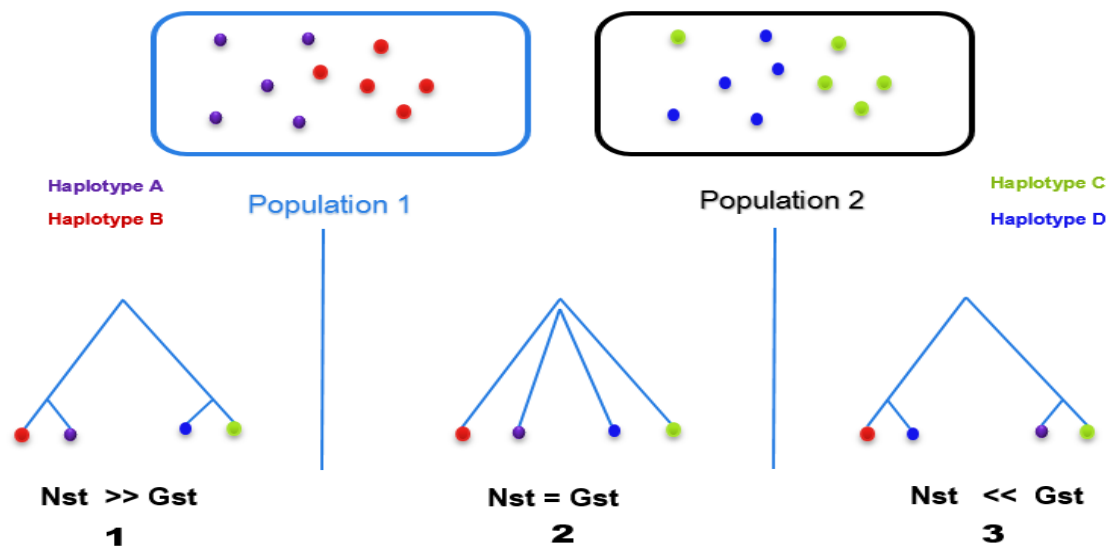
✓ enfin, quand les haplotypes les plus fortement reliés sont plus souvent dans des populations différentes, alors  $N_{ST} < G_{ST}$ .

En effet, la correspondance entre proximité géographique et proximité phylogénétique définit la structure phylogéographique puisque les populations géographiquement proches sont parfois aussi génétiquement proches et appartiennent aux mêmes lignées phylogénétiques (Avisé *et al.*, 1987).

Ces paramètres et le test d'AMOVA sont déterminés en utilisant les logiciels ARLEQUIN ver. 3.5.2.2 (Excoffier et Lischer, 2010) et PERMUT ver. 2.0 (Pons et Petit, 1996) respectivement.

Par ailleurs, Nei (1973) a proposé des indices basés sur la fréquence des allèles (haplotypes) qui décrivent le niveau et l'organisation de la diversité au sein et entre populations d'une même espèce. Ce sont la diversité génétique totale ( $H_T$ ), la diversité génétique intra-population ( $H_S$ ) et la diversité génétique inter-population ( $D_{ST}$ ), incluant les comparaisons des populations entre-elles. Les hétérozygoties  $H_T$  et  $H_S$  sont calculées sous l'hypothèse de l'équilibre d'Hardy-Weinberg et ne dépendent donc que des fréquences haplotypiques (alléliques) observées. Ces paramètres sont liés par les relations suivantes :

$$H_T = H_S + D_{ST} \quad \text{and} \quad G_{ST} = D_{ST} / H_T = 1 - H_S / H_T$$



**Figure 7 :** Correspondance entre la phylogénie des haplotypes et leur distribution géographique d'après Pons et Petit (1996).

### II.1.3.3.3.2 ANALYSE SPATIALE DE LA VARIANCE MOLECULAIRE

Le SAMOVA (Spatial Analysis of MOlecular Variance) ou Analyse spatialisée de la variance moléculaire ver. 2.0 (Dupanloup *et al.*, 2002) ou AMOVA spatialisée est une analyse à *posteriori* qui permet de définir des groupes de populations géographiquement homogènes et entre lesquels la différenciation génétique est maximale. Cette analyse définit des groupes en fonction de leurs similarités génétiques et cherche à détecter des barrières génétiques. Ainsi, il peut conduire à la mise en évidence de barrières génétiques entre ces groupes. La méthode est basée sur une procédure qui vise à maximiser la variance génétique totale associée à la différence entre groupes de populations. L'existence de  $n$  groupes est testée dans le jeu de données spatialisé, *i.e.* où chaque individu est renseigné par ses coordonnées géographiques. La variation de la différenciation intergroupe ( $F_{CT}$ ) en fonction de  $K$  (nombre de groupes génétiques) est explorée pour inférer le nombre « correct » de groupes avec l'idée que  $F_{CT}$  est maximale pour le bon  $K$  ( $K$  variant de 2 à 5).

**Remarques :** Nous avons utilisé ces 3 méthodes différentes mais complémentaires dans la détermination de la structure génétique et phylogéographique. En effet :

- les tests d'AMOVA permettent de caractériser la partition de la variance génétique sous des hypothèses biogéographiques et de voir quel regroupement expliquerait une meilleure séparation entre les regroupements définis à *priori* (structure des populations) ;
- les analyses de SAMOVA sont utilisées pour tester une séparation génétique de populations géographiquement isolées c'est-à-dire s'il existe un ensemble d'haplogroupes séparés géographiquement ; elles permettent de mettre en évidence l'existence de barrières génétiques.
- les analyses BAPS (Bayesian Analysis of Population Structure) permettent de réaliser un *clustering* des haplotypes échantillonnés dans des *clusters* monophylétiques (haplogroupes). Elles donnent alors le nombre de *clusters* ( $K$ ) existant dans un jeu de données et permettent de déterminer le meilleur pattern de séparation.

### II.1.3.3.4 STRUCTURE DES POPULATIONS ET PROCESSUS HISTORIQUES

La phylogéographie permet de mettre en exergue l'impact des processus historiques et démographiques dans la distribution géographique des polymorphismes génétiques (neutres). Ainsi, à partir du réseau d'haplotypes, une analyse de clades imbriqués/emboîtés (NCPA, Nested-Clade Phylogeographic Analysis) a été utilisée afin de tester l'association géographique entre les haplotypes (ou des clades imbriqués) et la distribution géographique de Templeton *et al.* (1995). Le NCPA (Templeton *et al.*, 1987 ; Templeton et Sing, 1993 ; Crandall, 1996) permet ainsi de différencier la structure des populations (flux de gènes, ...) de l'histoire des populations (fragmentation, colonisation, ...) (Posada *et al.*, 2000). Selon Donnelly et Tavaré (1986), les clades externes devraient être plus récents que les clades internes. Il est réalisé à l'aide du programme Geodis ver. 2.5 (Posada *et al.*, 2000). L'hypothèse nulle stipule que les haplotypes ou clades d'haplotypes sont distribués géographiquement au hasard (Templeton *et al.*, 1995). Malgré de nombreuses critiques dont les principales sont basées sur l'absence d'évaluation statistique de l'incertitude et liées aux inconsistances des inférences des histoires phylogéographiques complexes, particulièrement celles impliquant des taux de migration élevée (Knowles et Maddison, 2002 ; Panchal et Beaumont, 2007 ; Petit, 2007 ; Beaumont and Panchal 2008). Cependant, la pertinence de cette méthode a été défendue par l'un des auteurs initiateurs à savoir Templeton (2008 ; 2010a ; 2010b). Nous avons utilisé la clé d'inférence Templeton (2011) afin de déterminer les inférences du NCPA. Trois critères régissent le clustering des haplotypes jusqu'à ce que le dernier niveau d'imbrication englobe l'ensemble de l'arbre : le critère de fréquence dans lequel les haplotypes sont censés être connectés à des haplotypes fréquents plutôt qu'aux haplotypes rares ; le critère de topologie dans lequel les haplotypes sont censés être connectés à l'intérieur plutôt qu'aux haplotypes externes ; et le critère géographique, dans lequel les nouveaux haplotypes devraient se retrouver au même endroit que l'haplotype ancestral (Templeton et Sing, 1993).

### II.1.3.4 EVOLUTION DEMOGRAPHIQUE

L'inférence de l'évolution démographique des populations a été faite grâce à 3 approches : les diversités haplotypique et nucléotidique (Grant et Bowen, 1998 ; Avise, 2000) dont l'avantage est que le résultat est indépendant de la longueur du fragment et de la taille de l'échantillon (Nei et Li 1979 ; Nei 1987) ; les tests démogénétiques du  $F_S$  de Fu (Fu, 1997), du  $D$  de Tajima (Tajima, 1989a, b) et du  $R^2$  (Ramos-Onsín et Rozas, 2002) ; et aussi les indices démographiques du mismatch distribution SSD et  $R_g$  (Harpending, 1994) combinés aux courbes de mismatch distribution (Rogers et Harpending, 1992) et enfin le coefficient d'expansion  $S/K$  (Peck et Congdon, 2004).

Les tests basés sur le spectre de fréquence allélique déterminent si le spectre de fréquence des mutations est conforme aux attentes du modèle standard de neutralité. Ce sont :

- **TESTS DEMOGENETIQUES** : Ces tests sont plus exactement des tests de neutralité sélective et



d'équilibre des populations. Ce sont :

- D** de Tajima (Tajima, 1989a, b) : c'est la différence du nombre total de sites polymorphes observés (**S**) et la moyenne du nombre de différences observées entre paires de séquences (**K**) ;
- F<sub>s</sub>** de Fu (Fu, 1997) : il compare le nombre moyen de différences observées entre paires de séquences (**K**) avec le nombre d'haplotypes (**H**) dans une population ;
- R<sup>2</sup>** (Ramos-Onsin et Rozas, 2002) : C'est une statistique complémentaire qui est basée sur les différences entre le nombre de mutations de type singletons (**S<sub>s</sub>**) et le nombre moyen de différences observées entre paires de séquences (**K**).

Pour des valeurs de **D** de Tajima et de **F<sub>s</sub>** de Fu :

- (1) proches de zéro, il y a une stabilité démographique et une absence de sélection ;
- (2) significativement négatives, cela suggère une expansion démographique (récente) avec un excès d'allèles rares peu divergents (une ou deux bases de divergence) et correspond à une sélection négative ou purificatrice (disparition des mutations délétères) ou un balayage sélectif (sélection positive : apparition et propagation d'un mutant avantageux dans la population) qui peut être récent si excès de singletons ;
- (3) significativement positives, il évoque une réduction drastique de l'effectif de la population (goulot d'étranglement) avec un déficit de variants rares mais beaucoup d'hétérozygotie ; il concorde avec une sélection balancée (maintien d'allèles avantageux ancestraux), mais aussi elles indiquent la présence d'une forte structure de population.

Fu (1997) a constaté que la statistique **F<sub>s</sub>** est très sensible à l'expansion démographique de la population, ce qui conduit généralement à de grandes valeurs négatives du **F<sub>s</sub>**.

Il faut noter que le **F<sub>s</sub>** de Fu et le **R<sup>2</sup>** sont les tests les plus puissants pour détecter la croissance des populations. Cependant, le **R<sup>2</sup>** est meilleur pour les populations de tailles réduites et alors que le **F<sub>s</sub>** est plus adapté pour les populations de grandes tailles (Ramos-Onsins et Rozas, 2002).

Concernant la valeur de **R<sup>2</sup>**, quand elle est faiblement positive et significative, il y a expansion démographique récente ; et si **R<sup>2</sup>** est significativement négatif alors la population est constante.

La signification des statistiques de **D**, **F<sub>s</sub>** et **R<sup>2</sup>** ont été testées en générant des échantillons aléatoires sous l'hypothèse de la neutralité sélective et de l'équilibre de la population, à l'aide d'un algorithme de simulation par coalescence adapté de Hudson (1990) grâce aux logiciels ARLEQUIN (statistiques de **D** et **F<sub>s</sub>**) et DNAsp (statistique de **R<sup>2</sup>**).

- **ANALYSE DE "MISMATCH DISTRIBUTION"** (Hudson et Slatkin, 1991 ; Rogers et Harpending, 1992) ou distribution des disparités (*i.e.* distribution des différences entre séquences prises deux à deux) repose sur un principe assez simple : plus deux séquences ont divergé depuis longtemps, plus la différence de nombre de sites polymorphes entre les deux sera grande. Cette analyse a été réalisée à l'aide du logiciel Dnasp ; sous l'hypothèse nulle d'une expansion démographique. Ainsi, une distribution unimodale est le signal d'une population en expansion démographique ; alors qu'une distribution

multimodale infère une stabilité démographique (Rogers et Harpending, 1992 ; Harpending *et al.*, 1993).

- **INDICES DEMOGRAPHIQUES** : SSD (Sum of Squared Deviation) et Rg (Raggedness statistic, Harpending *et al.*, 1993 ; Harpending, 1994) sont estimés entre les distributions observées et attendues afin de tester la validité de l'hypothèse d'expansion de l'analyse "mismatch distribution". Leur significativité est testée et si :

- les *P-values* de SSD et Rg sont significatives : il y a écart entre les valeurs observées et celles simulées, la population est en stabilité démographique ou en début d'expansion démographique ;

- les *P-values* de SSD et Rg ne sont pas significatives : pas d'écart entre les valeurs observées et celles simulées, alors la population est en expansion démographique.

Par ailleurs, nous avons aussi utilisé le ratio S/K (S : nombre de sites polymorphes et K : nombre moyen de différences nucléotidique entre paire de séquences) ou coefficient d'expansion (Peck et Congdon, 2004 ; Von Haeseler *et al.*, 1996). Dès lors, des populations avec de forts ratios suggèrent l'expansion démographique au cours du temps alors que de faibles ratios indiquent une démographie constante à long terme (Peck & Congdon 2004).

### II.1.3.5 ANALYSES DES RELATIONS PHYLOGENETIQUES

L'analyse phylogénétique permet de relier les relations de parenté existant entre les différentes lignées de *S. zeamais* sous l'hypothèse que la phylogénie des gènes reflète celle des organismes.

#### II.1.3.5.1 RESEAUX DES HAPLOTYPES

Les réseaux d'haplotypes représentent une application de la méthode de liens médians pour montrer les relations phylogénétiques entre les différents haplotypes. Un réseau minimum d'haplotype se caractérise par des nœuds (cercles) et des branches (liens) qui relient les nœuds. Chaque nœud correspond à un haplotype dont la taille est proportionnelle à la fréquence de l'haplotype dans le jeu de données. Ils sont soit des séquences de l'ensemble de données (haplotypes), ou des vecteurs médians. Les liens sont les différences de caractères. Un vecteur médian est une séquence hypothétique (souvent héréditaire) qui est nécessaire pour connecter des séquences existantes dans le réseau avec une parcimonie maximale. Liens et vecteurs médians constituent respectivement le nombre de pas mutationnels et la barrière du flux génétique qui séparent les différents nœuds. Les différents réseaux d'haplotypes sont construits avec NETWORK ver. 5.0.0.0 en utilisant la méthode Median-Joining (Bandelt *et al.*, 1999) pour montrer les relations phylogénétiques entre les différents haplotypes. La construction de ces réseaux est basée sur la théorie de la coalescence (Kingman, 2000) qui consiste en une approche rétrospective qui décrit mathématiquement le processus de fusion binaire de tous les lignages généalogiques d'un échantillon de gènes jusqu'à leur plus proche ancêtre commun.

### II.1.3.5.2 RECONSTRUCTIONS PHYLOGENETIQUES

Un préalable à la reconstruction phylogénétique est la détection de la présence d'un signal phylogénétique des séquences par estimation du taux de saturation des sites dans les différentes matrices en faisant un test de saturation des sites de substitution basé sur la méthode de parcimonie de Steel *et al.* (1993) et celle de la méthode de l'information basée sur l'entropie (Xia *et al.*, 2003 ; Xia et Lemey, 2009) grâce au programme DAMBE ver. 5.6.8 (Xia, 2013). La saturation de substitution diminue l'information phylogénétique contenue dans les séquences, et biaise l'analyse phylogénétique impliquant les branches profondes, telles une majorité de groupes d'arthropode (Xia *et al.*, 2003). Ainsi, il y a présence d'un signal phylogénétique si l'indice de saturation de substitution (Iss) est plus petit que celui de sa valeur critique (Iss.c). Le test de Steel (1993) a été réalisé par l'échantillonnage de tous les quartets possibles dans une matrice et a été évalué par une utilisation de la moyenne  $\phi$ -value de chaque séquence (Xia et Lemey, 2009). Les séquences avec une  $\phi$ -value plus petite que 0,04 peuvent être interprétées comme présentant une absence de signal phylogénétique.

D'une autre part, nous avons vérifié entre les 3 matrices (*Cyt-B*, *COI* et Séquences concaténées), laquelle explique le mieux l'histoire évolutive des populations. En effet, l'approche par concaténation des gènes stipule l'hypothèse selon laquelle les Séquences concaténées possèdent les mêmes propriétés évolutives, en particulier les vitesses d'évolution aussi bien aux niveaux des différents sites que dans les différentes branches de l'arbre. Or la vitesse d'évolution du *Cyt-B* est plus rapide que celle du *COI* (Ndong *et al.*, 2015). Selon Yang (1996), il est, *a priori*, préférablement d'avoir des paramètres de modèle d'évolution différents pour des gènes différents. Ainsi, nous avons adopté la méthode préconisée par Brochier *et al.* (2002) en comparant la vraisemblance des 3 matrices (plus exactement la vraisemblance des Séquences concaténées et la somme des vraisemblances de chaque gènes) sans ou avec la loi gamma qui modélise la vitesse d'évolution des sites afin de déterminer la meilleure phylogénie présentée par les 3 matrices.

En plus, les reconstructions ont été enracinées avec une séquence de *Sitophilus oryzae* (SO). La crédibilité des relations phylogénétiques mises en exergue par les arbres dépend essentiellement de la séquence *out-groupe*. Il existe plusieurs critères afin de bien choisir le bon *out-groupe*. Lyons-Weiler *et al.* (1998) stipulent que le signal phylogénétique présent dans une analyse avec un arbre enraciné, donc comprenant l'*out-groupe*, doit être supérieur à celui trouvé avec l'analyse de l'arbre non enraciné. D'autre part, les différences généralement importantes entre les valeurs de bootstrap non-paramétriques (BP<sub>ML</sub>) et celles de probabilités postérieures bayésiennes (PP) et de probabilités postérieures bayésiennes obtenues après re-échantillonnage des données par bootstrap non-paramétrique (BP<sub>Bay</sub>) mènent souvent à des résultats contradictoires difficilement interprétables (Douady *et al.*, 2003 ; Suzuki *et al.*, 2002). Ainsi, après ces analyses phylogénétiques, les différences importantes observées pour un même nœud entre les indices de confiance bayésiens et les valeurs de bootstrap non-paramétrique et les différences topologiques ont été évalués en utilisant le *likelihood-*

*maps* (Strimmer et Von Haeseler, 1997) implanté dans le logiciel Tree-Puzzle ver. 5.2 (Schmidt *et al.*, 2002).

La reconstruction des arbres phylogénétiques du Neighbor-Joining (Saitou et Nei, 1987) sont basée sur un calcul de matrice de distances génétiques qui a pour objectif de minimiser la longueur totale de l'arbre et qui utilise le modèle Kimura-2-P (= 2 paramètres) qui postule que les transitions et les transversions ne sont pas équiprobables lors de l'évolution des séquences, notamment pour l'ADN mitochondrial où les transitions peuvent représenter 90% des mutations (Kimura, 1980). Cette méthode, en plus du maximum de vraisemblance (Cavalli-Sforza et Edwards, 1967 ; Felsenstein, 1981) ont été construits avec les logiciel PHYLIP ver. 3.6 (Felsenstein, 2005) et PhyML ver. 3.0 (Guindon et Gascuel, 2003) respectivement. La robustesse des nœuds a été évaluée par le *bootstrap* qui est une technique statistique utilisée pour évaluer la robustesse d'une reconstruction phylogénétique (Felsenstein, 1985). Un nœud avec un soutien égal ou supérieur à 70 % est considéré comme significatif (Hillis et Bull, 1993 ; Soltis et Soltis, 2003). Mahé (2009) considère qu'un seuil de 1000 réplicats est un bon compromis entre précision et temps de calcul. Par ailleurs, l'inférence bayésienne est estimée à l'aide du programme MrBAYES ver. 3.2.5 (Huelsenbeck et Ronquist, 2001). Cette dernière méthode et le maximum de vraisemblance sont qualifiés de probabilistes et ce sont elles qui atteignent les meilleurs résultats sur les jeux de tests (Mahé, 2009). De plus, le choix du modèle évolutif le plus adapté est primordial. En effet, la qualité de l'ajustement d'un modèle statistique désigne le degré d'ajustement du modèle aux données observées et constitue une importance cruciale pour la reconstruction phylogénétique (Cunningham *et al.*, 1998 ; Posada et Crandall, 2001). Toutes les méthodes phylogénétiques se fondent sur des hypothèses au sujet du processus des substitutions de l'ADN ou des acides aminés. La confiance dans une analyse phylogénétique dépend de la qualité de l'ajustement, c'est-à-dire, comment le modèle pour décrire les données est approprié (Schmidt et Von Haeseler, 2003). Nous avons utilisé l'AICc (*Akaike Information Criterion* corrected) (Hurvich et Tsai, 1989) qui est la version corrigée de l'AIC (*Akaike Information Criterion* Akaike ; Akaike, 1974) concernant la taille de l'échantillon. Le modèle avec l'AIC le plus faible est sélectionné. En outre, l'AIC est considéré par Posada et Buckley (2004) comme présentant la stratégie qui s'avère la plus pertinente. L'AICc est donc utilisé selon les recommandations de Burnham et Anderson (2002) : lorsque n (la longueur de l'alignement) est petit et/ou k (le nombre de paramètres k du modèle le plus complexe) est grand ; c'est-à-dire lorsque le ratio n/k est petit (c'est-à-dire  $\leq 40$ ) (Burnham et Anderson, 2004 ; Symonds et Moussalli, 2011). En effet, utiliser l'AIC plutôt que de l'AICc quand n n'est pas beaucoup plus grand que k augmente la probabilité de sélectionner un modèles avec un trop grand nombre de paramètres, c'est-à-dire de sur-ajuster. De ce fait, nous avons combiné l'AICc et le likelihood ratio test (LRT) (Swofford *et al.*, 1996 ; Huelsenbeck et Crandall, 1997 ; Huelsenbeck et Rannala, 1997) afin d'inférer sur le meilleur modèle évolutif de notre jeu de données. Ces indices sont déterminés avec le logiciel Jmodeltest ver. 2.1.10

(Darriba *et al.*, 2012) pour l'arbre du maximum de vraisemblance et pour l'inférence bayésienne. Le modèle évolutif sélectionné peut suivre une distribution gamma qui décrit le taux de substitution qui change souvent d'un site à un autre à l'intérieur d'une séquence et la forme de cette distribution est déterminée par une valeur connue sous le nom de « gamma parameter » ( $\Gamma$ ) ou « shape parameter » alors que d'autres sites peuvent rester invariables (+I). Des valeurs de gamma  $< 1$  décrivent une forte hétérogénéité, alors que de grandes valeurs décrivent une homogénéité (Gu *et al.*, 1995 ; Page et Holmes, 1998).

Pour l'inférence bayésienne, la distribution des probabilités postérieures des arbres a été estimée par 4 chaînes MCMC (dont trois ont été « chauffées » de façon graduelle et une froide). 5.000.000 générations ont été réalisées pour chacune des chaînes en échantillonnant les différents paramètres toutes les 1.000 générations. Le degré de convergence des chaînes peut être vérifié en examinant l'évolution de la fonction de vraisemblance pendant le parcours de la chaîne « froide » afin de déterminer la période d'allumage. Les générations réalisées pendant la période d'allumage sont éliminées des analyses. Donc de manière conservatrice, les 1.250.000 premières générations ont été éliminées (25%) et les inférences sont alors réalisées sur les 3.750.000 générations suivantes. La fiabilité d'un arbre se ramène donc à la fiabilité de ses branches internes (nœuds).

### II.1.3.5.3 DATATIONS MOLECULAIRES

Une bonne datation moléculaire est tributaire de la trilogie : choix rigoureux de modèle de substitution de nucléotide, de modèle d'horloge moléculaire adapté et de bons points de calibration (Ho et Lo, 2013). Ainsi, afin de dater l'émergence et la diversification des lignées à travers l'analyse phylogénétique, nous avons procédé à des analyses de datations moléculaires par une approche bayésienne avec le package BEAST ver. 1.8.2 (Drummond *et al.*, 2015). Ces analyses ont été réalisées avec le *Cyt-B*, le *COI* et les séquences concaténées sous le modèle d'évolution de GTR (Rodriguez *et al.*, 1990) avec une distribution gamma ( $\Gamma$ ) et une proportion de sites invariables (+I). L'analyse du MCMC a été exécutée avec 10.000.000 générations en échantillonnant toutes les 1.000 générations ; ensuite 1.000.000 premières générations ont été éliminées (10%) et les inférences sont alors réalisées sur les générations suivantes. Enfin, tenant compte des spécificités respectives des gènes *Cyt-B* et *COI* (Cox et Hebert, 2001 ; Simmons et Weller, 2001 ; Wares et Cunningham, 2001a ; 2001b ; Decaens *et al.*, 2013 ; Ndong *et al.*, 2015), les divergences intraspécifiques vont être estimées avec le *Cyt-B* et le *COI* ; alors que les divergences interspécifiques uniquement avec le *COI*. L'hypothèse de l'horloge moléculaire suppose que les substitutions dans une séquence s'effectuent à un taux constant. Le nombre de substitutions observées permettrait donc de connaître la date de divergence entre deux espèces. Mais les horloges moléculaires sont plus ou moins rapides selon les organismes, ou même à l'intérieur d'un organisme, sous l'influence du temps de génération, des systèmes de réparation et de la pression de sélection. Même si nous avons testé l'hypothèse de

l'horloge moléculaire avec MEGA en utilisant la méthode du Maximum de vraisemblance, nous avons adopté 2 types d'estimations :

(i) la méthode d'une horloge moléculaire constante qui postule que le taux d'évolution des lignées est constant. Dès lors, nous avons daté les lignées en considérant un taux d'évolution chez les insectes pour l'*ADNmt* de 1,5% ( $1,5 \cdot 10^{-2}$  substitution/site/million d'année) pour le *COI* et d'environ 2,3% ( $2,3 \cdot 10^{-2}$  substitution/site/million d'année) pour le *Cyt-B* (Brower, 1994 ; Farrell *et al.*, 2001).

(ii) la méthode d'une horloge moléculaire relâchée (Relaxed clock: uncorrelated lognormal) d'un modèle de variation au sein d'un ensemble de données qui suppose que les taux de substitution associés à chaque branche sont tirés indépendamment d'une seule distribution log-normale non-corrélée (Drummond *et al.*, 2006) pour estimer le taux d'évolution moléculaire et les temps de divergence entre les différents clades détectés.

La méthode d'horloge moléculaire relâchée présente trois avantages :

(i) la non nécessité de vérifier l'hypothèse d'un taux évolutif constant à travers l'histoire des eucaryotes ;

(ii) l'incorporation de contraintes à *priori* sur le temps de divergence est préférée par l'utilisation des points de calibrages afin de manipuler des incertitudes inhérentes aux données paléontologiques (les fossiles ne s'apparient jamais exactement aux nœuds des arbres phylogénétiques) ;

(iii) plusieurs points de calibrages indépendants peuvent être utilisés simultanément (Douzery *et al.*, 2004).

## II.2 RESULTATS

### II.2.1 DIVERSITE GENETIQUES DES POPULATIONS

#### II.2.1.1 POLYMORPHISME ET EVOLUTION GENETIQUES

Cent vingt (117) individus pour le *Cyt-B* et cent dix (112) pour le *COI* ont été séquencés ; cependant, cent douze (112), Cent neuf (109) et quatre-vingt-dix-neuf (99) individus ont constitué le jeu de données à analyser après alignements et corrections respectivement du *Cytochrome B (Cyt-B)*, du *Cytochrome oxydase I (COI)* et des séquences concaténées. Dans la population globale, il y a, respectivement pour le *Cyt-B*, le *COI*, et des Séquences concaténées : 442, 415 et 856 paires de bases (pb) selon le meilleur cadre de lecture. Il a été noté 384 (86,88%), 406 (97,83%) et 800 (93,46%) sites invariables ; 58 (13,12%), 9 (2,17%) et 56 (6,154%) sites variables dont 36 (62,07%), 1 (11,11%) et 37 (66,07%) sites informatifs en parcimonie et enfin 22 (37,93%), 8 (88,88%) et 19 (33,93%) sites singletons (Tableau III).

**Tableau III** : Polymorphisme génétique des matrices

Gènes	Nombre de taxa	Nombre de sites	Nombre de sites invariables	Sites polymorphes	Sites informatifs en parcimonie	Singletons	dN - dS	dN/dS
<i>Cyt-B</i>	112	442	389	53	41	12	-2,94	0,39
<i>COI</i>	109	415	406	09	01	08	-0,948	0,17
<i>Séquences concaténées</i>	99	856	800	56	37	19	-2,62	0,93

Les meilleurs modèles évolutifs déterminés par l'AICc (*Akaike Information Criterion corrected*), le ratio  $n/k=1,94 < 40$  (Burnham and Anderson, 2002 ; 2004) sont le TrN (Tamura et Nei, 1993), le F81 (Felsenstein, 1981) et le TrN avec une distribution gamma ( $\Gamma$ ) et une certaine proportion de sites invariables (+I). Cependant, grâce au LRT (likelihood ratio test), le meilleur ajustement des données avec l'addition de paramètres supplémentaires biologiquement significatifs est le TrN + G + I ; le TrNef + G + I et le GTR + G respectivement pour le *Cyt-B*, le *COI* et les Séquences concaténées (Tableau IV).

**Tableau IV** : Matrices analysés : modèles évolutifs et polymorphisme nucléotidique

Gènes	Modèles	- lnL	P	LRT	fA	fT	fC	fG	p-inv	Shape	R	nst
<i>Cyt-B</i>	TrN + I	1119,5696	228	*	0,3059	0,4047	0,1844	0,1048	0,7950	---	0,61	06
	TrN + I + G	1121,5074	229		0,3028	0,4041	0,1862	0,1069	0,3230	0,1610	0,67	06
<i>COI</i>	F81	610,8868	219	***	0,2759	0,3856	0,1953	0,1433	---	---	1	01
	TrNef + I + G	637,8310	220		0,25	0,25	0,25	0,25	0,1820	76,6290	---	06
<i>Séquences concaténées</i>	TrN + I + G	1719,5877	203	***	0,2906	0,3977	0,1913	0,1204	0,7520	0,3480	0,80	06
	GTR + G	1735,4722	205		0,2893	0,3968	0,1934	0,1205	---	0,0290	0,93	06

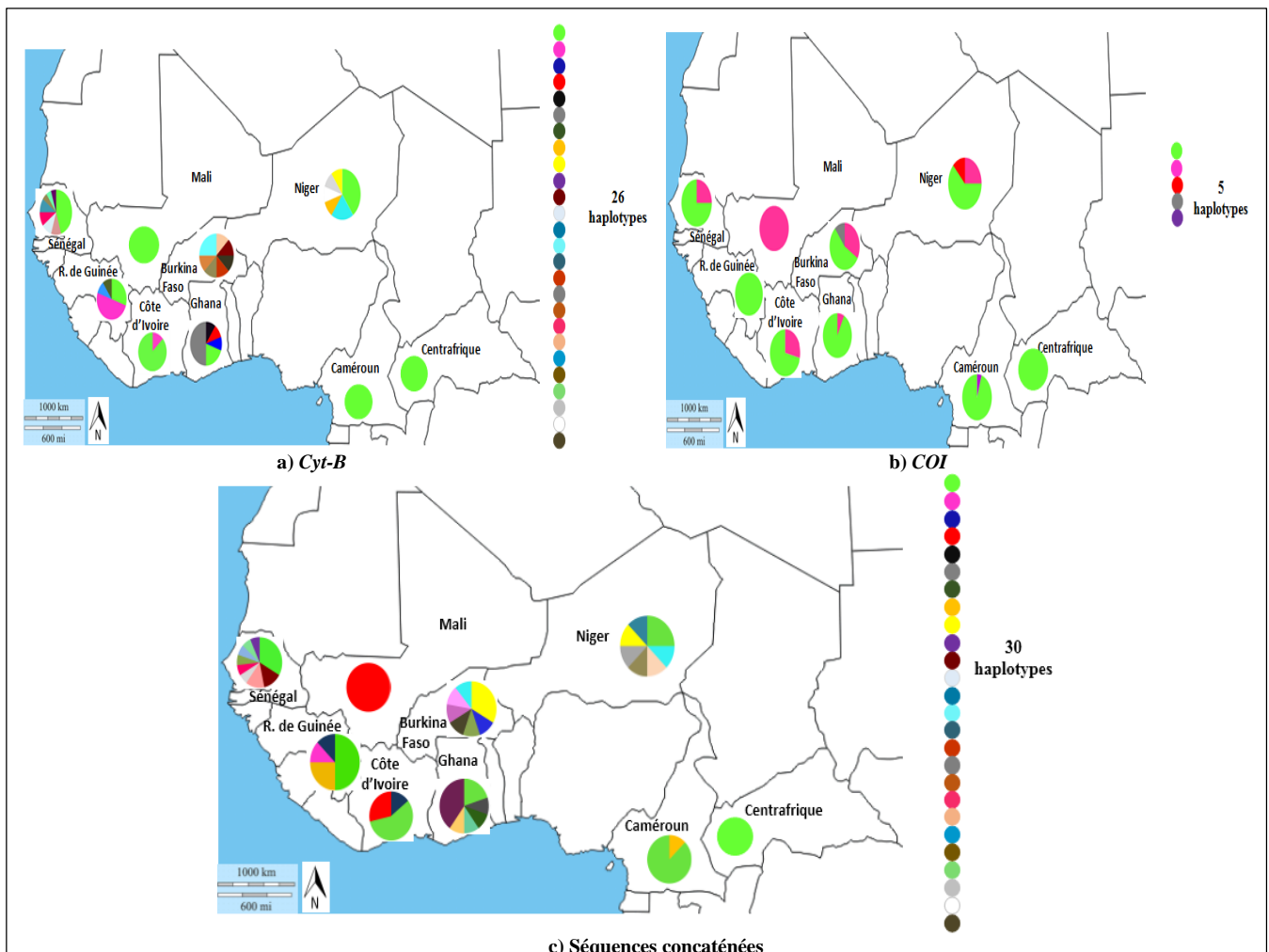
M1 : modèle simple ; M2 : modèle complexe ; lnL : likelihood scores ; P : nombre de paramètres ; P<sub>LRT</sub> : significativité du LRT (*likelihood ratio test*) ; \* : P<0,05 ; \*\* : 0,05<P<0,01 ; \*\*\* : 0,01<P<0,001 ; NS : non significatif (P > 0,05) ; f (A, T, C et G) : fréquences des nucléotides ; p-inv : proportion de sites invariables ; Shape : valeur du gamma ( $\Gamma$ ) ; R : ratio transitions/transversions ; nst : number of substitution types (nombre de type de substitutions).

Le contenu G+C est de 29, 34 et 31,5% contre 70, 66 et 68,5% pour celui d'A+T ; alors que le taux de mutations est en faveur des transversions qui sont de 52, 51 et 50,03% contre 48, 49 et 50,07% pour les transitions respectivement pour le *Cyt-B*, le *COI* et les séquences concaténées.

Le ratio global dN/dS sur le cadre de lecture ouvert pour la population globale a été estimé à 0,39, 0,17 et 0,93 respectivement pour le *Cyt-B*, le *COI* et les séquences concaténées.

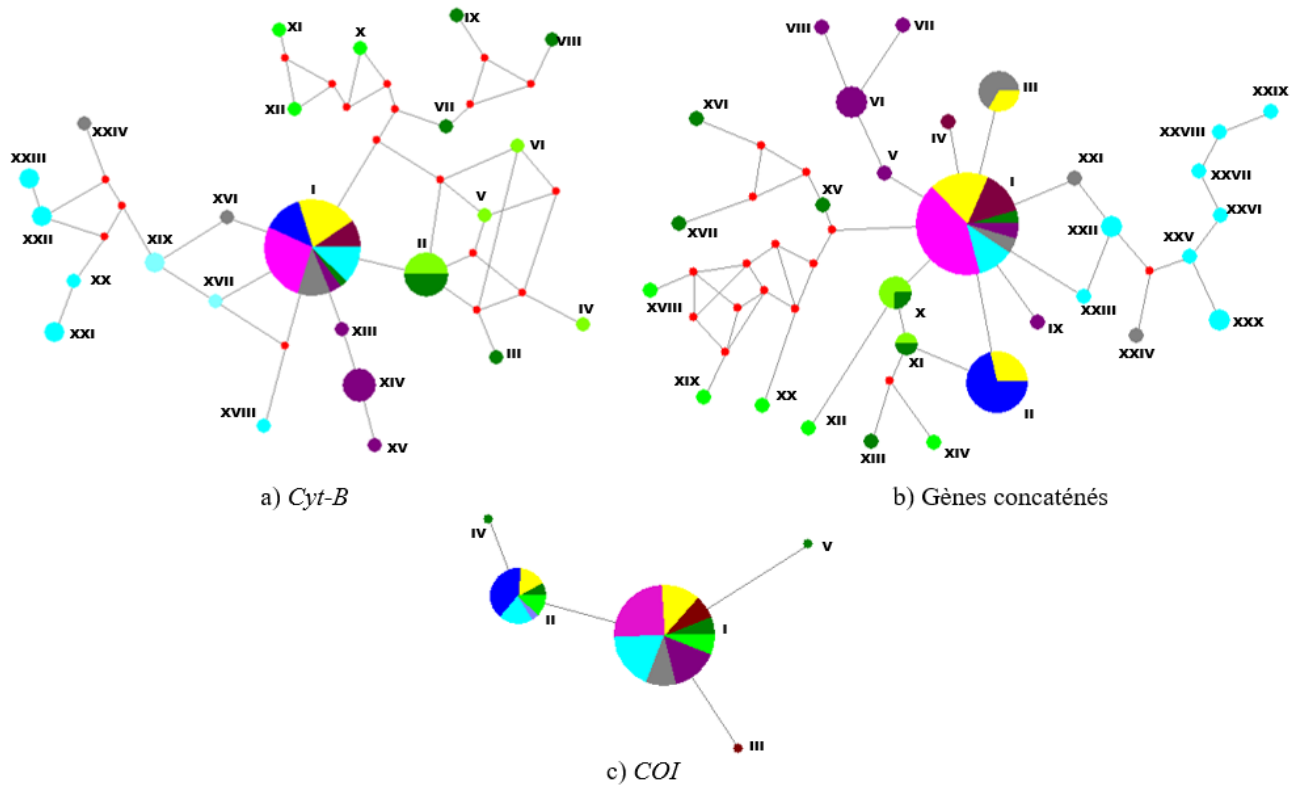
### II.2.1.2 DIVERSITE ET REPARTITION DES HAPLOTYPES

Il y a 26, 5 et 30 haplotypes observés respectivement pour le *Cyt-B*, le *COI* et les séquences concaténées. L'haplotype majoritaire représente 59, 74 et 43% des individus séquencés respectivement pour le *Cyt-B*, le *COI* et les séquences concaténées. L'haplotype majoritaire et central est présente partout sauf au Burkina Faso (*Cyt-B*), au Mali (*COI*), au Burkina Faso et au Mali (Séquences concaténées) (Figure 8). Il est remarquable de noter que les populations partagent des haplotypes en communs mais sont aussi caractérisées par des haplotypes privés. Le réseau d'haplotypes révèle une structuration plus ou moins étoilée. Il est noté la présence de plusieurs vecteurs médians et un nombre de pas de mutation très important dans les populations du Burkina Faso et du Niger (Figure 9).



**Figure 8 :** Cartes de distribution spatiale des haplotypes de *Sitophilus zeamais*. Les différents haplotypes rencontrés sont labellisés par couleur.





**Figure 9 :** Réseaux d'haplotypes de *S. zeamais*.

**Légende :** Répartition géographique et affinités phylogénétiques entre haplotypes.

Un haplotype est labellisé en fonction de différentes couleurs représentant les pays où il est rencontré

<span style="color: green;">●</span> Burkina Faso	<span style="color: brown;">●</span> Camérone	<span style="color: magenta;">●</span> Centrafrique	<span style="color: yellow;">●</span> Côte d'Ivoire	<span style="color: purple;">●</span> Ghana
<span style="color: blue;">●</span> Mali	<span style="color: darkgreen;">●</span> Niger	<span style="color: cyan;">●</span> Sénégal	<span style="color: grey;">●</span> R. de Guinée	<span style="color: red;">●</span> Vecteur médian

### II.2.1.3 DIVERSITE GENETIQUE DES POPULATIONS

Les valeurs de diversité nucléotidique sont assez faibles ( $\pi < 0,5\%$  ; Grant et Bowen, 1998) pour la population globale (*COI* et séquences concaténées), le groupe des populations de la zone sahélo-soudanienne (*COI*) et celui de la zone guinéo-congolaise (*Cyt-B*, *COI* et séquences concaténées). Cependant, ces valeurs de diversité nucléotidique sont assez fortes ( $\pi > 0,5\%$ ) pour la population globale (*Cyt-B*) et le groupe des populations de la zone sahélo-soudanienne (*Cyt-B* et séquences concaténées) (Tableau V). Parallèlement, la diversité haplotypique est forte ( $h > 50\%$ ) pour la population de la zone sahélo-soudanienne (*Cyt-B* et *COI*) et la population globale (*Cyt-B*) alors qu'elle est faible pour la population de la région guinéo-congolaise (*Cyt-B* et *COI*). En effet, il y a 22 haplotypes contre 6 ; et 51 sites variables contre 12 pour les régions biogéographiques sahélo-soudanienne et guinéo-congolaise avec le *Cyt-B* respectivement alors qu'il n'y a pas une grande différence en valeur absolue avec le *COI*. Notons l'exception des populations malienne et centrafricaine qui présentent une diversité génétique nulle pour les 2 gènes. Le test du Chi 2 révèle qu'il y a une différence significative à très significative de la diversité nucléotidique entre les 2 régions biogéographiques pour ces deux gènes.

**Tableau V : Indices standards de diversité génétique**

Groupes/Populations géographiques	Climat	N	N $\mu$	S	K	H $_T$	HP	%HP	Hd	$\pi$	
GSA	semi-aride	60	58	51	4,9898	22	20	90,91	0,8203	0,01129	
		55	04	04	0,5899	04	02	50	0,5165	0,00142	
		50	48	43	5,763	24	21	87,50	0,922	0,00673	
		Burkina Faso	10	18	15	5,9111	07	06	85,71	0,8667	0,01337
			09	02	02	0,7222	03	01	33,33	0,6389	0,00174
			09	20	17	6,972	07	-	-	0,917	0,00815
		Mali	10	00	00	0,0000	01	01	100	0,0000	0,00000
			10	00	00	0,0000	01	00	00,00	0,0000	0,00000
			10	00	00	0,0000	01	-	-	0,0000	0,00000
		Niger	10	24	22	7,0889	06	04	06,66	0,8444	0,01604
			08	03	03	1,0357	03	01	33,33	0,6071	0,00250
			08	27	25	9,571	07	-	-	0,964	0,01118
		République de Guinée	10	06	06	1,6889	04	02	50	0,7111	0,00382
			08	00	00	0,0000	01	00,00	00,00	0,0000	0,00000
			08	06	06	1,964	04	-	-	0,750	0,00229
Sénégal	20	24	22	4,9316	08	07	87,50	0,7895	0,01116		
	20	01	01	0,3947	02	00,00	00,00	0,3947	0,00095		
	15	11	11	4,419	09	-	-	0,886	0,00516		
GGG	Humide à sub-humide	52	12	12	2,3718	06	04	66,66	0,3424	0,00537	
		54	06	06	0,3564	03	01	33,33	0,2047	0,00086	
		49	18	18	2,889	09	06	66,66	0,512	0,00337	
		Cameroun	07	00	00	0,0000	01	00,00	00,00	0,0000	0,00000
			07	05	05	1,4286	02	01	100	0,2857	0,00345
			07	05	05	1,429	02	-	-	0,286	0,00167
		Centrafrique	20	00	00	0,0000	01	00,00	00,00	0,0000	0,00000
			20	00	00	0,0000	01	00,00	00,00	0,0000	0,00000
			18	00	00	0,0000	01	-	-	0,0000	0,00000
		Côte d'Ivoire	15	01	01	0,2476	02	00,00	00,00	0,2476	0,00056
			14	01	01	0,4396	02	00,00	00,00	0,4396	0,00106
		Ghana	14	02	02	0,703	03	-	-	0,615	0,00082
			10	11	11	5,1333	05	04	80,00	0,7556	0,01161
			13	01	01	0,1538	02	00,00	00,00	0,1538	0,00037
		Population globale	112	68	58	3,9574	26	-	-	0,6448	0,00895
109	09		09	0,5134	05	-	-	0,3986	0,00124		
99	63		56	4,602	30	-	-	0,789	0,00538		
Test du Chi2 (GSA vs GDG)					Cyt-B <sup>NS</sup>					Cyt-B <sup>NS</sup>	Cyt-B <sup>**</sup>
					COI <sup>NS</sup>					COI <sup>NS</sup>	COI <sup>***</sup>
					Gènes concaténés <sup>NS</sup>					Gènes concaténés <sup>NS</sup>	Gènes concaténés <sup>**</sup>

Cyt-B (gras) ; COI (italique) ; Séquences concaténées (gras et italique) ; N : nombre de séquences ; N $\mu$  : nombre de mutations ; S : nombre de sites polymorphes ; H : nombre d'haplotypes ; K : nombre moyen de différences nucléotidiques entre paires de séquences ; Hd : diversité haplotypique ;  $\pi$  : diversité nucléotidique ; GSA vs GGC : Groupe des Pays du Sahel versus Groupe des pays du domaine guinéo-congolais ; Valeurs de significativité : \* : P<0,05 ; \*\* : 0,05<P<0,01 ; \*\*\* : 0,01<P<0,001 ; NS : non significatif (P > 0,05).

### II.2.1.4 ACP DES INDICES DE DIVERSITE GENETIQUE DES POPULATIONS

L'analyse en composantes principales (ACP) des indices standards de la diversité montre que plus de 90% de la diversité génétique est expliquée par 2 axes F1 et F2. Elle indique un changement de configuration des groupes de populations en fonction du marqueur. Elle met en évidence une affinité particulière entre les populations du Burkina Faso et du Niger d'une part ; et d'autre part entre la République de Guinée et le Mali et enfin un isolement plus ou moins important des individus issus du Sénégal. Il est également noté un certain degré d'introgession entre les deux régions biogéographiques (GSA et GGC) (Figure 10). La formation de ces groupes est déterminée principalement par les paramètres : diversité nucléotidique, nombre d'haplotypes et nombre moyen de différences entre paires de séquences.

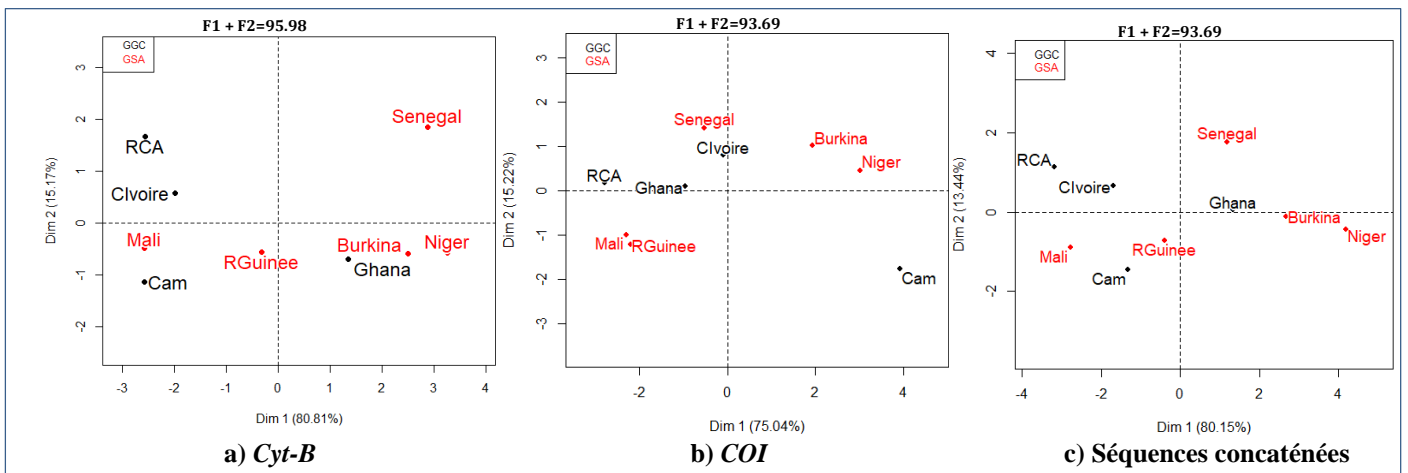


Figure 10 : Analyse en composantes principales des populations

II.2.2 STRUCTURE GENETIQUE ET RELATIONS PHYLOGEOGRAPHIQUES DES POPULATIONS

II.2.2.1 DIVERGENCE ET DIFFERENCIATION GENETIQUES DES POPULATIONS

L’analyse des matrices de divergence et de différenciation génétiques (Annexes 2, 3, 4, 5 et 6) par une classification ascendante hiérarchique (CAH), qui permet d’identifier des groupes homogènes en fonction des variables, donne une topologie variable selon le marqueur. Elle met en évidence une grande affinité entre le Burkina Faso et le Niger d’une part ; et d’autre part entre le Sénégal et la République de Guinée par de fortes valeurs d’AU (*Approximately Unbiased*) *p-values* (en rouge) et de BP (Bootstrap Probability) *p-values* (en vert). Elle indique une différenciation et un isolement de la population malienne des autres (*COI* et séquences concaténées) et celui du Ghana des autres (*Cyt-B*) (Figure 11). Elle révèle aussi des processus d’introgession entre régions biogéographiques.

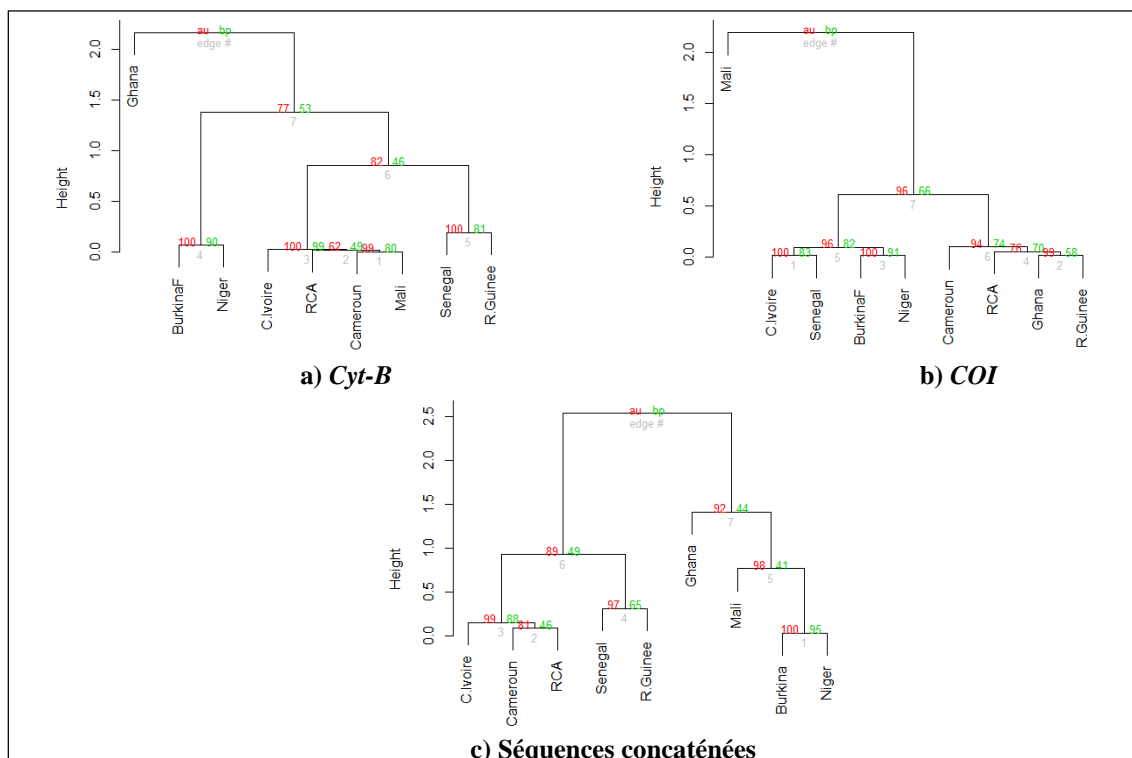
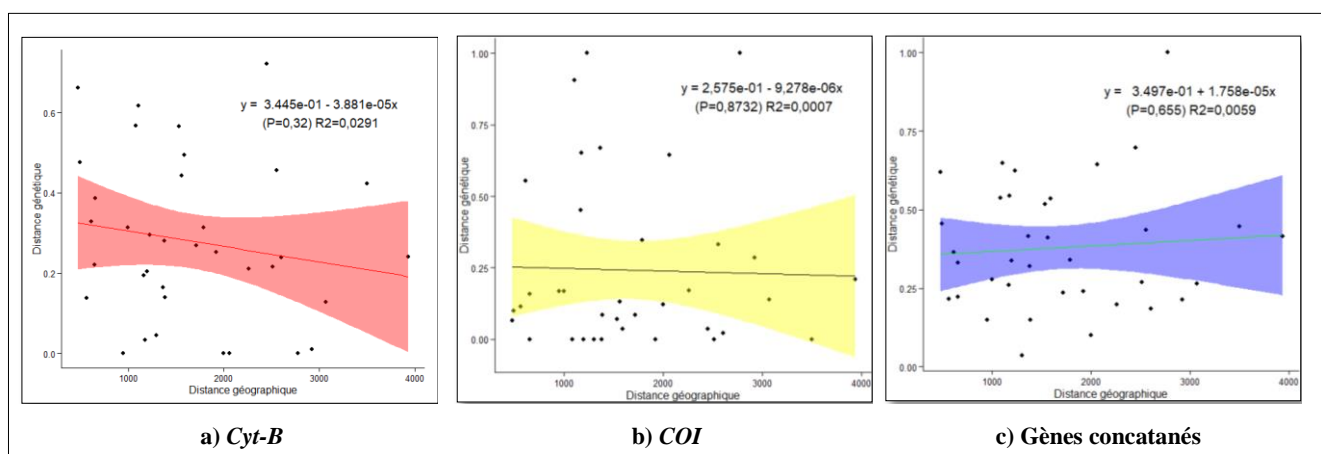


Figure 11: Classification hiérarchique ascendante des populations

### II.2.2.2 ISOLEMENTS GENETIQUES

Les hypothèses de linéarité, de normalité et d'homogénéité et d'indépendance des résidus ont été acceptées. Alors, les résultats de la régression linéaire sont valides. Le test de Mantel (Figure 12) révèle donc qu'il y a une corrélation entre la distance génétique et la distance géographique qui est négative avec le *Cyt-B* ( $r = -0,170$  ;  $P = 0,320$ ) et le *COI* ( $r = -0,027$  ;  $P = 0,873$ ) alors qu'elle est positive avec les séquences concaténées ( $r = 0,077$  ;  $P = 0,655$ ). Néanmoins, cette corrélation n'est pas statistiquement significative.



**Figure 12:** Test de Mantel (Distances géographiques en fonction des distances génétiques)

### II.2.2.3 DISTRIBUTION DE LA VARIANCE GENETIQUE ET REPARTITION SPATIALE

#### II.2.2.3.1 SOURCE DE LA VARIANCE MOLECULAIRE (AMOVA)

L'analyse de la population globale par AMOVA indique que la source de la variance moléculaire est due principalement à 2 éléments (Tableau VI) :

- la différenciation génétique moyenne entre population représentant 31 à 35% de la variation totale ;
- les caractéristiques différentielles individuelles, c'est-à-dire la différenciation entre individus d'une même population qui constitue plus de 60 à 68% de la variance moléculaire totale.

En considérant la pluviométrie comme facteur de structuration, 2 groupes de populations ont été définies à savoir le groupe des populations de la zone sahélo-soudanienne (GSA) et le groupe des populations de la zone guinéo-congolaise GGC) (Tableau VI), l'analyse confirme que l'importance de la variance moléculaire est due essentiellement à une diversité génétique au sein d'une population qui accumule 85 à 91% de la variation totale. Cependant, la différenciation entre les deux groupes est faible mais significativement importante et allant de 8 à 14% de la variation totale.

Tableau VI : Source et répartition de la variance moléculaire

Source de la Variance	Degrés de liberté			Somme des carrés de déviation			Composantes de la variance			Pourcentage de la variance			Indices de fixation (Test de significativité)
<b>Répartition de la variance moléculaire entre zones biogéographiques</b>													
Entre Zone	1	1	1	11,953	2,353	14,958	0,181	0,0388	0,25836	8,73	14,07	10,64	$F_{ST}=0,0873$ *** $F_{ST}=0,14071b$ *** $F_{ST}=0,1086$ ***
Au sein d'une zone	110	107	97	207,68	25,37	210,55	1,888	0,2371	2,17059	91,27	85,93	89,36	
Total	111	108	98	219,63	27,73	225,51	2,069	0,2760	2,42895	100	100	100	
<b>Répartition de la variance moléculaire entre populations</b>													
Entre populations	8	8	8	81,851	9,395	93,451	0,726	0,083	0,941	35,19	31,28	39,070	$F_{ST}=0,3519$ *** $F_{ST}=0,3128$ *** $F_{ST}=0,3907$ ***
Au sein des populations	103	100	90	137,78	18,33	132,05	1,3377	0,1833	1,467	64,81	68,72	60,930	
Total	111	108	98	219,63	27,72	225,50	2,0641	0,2667	2,408	100	100	100	

Résultats : *Cyt-B* (gras), *COI* (italique) et Séquences concaténées (gras et italique) ; \* : Niveau de significativité après 1023 permutations.

**II.2.2.3.2 ANALYSE SPATIALE DE LA VARIANCE MOLECULAIRE**

L'analyse SAMOVA (Spatial Analysis of MOlecular VARIance) a mis en évidence deux groupes génétiques statistiquement non significatifs ( $P > 0,05$ ) pour chaque marqueur. En effet, il y a le groupe génétique constitué uniquement par le Ghana d'une part et d'autre part le groupe composé du Burkina Faso, du Camérout, de la Centrafrique, de la Côte d'Ivoire, du Ghana, de la République de Guinée, du Mali, du Niger et du Sénégal (*Cyt-B* et les Séquences concaténées ; Figure 13 ; Tableau VII). Il y a aussi le groupe formé par le Burkina Faso, le Camérout, la Centrafrique, la Côte d'Ivoire, le Ghana, la République de Guinée, le Niger et le Sénégal et celui formé par le Mali (*COI*) ; Figure 13 ; Tableau VII).

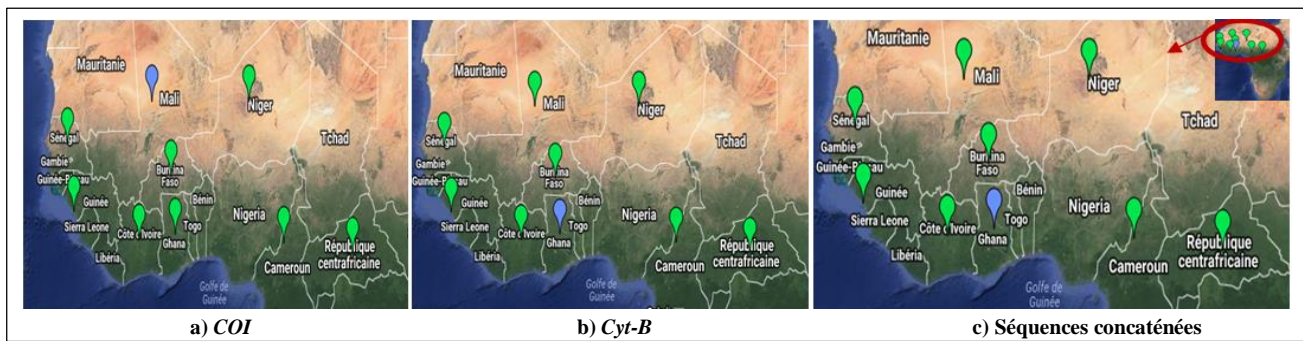


Figure 13: Cartes de distribution spatiale des groupes génétiques *Sitophilus zeamais*. Les groupes génétiques étant définis en fonction de la coloration des icônes.

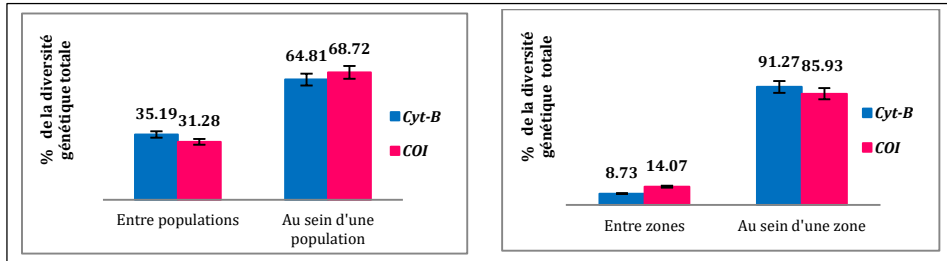
Tableau VII : Analyse de la variance moléculaire spatialisée

Source de la variance	Degrés de liberté			Somme des carrés de déviation			Composantes de la variance			Pourcentage de la variance			Indices de fixation		
Entre groupe de populations	1	1	1	40,279	6,39	40,21	1,928	0,330	1,835	53,06	61,85	47,51	$F_{SC} = 0,216$ *** $F_{ST} = 0,632$ *** $F_{CT} = 0,531$ NS	$F_{SC} = 0,100$ *** $F_{ST} = 0,657$ * $F_{CT} = 0,618$ NS	$F_{SC} = 0,277$ *** $F_{ST} = 0,620$ *** $F_{CT} = 0,475$ NS
Entre populations	7	7	7	41,572	3,00	53,24	0,368	0,020	0,56	10,12	3,81	14,52			
Au sein des populations	103	100	90	137,78	18,33	132,05	1,338	0,183	1,467	36,82	34,34	37,98			
Total	111	108	98	219,63	27,73	225,51	3,634	0,534	3,864	100	100	100			

Résultats : *Cyt-B* (gras), *COI* (italique) et Séquences concaténées (gras et italique) ; \* : Niveau de significativité après 1000 permutations.

**II.2.2.4 RESUME DE LA REPARTITION DE LA DIVERSITE GENETIQUE**

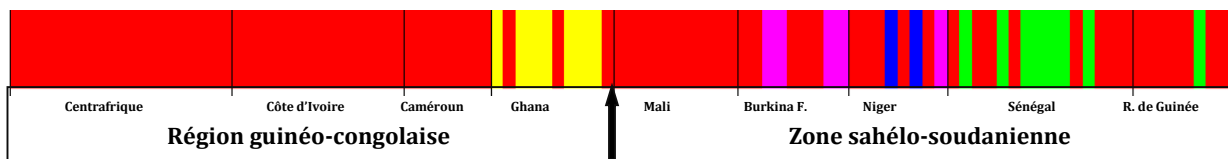
La différenciation entre individus constituant une population ou un groupe de populations inféodées à des régions biogéographiques est la principale source de la variation moléculaire (Figure 14).



**Figure 14 :** Répartition de la diversité génétique chez *S. zeamais* en Afrique de l'Ouest

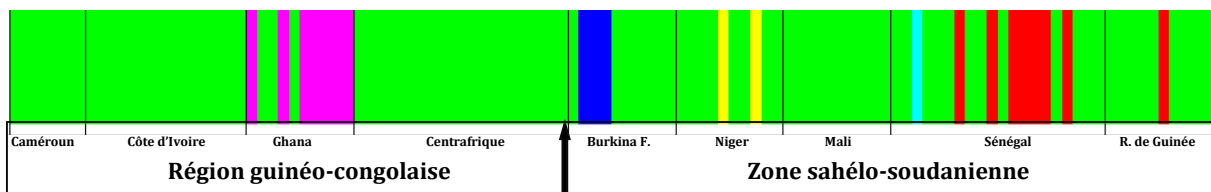
**III.2.2.5 CLUSTERING PAR ANALYSE BAYESIENNE DE LA STRUCTURE GENETIQUE**

Les figures (15, 16 et 17) indiquent les résultats de l'analyse du BAPS (Bayesian Analysis Population Structure) et les *clusters* sont labellisés par couleur. Avec les Séquences concaténées (Figure 15), l'analyse de *clustering* hiérarchique avec le modèle mixture permet d'identifier 5 *clusters* avec une probabilité de 99,99% contre 4 groupes par le modèle d'admixture.



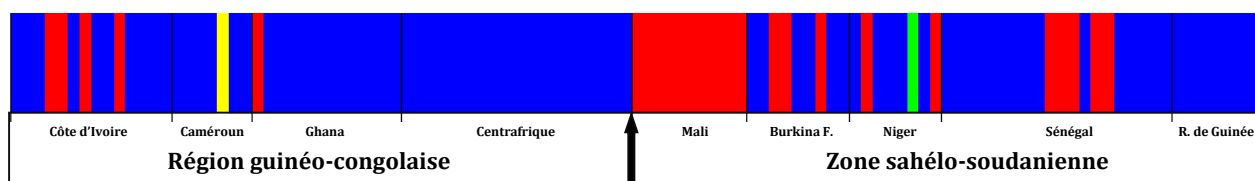
**Figure 15:** *Clustering* des populations par le modèle mixture avec les Séquences concaténées

Le modèle mixture du *Cyt-B* (Figure 16) permet de recenser 6 *clusters* avec une probabilité de 85,80% alors que le modèle admixture indique 3 groupes.



**Figure 16:** *Clustering* des populations par le modèle mixture avec le *Cyt-B*

Les deux modèles présentent la même topologie avec le *COI* (Figure 17). Il a été dénombré 4 *clusters* avec une probabilité de 75,50% avec le modèle mixture contre deux *clusters* à l'aide du modèle admixture.



**Figure 17 :** Clustering des populations par le modèle mixture avec le *COI*

Globalement, il a été constaté une différenciation génétique entre les individus de la zone sahélo-soudanienne et ceux de la zone guinéo-congolaise. Cependant, cette dernière présente moins de diversité génétique que la première.

### II.2.2.6 INDICES DE DIFFERENCIATION GENETIQUE ET DETECTION DE SIGNAL PHYLOGEOGRAPHIQUE

Aussi bien entre les populations (pays) qu'entre les groupes de populations (régions biogéographiques), les valeurs des  $N_{ST}$  ne sont significativement pas plus grandes que celles des  $G_{ST}$ . Néanmoins, la différenciation entre les 2 régions biogéographiques est assez importante ( $14 \leq G_{ST} \leq 22$ ) même si la structure génétique des populations est plus marquée ( $35 \leq G_{ST} \leq 38$ ) (Tableau VIII).

**Tableau VIII :** Indices de signal phylogéographique

	Zones	$G_{ST} \pm (SE)$	$N_{ST} \pm (SE)$	$G_{ST}$ vs $N_{ST}$	$H_T \pm (SE)$	$H_S \pm (SE)$	$D_{ST}$	$N_m$
<i>Cyt-B</i>	GSS vs GGC	0,137 (NC)	0,090 (NC)	NS	0,674 (0,0000)	0,581 (0,2390)	0,093	5,225
	Entre population	0,365 (0,0542)	0,366 (NC)	NS	0,727 (0,1411)	0,462 (0,1297)	0,265	0,4359
<i>COI</i>	GSS vs GGC	0,167 (NC)	0,141 (NC)	NS	0,433 (0,0000)	0,361 (0,1559)	0,153	3,053
	Entre population	0,381 (0,2325)	0,280 (0,1910)	NS	0,452 (0,0983)	0,280 (0,0854)	0,172	0,4070
Gènes concaténés	GSS vs GGC	0,221 (NC)	0,066 (NC)	NS	0,854 (0,0000)	0,665 (0,2659)	0,189	4,201
	Entre population	0,350 (0,1353)	0,369 (0,0526)	NS	0,846 (0,0902)	0,550 (0,1312)	0,296	0,4645

Indices de différenciation génétique ( $F_{ST}$ ,  $N_{ST}$ ,  $G_{ST}$ ), Indices de Nei ( $H_T$ ,  $H_S$  et  $D_{ST}$ , 1973 ; 1987), Flux géniques et nombre de migrants ( $N_m$ ) : **GSA vs GGC** : Groupe des pays Sahélo-soudaniens versus Groupe des pays du domaine guinéo-congolais ; Les valeurs significatives de significativité : \* :  $P < 0,05$  ; \*\* :  $0,05 < P < 0,01$  ; \*\*\* :  $0,01 < P < 0,001$  ; NS : non significatif ( $P > 0,05$ ) ; SE : Standard error.

### II.2.2.7 PROCESSUS HISTORIQUES ET STRUCTURE DES POPULATIONS

Dans le cadre de l'analyse des clades emboîtés ou NCPA (nested-clade phylogeographic analysis) selon Templeton *et al.* (1995) et Templeton (1998), les cercles avec les nombres romains (I à XXI) se rapportent à des haplotypes (0-step clades) ; les boîtes en ligne simple sont marquées 1–1 à 1–4 et représentent des 1-step clades ; et enfin les lignes en pointillés sont marquées 2–1 à 2–2 indiquent des 2-step clades (Figure 18). Une ligne fine et grise indique un raccordement entre haplotypes ou entre n-steps clades. L'analyse des résultats et l'inférence du réseau des clades emboîtés faites selon Templeton (2011) sont présentés dans le Tableau IX. Elles permettent de mettre en exergue les principaux éléments historiques responsables de la structure génétique actuellement observée.

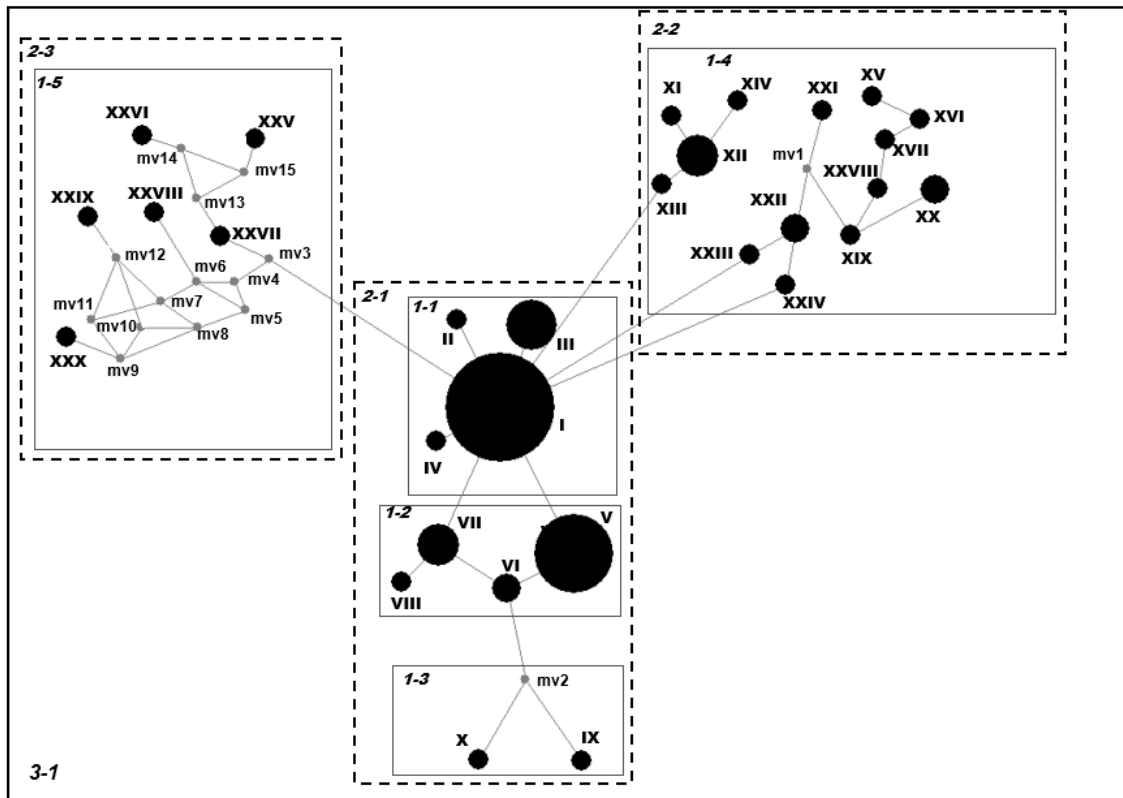


Figure 18 : Réseau d'haplotypes en clades emboîtés des Séquences concaténées

Tableau IX : Résultats d'analyse des clades emboîtés

Clades	Chaine d'inférence	Pattern inféré
Clade 1-1	1-2-3-4 : NO	Flux de gène restreint avec isolation par la distance
Clade 1-2	1-2-11-17 : NO	Résultats non concluants
Clade 1-3	1-2-11-17 : NO	Résultats non concluants
Clade 1-4	1-2-11-17 : NO	Résultats non concluants
Clade 1-5	1-2-11-17 : NO	Résultats non concluants
Clade 2-1	1-2-3-5-6-7 : OUI	Restriction de flux de gènes / Dispersion avec colonisation à longue distance
<b>Cladogramme total</b>	<b>1-2-3-4: NO</b>	<b>Flux de gène restreint avec isolation par la distance</b>

### II.2.3 EVOLUTION DEMOGRAPHIQUE DES POPULATIONS

Les valeurs des tests démogénétiques ( $D$  de Tajima et  $F_s$  de Fu) sont négatives et non significatives ; et celles du  $R^2$  sont positives et non significatives (Tableau X) pour la population globale. Similairement, les indices démographiques (SSD et  $R_g$ ) sont non significatifs avec une exception de la région Guineo-Congolaise où le SSD est significatif bien que les courbes de mismatch distributions du *Cyt-B* et des Séquences concaténées sont bimodales aussi bien pour la population globale que pour les régions biogéographiques (Figure 19, 20 et 21). En effet, les courbes de mismatch distributions de la population globale ont montré une distribution bimodale avec 2 pics approximativement à 7 et 12 différences (Figure 19, 20 et 21). Les mêmes résultats ont été notés pour la région Sahélo-Soudanienne, tandis que pour la région Guineo-Congolaise, le second pic est presque imperceptible.



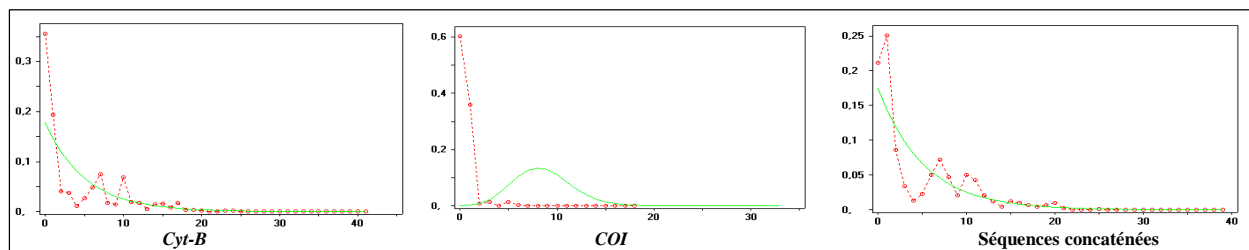


Figure 19 : Mismatch distributions par paires d'individus de la population globale

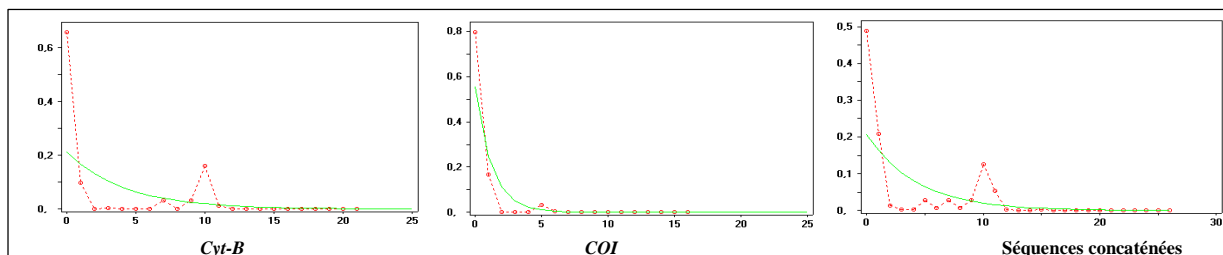


Figure 20 : Mismatch distributions par paires d'individus de la population guinéo-congolaise

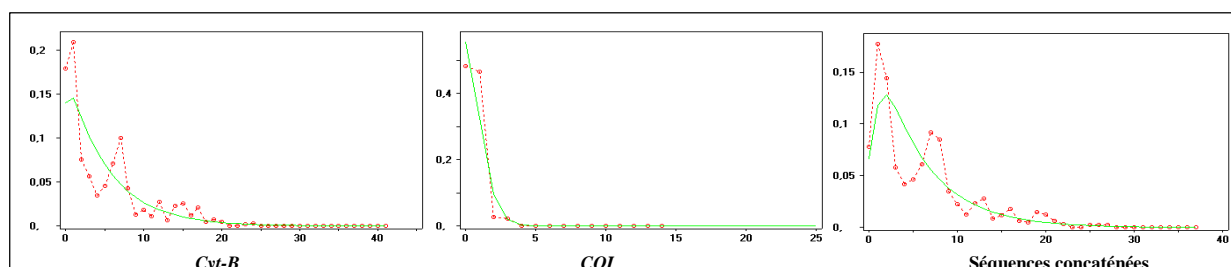


Figure 21 : Mismatch distributions par paires d'individus de la population sahélienne

Tableau X : Statistiques de la démographie du passé et tests de neutralité

Paramètres démogénétiques	Groupes Géographiques	N	D de Tajima	Fs de Fu	R <sup>2</sup>	SSD	Rg	S/K
Cyt-B	GSA	60	-1,8350 (0,0110)*	-5,7384 (0,0340)	0,1007 (0,5170)	0,7493 (0,0000)**	0,0270 (1,0000)	10,22
	GGC	52	-0,31392 (0,39600)	1,84322 (0,8060)	0,1063 (0,5350)	0,04557 (0,1000)	0,36355 (0,4400)	5,06
	Population globale	112	-1,07444 (0,2030)	-1,94760 (0,4200)	0,1089 (0,5600)	0,39741 (0,0500)	0,1953 (0,7200)	14,66
COI	GSA	55	-0,7230 (0,2880)	-0,37943 (0,3890)	0,1089 (0,5600)	0,0234 (0,0700)	0,19413 (0,0300)*	6,78
	GGC	54	-1,8285 (0,0100)*	-0,2619 (0,3400)	0,1083 (0,5190)	0,0023 (0,4200)	0,4240 (0,6100)	16,84
	Population globale	109	-1,27573 (0,1490)	-0,32066 (0,3645)	0,0917 (0,5350)	0,01282 (0,2450)	0,30905 (0,3200)	17,53
Séquences concaténées	GSA	50	-1,3668 (0,0670)	-7,8354 (0,0110)*	0,1058 (0,5360)	0,0154 (0,9500)	0,0234 (0,7400)	7,46
	GGC	49	-0,8974 (0,1790)	0,1722 (0,5810)	0,1044 (0,5790)	0,3270 (0,0000)**	0,1362 (1,0000)	6,23
	Population globale	99	-1,13210 (0,1230)	-3,83159 (0,2960)	0,0922 (0,5450)	0,1712 (0,4750)	0,0798 (0,8700)	12,17

N : nombre d'individus par population ; D de Tajima (Tajima, 1989a, b) ; R<sup>2</sup> (Ramos-Onsins et Rosas, 2002) ; Fs de Fu (Fu, 1997) ; SSD (Sum of Squared Deviation) ; Rg (Indice du Ruggedness de Harpending, 1994) ; S/K : Coefficient d'expansion ; GSA vs GGC : Groupe des Pays du Sahel versus Groupe des pays du domaine guinéo-congolais ; Pop. : Population. Les valeurs accompagnées du symbole \* sont statistiquement significatives où pour D de Tajima ; SSD et Rg, P-value ≤ 0,05 ; pour le Fs de Fu P-value ≤ 0,02.

## II.2.4 RELATIONS PHYLOGENETIQUES DES HAPLOTYPES

### II.2.4.1 ARBRES PHYLOGENETIQUES ET SIGNAL PHYLOGENETIQUE

Il existe un bon signal phylogénétique au sein des séquences analysées d'après les inférences de Lyons-Weilers *et al.* (1998) et le test de saturation des sites de substitution de Xia *et al.* (2003) qui met en évidence que le niveau de saturation des séquences utilisées est négligeable car l'indice de saturation de substitution (Iss) est plus petit que sa valeur critique (Iss.c) (Tableau XI). Ce qui est confirmé par la moyenne  $\phi$ -value du test de Steel (1993) qui était plus grande que 0,04.

En adoptant la méthode de Brochier *et al.* (2002), nous avons trouvé que :

$$\ln (\text{Séquences concaténées}, r) < \Sigma (\ln (\text{Cyt-B} - \text{COI}, r)).$$

La vraisemblance du modèle le moins riche en paramètres ( $\ln (\text{Séquences concaténées}, r ; 69 \text{ paramètres} : \ln L = -2795,4659)$ ) est nettement plus petite que celle du modèle le plus riche en paramètres ( $\Sigma (\ln (\text{Cyt-B}, \text{COI}, r ; 1080 \text{ paramètres} : \ln L = -2389,4273)$ ). Ce qui suggère que la différence soit significative et que la variation de la vitesse d'évolution entre branches a un effet plus important que la variation de vitesse d'évolution entre sites d'après les inférences de Brochier *et al.* (2002). Ainsi, l'analyse des gènes séparés avec une loi gamma ( $r$ ) est une bonne approximation, mais la meilleure phylogénie est obtenue par une analyse des Séquences concaténées avec une loi gamma ( $r$ ) (Tableau XII). Après une évaluation des conflits topologiques des différentes méthodes de reconstructions phylogénétiques (Figure 22) par le *likelihood-mapping* (cartographie de vraisemblance) (Strimmer et Von Haeseler, 1997), les résultats des méthodes probabilistes de l'approche cladistique (maximum de vraisemblance, ML pour maximum likelihood ; et inférence bayésienne, IB pour bayesian inference) avec les Séquences concaténées et le *Cyt-B* présentent globalement 5 haplogroupes correspondant aux *clusters* déjà identifiés par l'analyse bayésienne du BAPS et labellisés selon la couleur (Figures 23 à 26). En effet, le *likelihood-maps* montre que 79% des résultats appuient la topologie des haplogroupes identifiés par les différentes méthodes de reconstructions phylogénétiques (A et B).

Les arbres sont représentés avec des feuilles (haplotypes identifiés) coloriées en fonction de leur assignement dans les différents *clusters* détectés par le modèle de mixture de BAPS.

La topologie des arbres révèle principalement 5 clades (haplogroupes ou *clusters*) :

- 3 haplogroupes communs : 1 haplogroupe regroupant des individus originaires du Sénégal et de la République de Guinée et deux autres haplogroupes constitués chacun d'échantillons du Burkina Faso et du Niger ;
- 2 haplogroupes privés issus respectivement du Ghana et du Niger.

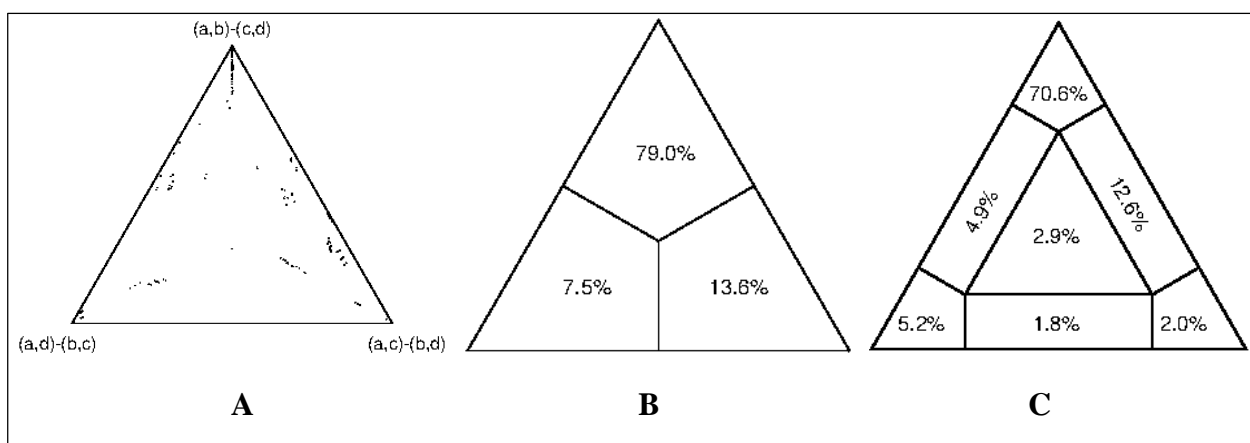
Cependant, il existe un nombre important d'haplotypes non résolutifs formant un râteau.

**Tableau XI** : Qualité du signal phylogénétique des séquences analysées

Gènes	Signal phylogénétique/ arbre non-enraciné (%)	Signal phylogénétique/ arbre enraciné (%)	Iss	Iss.c	Test de significativité de l'Iss
<i>Cyt-B</i>	83,94	99,90	0,0347	0,7014	***
<i>COI</i>	81,50	99,78	0,0111	0,7560	***
Séquences concaténées	84,02	99,98	0,0169	0,7432	***

**Tableau XII** : Analyse de la meilleure phylogénie obtenue parmi les matrices analysées

Gènes	Modèles sélectionnés	- lnL
Séquences concaténées	GTR	2835,9925
	GTR + G	2795,4659
<i>Cyt-B</i>	GTR	1581,1287
	GTR + G	1550,2416
<i>COI</i>	GTR	842,1218
	GTR + G	839,1857
<i>Cyt-B</i> + <i>COI</i>		$\sum \ln(Cyt - B, COI) = 2423,2505$
		$\sum \ln(Cyt - B, COI, r) = 2389,4273$



**Figure 22** : Diagrammes triangulaires du test « Four-cluster Likelihood-mapping » des séquences concaténées

**A** : Diagramme de cartographie de vraisemblance du modèle de branchement des haplotypes (a : cluster en vert ; b : cluster en marron ; c : cluster en rose ; d : cluster en bleu). **B** : Mesure du signal phylogénétique. **C** : Résultat du « Four-cluster Likelihood-mapping »

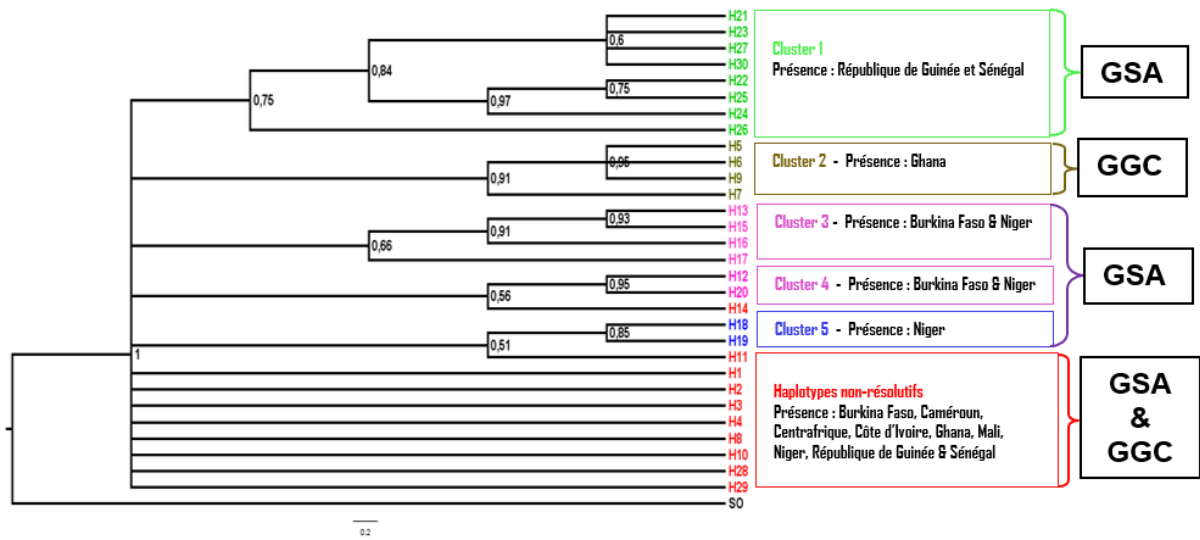


Figure 23: Inférence bayésienne des haplotypes des séquences concaténées

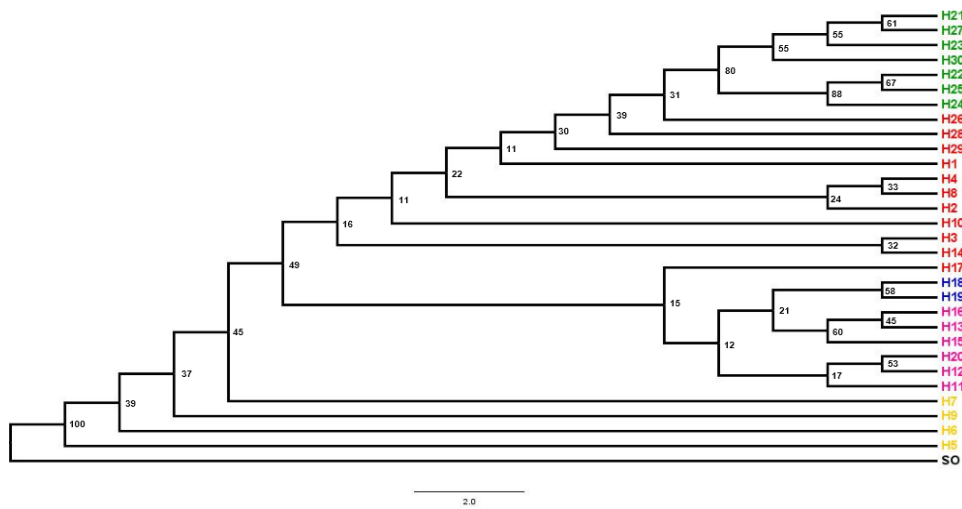
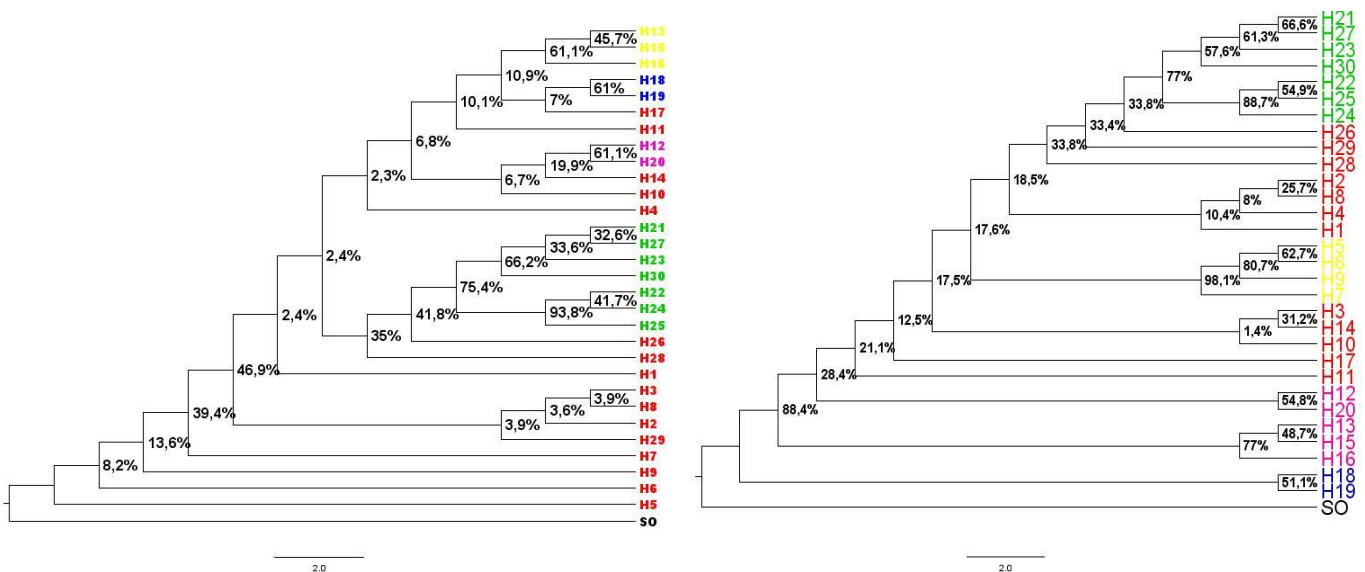


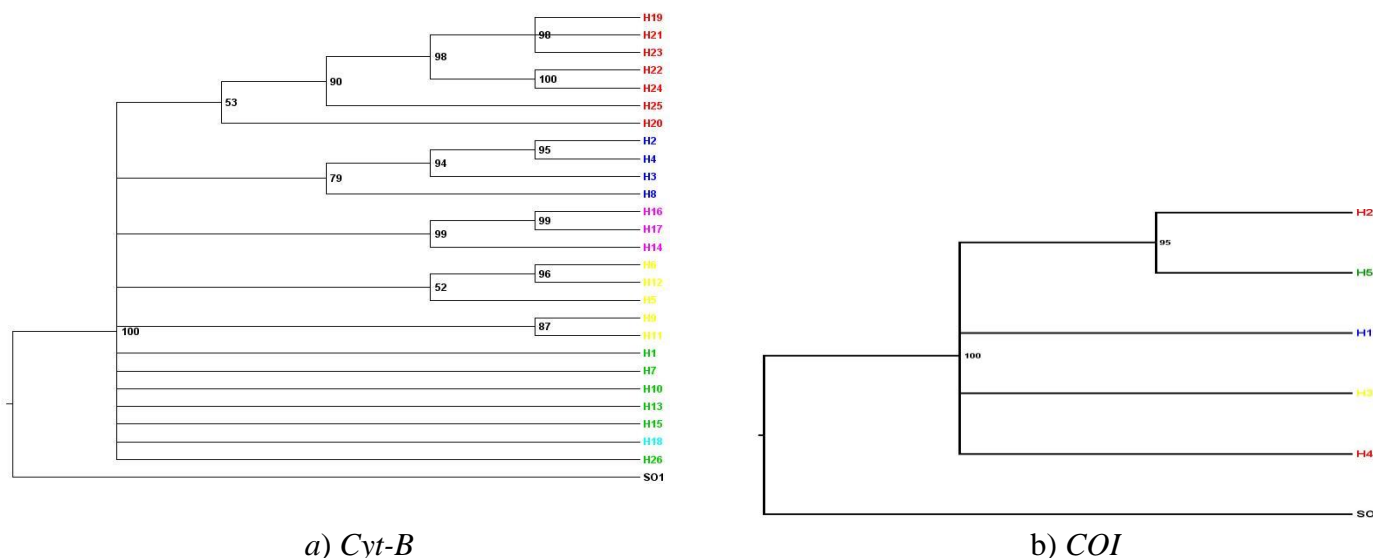
Figure 24 : Neighbor-joining des haplotypes des séquences concaténées



a) Méthode du maximum de parcimonie

b) Méthode du maximum de vraisemblance

Figure 25 : Phylogénie des haplotypes des séquences concaténées par le MP et MV



**Figure 26** : Inférence bayésienne des haplotypes par le *Cyt-B* et le *COI*

**II.2.4.2 DIVERGENCE ET STRUCTURE DES HAPLOGROUPES IDENTIFIES**

La divergence génétique entre les haplogroupes est de 1 à 3,4% avec une moyenne de 3,5% pour le *Cyt-B* et de 0,5% avec le *COI*. La diversité haplotypique est de 1 pour tous les *clusters* cependant la plus importante diversité nucléotidique est trouvée au sein de l’haplogroupe privé du Niger avec 2,9% (Tableau XIII et XIV). L’ancienneté d’un haplogroupe par rapport à un autre est déterminée par la diversité nucléotidique qui intègre une dimension évolutive (Tableau XIII et XIV),) et il a été utilisé pour déterminer l’origine probable de l’infestation et de la dispersion. Il indique la population du Niger comme probablement le lieu de l’infestation originelle (Figure 27).

**Tableau XIII** : Indices de diversité génétique

Indices de diversité		N	S	h	K	Hd	π	Dgm intra-haplogroupes	Dgm inter-haplogroupes			Dgm entre espèces SZ-SO
<i>Cyt-B</i>	Haplo-groupe I (Burkina Faso)	04	14	04	8,17	1,000	0,0185	0,020	0,000			
	Haplo-groupe II (Niger)	02	13	02	13,00	1,000	0,0294	0,034	0,023			
	Haplo-groupe III (Burkina Faso-Niger)	02	04	02	4,00	1,000	0,0005	0,009	0,008	0,022		
	Haplo-groupe IV (Ghana)	07	04	03	2,67	1,000	0,0060	0,006	0,030	0,051	0,026	
	Haplo-groupe V (Sénégal-R. Guinée)	10	09	06	4,13	1,000	0,0094	0,010	0,017	0,038	0,018	0,030
	Moyennes des haplogroupes									0,035		
<i>COI</i>	Haplo-groupe unique	26	02	02	2,00	1,000	0,00482	0,005				0,4264

N : Nombre d’individus ; K : Nombre moyen de différences ; h : Nombre d’haplotypes ; nucléotidiques entre séquences ; Hd : Diversité haplotypique ; π : Diversité nucléotidique ; Dgm : Distance génétique moyenne ; SZ : *Sitophilus zeamais* ; SO : *Sitophilus oryzae*.

Par ailleurs, l’analyse AMOVA indique qu’environ 70% de la variance moléculaire est due à une différenciation entre les 5 haplogroupes détectés par les Séquences concaténées (Tableau XV). Il est aussi important de noter que les différenciations entre l’haplogroupe I et l’haplogroupe II ; et entre l’haplogroupe I et l’haplogroupe III ne sont pas significatives d’où la présence de 4 haplogroupes (*clusters*) significativement différenciés (Tableau XVI). Il y a un fort flux génique entre les populations du Burkina Faso et du Niger.

**Tableau XIV** : Indices de diversité génétique des haplogroupes avec les Séquences concaténées.

Indices de diversité haplogroupes	N	S	h	K	Hd	$\pi$	Dgm		Dgm		Dgm entre SZ et SO	
							intra-haplogroupes		inter-haplogroupes			
Haplo-groupe I (Niger)	02	16	02	16,00	1,00	0,019	0,031		0,000		0,4264	
Haplo-groupe II (Burkina Faso-Niger)	07	16	04	9,17	1,00	0,011	0,015		0,022			
Haplo-groupe III (Burkina Faso-Niger)	02	23	04	4,00	1,00	0,005	0,005		0,018	0,006		
Haplo-groupe IV (Ghana)	07	05	04	2,50	1,00	0,003	0,003		0,053	0,025		0,016
Haplo-groupe V (Sénégal-R. de Guinée)	10	10	8	4,36	1,00	0,005	0,006		0,037	0,014		0,014
Moyennes des haplogroupes										0,027		

N : Nombre d'individus ; K : Nombre moyen de différences ; h : Nombre d'haplotypes ; nucléotidiques entre séquences ; Hd : Diversité haplotypique ;  $\pi$  : Diversité nucléotidique ; Dgm : Distance génétique moyenne ; SZ : *Sitophilus zeamais* ; SO : *Sitophilus oryzae*.

**Tableau XV** : Répartition de la variance moléculaire des haplogroupes des Séquences concaténées.

Source de la Variance	d.f.	Sommes des carrés	composantes de la Variance	Pourcentage de la variance	Indices de fixation
Entre haplogroupes	4	108,784	5,495 Va( $F_{ST}$ )	69,99	$F_{ST} = 0,6999$ ***
Au sein des haplogroupes	20	47,136	2,357 Vb	30,01	
Total	24	155,920	7,852	100	

d.f: degrees of freedom ; SSD : sum of squares ; \* : Niveau de significativité après 1023 permutations.

**Tableau XVI** : Différenciation génétique entre haplogroupes des Séquences concaténées.

Haplogroupes	1	2	3	4	5
1. Haplo-groupe I (Niger)	0,0000				
2. Haplo-groupe II (Burkina Faso-Niger)	0,2577 <sup>Ns</sup>	0,0000			
3. Haplo-groupe III (Burkina Faso-Niger)	0,4286 <sup>Ns</sup>	0,4272	0,0000		
4. Haplo-groupe IV (Ghana)	0,8720**	0,7510	0,8519	0,0000	
5. Haplo-groupe V (Sénégal-R. Guinée)	0,6280*	0,53678**	0,7315	0,8002	0,0000

\* : Niveau de significativité après 1000 permutations ; Ns : non significatif ( $P > 0,05$ ).

### II.2.4.3 ANALYSE DE LA COALESCENCE ET TEMPS DE DIVERGENCE

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. La divergence entre deux populations établie sur des séquences d'ADN augmente en fonction du temps et du taux d'évolution ou de mutation ( $\mu$ ) de la séquence ( $d=2\mu T$  ; Nei, 1972) selon le principe de l'horloge moléculaire (Zuckerandl et Pauling, 1962). Ainsi, le temps de divergence des haplogroupes selon l'horloge moléculaire est de 0,760870 MA en moyenne et celui entre les deux espèces jumelles (*Sitophilus zeamais* versus *Sitophilus oryzae*) est environ 14 MA (Tableau XVII). Alors que les résultats de la datation bayésienne de l'horloge moléculaire montrent un temps de divergence compris entre environ 0,6-0,85 Ma (*COI* et Séquences concaténées) à plus de 1Ma (*Cyt-B*) entre haplogroupes et plus de 6 Ma pour la divergence entre les espèces jumelles. Par ailleurs, l'hypothèse nulle, que le taux d'évolution est homogène entre les branches, est rejetée. Les taux de substitution varient significativement entre les branches et une phylogénie sans horloge moléculaire

est plus appropriée. Cependant, la méthode bayésienne d'une horloge moléculaire relâchée (relaxed clock : uncorrelated lognormal) où le taux de mutation varie en fonction des branches des arbres phylogénétiques indiquent que le temps de divergence entre haplogroupes est de 1 à 2 Ma (*Cyt-B* et *COI*) et celui entre espèces est d'environ 4Ma (Tableau XVIII).

**Tableau XVII :** Datation selon Nei (1972) avec  $T=D/2\mu$

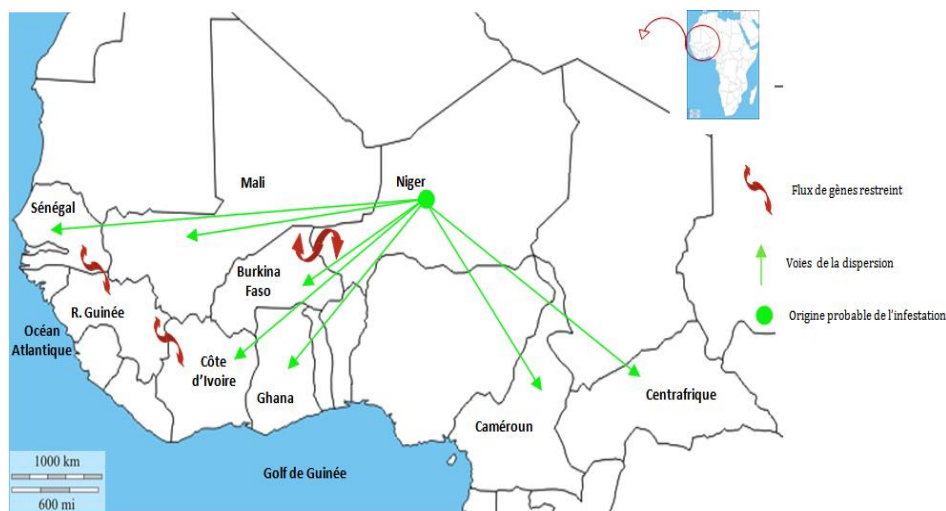
Marqueurs	<i>Cyt-B</i>	<i>COI</i>
<b>Temps de divergences</b>		
Au sein des haplogroupes	0,130435 – 0,739130 Ma	0,333333 Ma
TMRCA des haplotypes	0,173913 – 1,108696 Ma	---
Moyennes entre haplogroupes	0,760870 Ma	---
TMRCA entre SZ-SO	---	14,21333 Ma

TMRCA: Time of Most Recent Common Ancestor; SZ: *Sitophilus zeamais*; SO: *Sitophilus oryzae*.

**Tableau XVIII :** Datation bayésienne

Marqueurs	Horloge moléculaire			Horloge moléculaire relâchée		
	<i>Cyt-B</i>	<i>COI</i>	Séquences concaténées	<i>Cyt-B</i>	<i>COI</i>	Séquences concaténées
TMRCA des haplotypes 95% HPD	1,7196 [1,145 – 2,3295]	0,5718 [0,2345 – 0,9476]	0,8521 [0,6048 – 1,0983]	2,4929 [1,1894 – 4,0979]	1,0767 [0,2012 – 2,4085]	27,5747 [1,4272 – 86,6587]
TMRCA SZ-SO 95% HPD	-----	6,6758 [4,5753 – 9,1545]	-----	-----	4,0377 [1,8656 – 7,2683]	-----

TMRCA: Time of Most Recent Common Ancestor; HPD: Highest Posterior Density intervals ou Densité à postériori la plus forte ; SZ : *Sitophilus zeamais* ; SO : *Sitophilus oryzae*



**Figure 27 :** Origine probable de l'infestation et voies de dispersion de *Sitophilus zeamais*

## II.3 DISCUSSION

L'objectif de ce chapitre est de caractériser la diversité et la structure génétiques, de définir les relations phylogéographiques et de mettre en exergue l'impact des zones biogéographiques sur la structure génétique des populations de la sous-région de l'Afrique de l'Ouest à l'aide des séquences d'ADN. Il est important de noter qu'un enjeu majeur de l'étude de la dynamique spatio-temporelle de la différenciation réside dans la détection des barrières à la dispersion, l'identification des contraintes environnementales qui régulent les flux de gènes entre populations et leur évolution dans l'espace et le temps (Fontaine, 2008). Par ailleurs, la combinaison de méthodes spatiales et temporelles permet d'interpréter de manière plus exhaustive la structuration de la diversité génétique observée. En effet, l'une des étapes cruciales de la lutte intégrée, une alternative à la voie chimique (coûteuse et dangereuse), est la connaissance du statut bioévolutif des populations impliquées (Sembène, 2000).

### III.3.1 DIVERSITE GENETIQUE DES POPULATIONS

Les valeurs des paramètres caractéristiques de la diversité génétique ( $S$ ,  $K$  et  $\pi$ ) de cette étude sont dans les gammes observées dans des travaux effectués chez certains modèles biologiques tels que *Drosophila melanogaster* (Andolfatto, 2001) ; et sur les insectes ravageurs tels que *Busseola fusca* (Sézoulin *et al.*, 2006) ; *Callosobruchus maculatus* (Kébé, 2016) ; *Tribolium castaneum* (Diome *et al.*, 2013) ; *Sitophilus zeamais* (Ndong *et al.*, 2015). La richesse en A+T est importante chez les deux gènes comme attendu chez les insectes. En effet, chez ces derniers, en particulier aux troisièmes positions de codons, il y a un niveau élevé de saturation A+T dans l'ADN mitochondrial (> 75%) (Perna et Kocher, 1995; Pedersen, 2002). Globalement, aussi bien dans la population globale que dans les sous-populations, le nombre de substitutions synonymes est supérieur au nombre de substitutions non-synonymes. Les valeurs de l'oméga global ( $d_N/d_S$ ) sur le cadre de lecture ouvert ont caractérisé une sélection purificatrice ( $\omega \leq 0,3$ ) pour le gène *COI* et une neutralité sélective ( $0,3 < \omega < 1$ ) pour *Cyt-B* et les Séquences concaténées selon la méthode  $\gamma$ -MYN (modification de YangNielsen) (Wang *et al.*, 2009 ; Peters *et al.*, 2012). Notons que la grande majorité des mutations sont neutres (c'est-à-dire n'ont pas d'effet phénotypique) ou nuisibles et que les mutations avantageuses sont très rares. Habituellement, seules des parties d'une protéine sont soumises à une sélection positive tandis que d'autres sont sous une forte sélection de purification afin de maintenir la structure et la fonction basiques d'une protéine. En effet, cette valeur d'oméga inférieure à 1 pourrait être due à l'évitement de fonctions délétères plutôt qu'à une perte de fonction. Dès lors, si un gène est exprimé, l'utilisation des codons, le biais des nucléotides et d'autres facteurs (toxicité des protéines) généreront une sélection purifiante. De plus, la sélection généralement positive ne fonctionne que pendant un certain temps évolutif.

Les analyses du Chi 2 sur la diversité nucléotidique ( $\pi$ ) entre la région guinéo-congolaise et la zone sahélo-soudanienne ont montré une différence très significative avec les 3 matrices. Ainsi, les



paramètres écologiques (précipitations, température, humidité, etc.) semblent avoir un effet sur la diversité évolutive du charançon du maïs entre les deux régions biogéographiques car leur différence fondamentale est liée aux paramètres écologiques, particulièrement la pluviométrie. En outre, l'analyse des principaux composants des indices standards de la diversité génétique ont montré que cette divergence est principalement due à la différenciation des populations du Burkina Faso et du Niger par rapport aux autres populations. Ces deux populations ont, en effet, des caractéristiques typiquement sahéliennes telles que la courte saison agricole, l'aridité du climat, *etc...* Dès lors, les précipitations sont plus importantes dans le Sud que dans le Nord. Les valeurs des diversités nucléotidiques et haplotypiques ont aussi suggéré que l'infestation serait beaucoup plus ancienne en Afrique de l'Ouest qu'Afrique centrale.

### III.3.2 STRUCTURE GENETIQUE ET PHYLOGEOGRAPHIQUE

Selon les inférences de l'AMOVA, une grande partie de la variance moléculaire est due essentiellement aux caractéristiques individuelles d'une population. Ce qui évoque probablement l'impact des caractéristiques propres aux populations, comme les systèmes agricoles, l'intensité des activités commerciales dans la structure génétique des insectes nuisibles. Tandis qu'une différenciation génétique ( $F_{ST}$ ) faible mais significative serait liée aux paramètres écologiques des régions biogéographiques, d'où une implication probable de l'effet climatique sur la structure de la population. Ce qui est en phase avec les résultats du Chi2 sur la diversité nucléotidique. D'après Wright *et al.* (2006), le double taux de substitution de l'ADN des arbres tropicaux par rapport aux espèces nichant aux plus hautes latitudes serait associé au climat qui influencerait le métabolisme. Cette partition de la variation moléculaire indiquerait aussi que les activités anthropiques auraient une influence beaucoup plus importante sur la structure génétique des populations de ravageurs que les paramètres écologiques. En outre, la différenciation génétique n'est pas expliquée par l'isolement par la distance (IBD) (Wright, 1943). Selon Bossart et Prowell (1998), l'isolement génétique observé peut être dû à un autre facteur que la distance géographique entre les populations. Ces auteurs ont souligné que l'hétérogénéité écologique d'un habitat d'une espèce implique souvent des barrières à la dispersion.

La comparaison des valeurs de  $N_{ST}$  et de  $G_{ST}$  (Pons et Petit, 1996 ; Grivet, 2002) a suggéré qu'il n'y a pas de correspondance entre les similarités des haplotypes et leur répartition géographique. Ces résultats associés à ceux des analyses de SAMOVA, des indices de Nei (1973 ; 1987) et à des analyses multivariées (CAH, ACP, *etc...*) et bayésienne (BAPS, *etc...*) ont corroboré une structure génétique sans pattern phylogéographique de *S. zeamais* en Afrique de l'Ouest et du Centre. Ils ont aussi évoqué un flux de gènes multidirectionnel qui serait une conséquence du commerce des grains. En effet, les indices de Nei (1973 ; 1987) ont révélé que la contribution de la diversité génétique moyenne intra-population ( $H_S$ ) par rapport à la diversité génétique totale ( $H_T$ ) est supérieure à celle de la diversité

génétique moyenne inter-population ( $D_{ST}$ ) aussi bien entre populations qu'entre régions biogéographiques. En outre, les populations rurales de la sous-région ouest-africaine sont caractérisées par un dynamisme commercial à travers les échanges de grains. Ce transfert de grains d'un pays à un autre pourrait être accompagné de transfert de larves, de cocons de pontes ou même d'adultes ; et donc d'une installation nouvelle des insectes ravageurs dans des zones qui n'étaient pas jusque-là touchées par ces pertes post-récoltes (Kébé, 2013 ; Ndiaye, 2014). Ce qui tend à atténuer tout isolement génétique.

### III.3.2 AFFINITES PHYLOGENETIQUES, BIOGEOGRAPHIE ET HISTOIRE EVOLUTIVE

Les résultats des analyses des tests de saturation de Xia *et al.* (2003) et de Steel (1993) ont indiqué que les populations étudiées ont présenté un bon signal phylogénétique. Ainsi, les effets de la saturation et de l'homoplasie sur les données de *Cyt-B* et *COI* sont faibles à négligeables (Kartavtsev, 2011a ; b). En plus, le réseau d'haplotype a montré un haplotype majoritaire et central qui pourrait être considéré comme un haplotype sous-régional voire interrégional car il a été localisé jusqu'en Afrique centrale où il est essentiellement le seul haplotype. En outre, d'après les analyses combinées des résultats des méthodes de reconstructions phylogénétiques, de l'évaluation des différences topologiques en utilisant le *likelihood-maps* de Strimmer et Von Haeseler (1997) et des datations moléculaires, il a été identifié 5 groupes phylogénétiques (haplogroupes) qui auraient divergé au Pléistocène, il y a environ 1 million d'années d'après les datations moléculaires bayésiennes. Par conséquent, la présence de *S. zeamais* serait antérieure à la domestication du maïs dans ces régions biogéographiques et donc de l'émergence de l'agriculture (10.000-6.000 ans BP). Cette période du Pléistocène a été caractérisée par un climat instable en Afrique subsaharienne avec alternance de saisons chaudes et humides, froides et sèches (Wagner, 2002) ; et ces oscillations climatiques ont joué un rôle important dans la diversité contemporaine de nombreuses espèces et communautés (Hewitt 2000 ; 2004). Ces inférences sont en accord avec celles de Sezonlin *et al.* (2006) concernant le foreur de la tige de maïs (*Busseola fusca*) dont la structure génétique est expliquée par ces événements de la même période. Cette même caractéristique est également observée chez les rongeurs dans la sous-région (Quéroutil *et al.*, 2003, Bryja *et al.*, 2010). Il est important de noter que même si la présence de *S. zeamais* est antérieure à l'exploitation du maïs, le commerce des grains en a joué un rôle important. En effet, l'importation d'haplotypes originaires d'autres régions expliquerait la présence d'un nombre excessif d'haplotypes peu fréquents. D'ailleurs, les voies originelles de dispersion de l'infestation viendraient du Nord vers le Sud, plus précisément au Niger selon les indices de diversité génétique et les inférences du NCPA. En effet, en plus des valeurs de diversité (haplotypique et nucléotidique) très fortes ( $Hd$  K et  $\pi$ ) trouvées au Niger, l'haplotype majoritaire est central avec plus de 43% des individus échantillonnés ; or il existe une relation probabiliste directe entre la fréquence d'un allèle et son ancienneté (Watterson et Gues, 1977;

Donnelly et Tavaré, 1986) car l'haplotype le plus fréquent serait en moyenne le plus ancien. D'autre part, les clades externes devraient être plus récents que les clades internes qui devraient être plus géographiquement répandus (Donnelly et Tavaré, 1986 ; Templeton *et al.*, 1995). Néanmoins, selon Dick *et al.* (2003), Petit *et al.* (2003) et Ley *et al.* (2014), ces grandes valeurs de diversité peuvent également révéler à la fois une zone de refuge et/ou une zone de croisement des voies de recolonisation, c'est-à-dire une zone de contact secondaire. Cependant, cette hypothèse de l'origine sahélienne de l'infestation n'est pas conforme avec celle des zones refuges de la savane humide (région Guinéo-Congolaise) (Maley, 1996 ; 2001 ; Léal, 2004) où un petit nombre de lignées survivraient et se réfugiaient pendant les périodes défavorables du Pléistocène (aridité, glaciation) dans des zones confinées et coloniseraient d'autres territoires pendant les périodes favorables (fortes humides) (Hewitt, 2000 ; 2004). Effectivement, les variations de l'humidité, en particulier, ont contribué aux changements dans la répartition des forêts tropicales (Hammond, 2005) ; et conduirait à l'existence de zones refuges dans les zones les plus humides (Dick *et al.*, 2003). Dans la sous-région de l'Afrique de l'Ouest, par exemple, il existe le cas de rongeurs (Nicolas *et al.*, 2009 ; Bryja *et al.*, 2010) ; d'ongulés (Lorenzen *et al.*, 2012), *etc.* Toutefois, durant les périodes défavorables du Pléistocène, il y a eu des espèces qui se réfugiaient sous les montagnes ou les massifs rocheux au lieu des zones humides considérées comme des «zones de sécurité». C'est le cas d'un rongeur du genre *Acomys* (*Rodentia*, *Muridae*) (Nicolas *et al.*, 2009) connu pour sa préférence pour des habitats rocheux (Volobouev *et al.*, 2002). Par conséquent, il serait nécessaire de faire des études supplémentaires, plus approfondies afin d'identifier, avec plus de probabilité, l'origine de l'infestation et les voies de dispersion des lignées après un retour aux conditions environnementales propices à l'expansion. En ce sens, la collaboration interdisciplinaire entre, par exemple, la génétique des populations de ravageurs et l'archéo-entomologie serait un effet synergique pour donner une réponse scientifique à cette question sur l'origine des populations de *S. zeamais* en Afrique Subsaharienne.

### III.3.3 EVOLUTION DEMOGRAPHIQUE

Des divergences intraspécifiques faibles ( $\leq 1\%$ ) définissent les populations étudiées. Ces résultats combinés à l'analyse des courbes « mismatch distribution » validée par les indices démographiques (SSD et  $R_g$ ) ainsi que les indices de diversité ( $H_d$  et  $\pi$ ) et les tests démogénétiques ( $D$  de Tajima,  $F_s$  de Fu et  $R^2$ ) ont révélé une population globale qui a subi une expansion démographique modérée depuis le Pléistocène (environ 1 Ma). Selon les inférences de Grant et Bowen (1998) et Avise (2000), cette expansion rapide s'est faite à partir d'une population ancestrale à faible effectif car issue d'un bottleneck ; et où le temps n'est pas suffisant pour retrouver une forte diversité entre les haplotypes. Selon Lefèvre *et al.* (2016), ces phases de forte réduction d'effectifs ont un impact durable sur la diversité génétique et accentue la dérive génétique d'où une disparition de nombreux variants qui seront régénérés lentement par mutation à la fin de ce phénomène. D'après

ces auteurs, la diversité génétique d'une espèce refléterait la date et l'intensité du dernier goulot qu'elle a subi et non pas son abondance maximale à l'équilibre démographique. Cette récente explosion démographique pourrait être due à la colonisation de nouveaux territoires après le dernier maximum glaciaire, comme cela a déjà été démontré chez de nombreuses espèces (Hewitt, 2000 ; Maley, 1991). Par ailleurs, la topologie des arbres phylogénétiques (Ramirez-Soriano *et al.*, 2008 ; Girod, 2010) et la structure du réseau d'haplotypes confirment l'expansion démographique de la population globale.

Spécifiquement, la caractéristique démographique de la zone sahélo-soudanienne est d'une population stable ayant une longue histoire évolutive ou le résultat d'un second contact de populations différenciées préalablement séparées (Maruyama et Fuerst, 1984 ; 1985 ; Grant et Bowen, 1998). Cette dernière inférence est appuyée par l'analyse des courbes mismatch. En effet, d'après Maltagliati *et al.* (2010), la bimodalité des distributions de mismatch serait la conséquence de la présence de différents haplogroupes et donc ne caractériserait pas une stabilité démographique. Tandis que les populations de la zone guinéo-congolaise, du Mali, du Cameroun et de la Centrafrique reflètent le résultat d'un signal d'un *bottleneck* démographique récent ou d'un effet fondateur (colonisation récente) par un petit nombre de lignées (Grant et Bowen, 1998 ; Avise, 2000). L'identification de cette colonisation récente d'insectes ravageurs des pays de l'Afrique centrale comme le Cameroun a été soulignée par Sézonlin *et al.* (2006) pour le foreur de tige et par Corrêa *et al.* (2016) avec le charançon du maïs.

Partant de ces résultats, on peut craindre deux menaces : celle d'une expansion des insectes, donc dépréciation massive des stocks de céréales et celle, très probable, d'émergence de souches résistantes aux méthodes de lutte, surtout la lutte chimique. Ceci serait favorisée par l'existence de cette forte diversité génétique au sein des populations mais aussi par l'importance des substitutions de types transversions. Selon Frankham (1996), la diversité génétique est positivement corrélée à la taille de la population ; plus la diversité génétique est importante, plus la taille de la population est grande. De plus, une forte diversité génétique est corrélée avec les traits de l'histoire de vie, telles que les espèces caractérisées par une stratégie R (MacArthur et Wilson, 1967 ; Galtier et Romiguier, 2015), produisant de nombreux descendants, comme les insectes. D'après Usher (2007), dans un environnement en évolution, l'écosystème nécessite un groupe d'individus ayant une constitution génétique variable. L'émergence de souches résistantes aux méthodes de lutte est suggérée par les taux de transversions qui sont plus élevés que ceux des transitions. Or dans l'ADN mitochondrial, les transitions peuvent représenter 90% des mutations (Kimura, 1980). Ces résultats évoqueraient donc une acquisition de la résistance car ce type de mutations entraîne un changement conformationnel des molécules (*Cyt-B* et *COI*) par le remplacement d'un noyau purine par un noyau pyrimidine, et vice-versa.

## CONCLUSION PARTIELLE

L'analyse de la distribution géographique de la diversité génétique dans les populations naturelles et l'explication des causes qui conduiraient à une corrélation ou non entre leur structure génétique et leur arrangement spatial constituent une base incontournable en biologie évolutive. En plus, l'étude de la structure génétique et phylogéographique est très importante car elle peut présenter un intérêt pratique pour la traçabilité des insectes nuisibles en fonction de leur diversité génétique et de leur origine géographique pour les stratégies de gestion.

Cette étude a révélé que les populations de *S. zeamais* présentent un haut niveau de polymorphisme avec des haplotypes communs et privés. La diversité génétique des individus d'une population concentre l'essentielle de la variance moléculaire tandis que celle entre régions naturelles est faible mais très significative. Il y a une forte diversité haplotypique, avec un nombre excessif d'haplotypes peu fréquents et quelque peu divergents. Ceci serait le résultat d'un flux multidirectionnel dû à l'importance des activités commerciales. En outre, le centre de l'infestation et de la dispersion se situerait dans la zone sahélo-soudanienne, plus particulièrement au Niger depuis le Pléistocène. Ainsi, les populations de *S. zeamais* sont caractérisées par une structure génétique sans signal phylogéographique. Une forte divergence génétique entre individus au sein des populations et des taux de transversions plus importants que ceux des transitions dans le génome mitochondrial peuvent être favorables à une émergence de souches résistantes aux méthodes de lutte particulièrement la lutte chimique. Les populations de *S. zeamais* peuvent donc être considérées comme ayant un pool génétique très important, ce qui offre un potentiel considérable d'adaptation et de survie. C'est un désavantage pour la sécurité alimentaire. En effet, il existe une compétition interspécifique pour les ressources végétales entre les insectes ravageurs et l'homme. D'où la nécessité de définir des politiques de surveillance des normes d'échanges et de coopération concernant les modalités d'échanges des denrées alimentaires pour atténuer voire supprimer l'infestation des cultures et denrées stockées.

### CHAPITRE III

## STRUCTURE GENETIQUE ET FLUX GENIQUE ENTRE POPULATIONS DE *Sitophilus zeamais* (MOTSCHULSKY, 1855), RAVAGEUR DE MAÏS STOCKE EN AFRIQUE DE L'OUEST

## **CHAPITRE III : STRUCTURE GENETIQUE ET FLUX GENIQUE ENTRE POPULATIONS DE *SITOPHILUS ZEAMAI*S (MOTSCHULSKY, 1855), RAVAGEUR DE MAÏS STOCKE EN AFRIQUE DE L'OUEST**

### **INTRODUCTION**

La génétique des populations étudie les fluctuations des fréquences des différentes versions d'un gène (allèles) au cours du temps dans les populations naturelles sous l'influence des forces évolutives que sont la sélection naturelle, la dérive génétique, les mutations, les migrations et les recombinaisons génétiques ; et cherche à expliquer l'adaptation et la spéciation (Fisher, 1930 ; Wright, 1931 ; Haldane, 1932). A ces forces s'ajoutent les forces écologiques qui peuvent être le système de reproduction, la sex-ratio, la démographie, la capacité de dispersion, la distance géographique, les ressources naturelles ; et les facteurs pluvio-climatiques et anthropiques. Cette discipline constitue, en fait, une matrice interdisciplinaire car réussissant à réconcilier la génétique mendélienne (hérédité des gènes) et la théorie darwinienne de l'évolution (évolution des gènes). De cette fusion prend racine une synthèse cohérente de conclusions qui ont donné naissance à la théorie synthétique de l'évolution nommée néodarwinisme (Dobzhansky, 1937 ; Huxley, 1942 ; Mayr, 1942) qui est à la base de toute la biologie moderne (Sembène, 2012b).

Par ailleurs, l'ensemble de ces contraintes constituent des pressions de sélection qui vont influencer la survie et le succès reproducteur des individus ; la source de l'évolution et la structuration de la diversité génétique. Dès lors, l'objectif de cette spécialité dans l'entomologie des insectes ravageurs de céréales, et donc dans le développement de pratiques de gestion durable des ressources agricoles, est d'investir l'impact spatio-temporal de la composante génétique dans le succès invasif des insectes ravageurs en mettant en exergue les facteurs génétiques, environnementaux et anthropogéniques qui gouverneraient la dynamique de ces populations. Or, ces informations cruciales restent encore à être évaluées pour la plupart des populations d'Afrique subsaharienne et l'approche indirecte *via* les marqueurs moléculaires proposée par cette discipline constitue une source d'informations permettant de répondre à ces questions d'importance capitale. De ce fait, le présent chapitre s'inscrit dans le contexte politico-économique de la sous-région ouest-africaine. En effet, elle se caractérise par une multitude d'organismes d'intégration politico- socio-économique facilitant ainsi les échanges commerciaux. En partant de ce contexte, nous émettons l'hypothèse qu'il existe un fort flux génique entre les populations du charançon du maïs sous l'impact des échanges commerciaux. Afin de vérifier cette hypothèse, nous a mené différentes activités, à savoir :

- estimer la variabilité et équilibre génétique des populations ;
- déterminer la différenciation et la structure génétiques des populations ;
- évaluer le flux de migrants entre populations ;
- identifier les facteurs impactant la structure génétique des populations.

### III.1 MATERIEL ET METHODES

#### III.1.1 ECHANTILLONNAGE

Nous avons échantillonné 15 individus issus des mêmes pays que dans le chapitre 2. Le choix de ces zones a été fait en fonction de l'importance de la production du maïs dans ces pays. Ce sont les populations du Burkina Faso, de la Côte d'Ivoire, du Ghana, du Niger, du Mali et du Sénégal qui ont été étudiées dans cette 2<sup>ème</sup> partie.

#### III.1.2 GENOTYPAGE DES MICROSATELLITES DE *SITOPHILUS ZEAMAI*S

##### III.1.2.1 INTERET ET LIMITES DES MICROSATELLITES

Les microsatellites sont particulièrement intéressants car, en raison de la petite taille de SSR, ils sont relativement stables dans l'ADN dégradé (Schneider *et al.*, 2004) et ils évoluent plus rapidement que d'autres portions du génome,  $10^{-6}$  à  $10^{-2}$  mutations par locus par génération et  $10^2$  à  $10^3$  fois plus vite que l'ADN nucléaire à copie unique (Ellegren, 2000a ; Li *et al.*, 2002 ; Thuillet *et al.*, 2002 ; Ellegren, 2004 ; Selkoe et Toonen, 2006 ; Bhargava et Fuentes, 2010). Ces taux leur permettent de constituer un outil puissant aussi bien sur une étude à échelle spatiale très restreinte que sur l'analyse des événements récents et contemporains (Ellegren, 2000b ; Giro, 2010). Les microsatellites nucléaires sont des marqueurs co-dominants (distinction entre homozygotes et hétérozygotes) hautement polymorphes, à hérédité biparentale (mendélienne) et avec une caractéristique majeure, la neutralité (Jarne et Lagoda, 1996 ; Chistiakov *et al.*, 2006). Une autre caractéristique intéressante des marqueurs de microsatellites est la présence des allèles privés (Corrêa *et al.*, 2012). Ils sont aussi très efficaces dans la détermination de la structuration intra- et inter-populations, de l'analyse de la généalogie ou du kinship (Queller *et al.*, 1993 ; Estoup, *et al.*, 1995 ; Jarne et Lagoda, 1996). L'utilisation des microsatellites en phylogéographie est très adaptée pour reconstituer les mécanismes à l'origine de la distribution génétique et géographique d'une espèce (Jarne et Lagoda, 1996 ; Koskinen *et al.*, 2002 ; Ellegren, 2004 ; Gum *et al.*, 2005). Par ailleurs, ils sont très résolutifs dans l'assignation d'individus (Zhang et Hewitt, 2003). Ils sont fréquemment utilisés pour déterminer l'histoire démographique des populations ; en effet, de nombreuses méthodes d'inférence ont été développées pour détecter à partir de la variation génétique actuelle observée dans les populations naturelles, des événements historiques tels que des contractions, des effets fondateurs ou des expansions (Excoffier et Heckel 2006 ; Girod *et al.*, 2011). Cependant, les microsatellites présentent des limites comme l'homoplasie ; ce qui peut fausser l'inférence en phylogénie des populations ou des espèces à étudier (Estoup *et al.*, 1995 ; Jarne et Lagoda, 1996 ; Orti *et al.*, 1997 ; Estoup *et al.*, 2002) ; d'où une sous-estimation de la diversité génétique et une surestimation de la différenciation génétique entre les populations influencées par la présence d'allèles nuls (Jarne et Lagoda, 1996 ; Orti *et al.*, 1997 ; Chapuis et Estoup, 2007). Certains locus microsatellites peuvent être sous sélection (Hendry *et al.*, 2000 ; Santiago et Caballero, 2005). Entre autre limite, la cherté du



processus de caractérisation des microsatellites qui est, en plus, longue.

### **III.1.2.2 CARACTERISATION DU PROCEDURE DE GENOTYPAGE DES POPULATIONS**

#### **III.1.2.2.1 EXTRACTION DE L'ADN GENOMIQUE**

L'extraction de l'ADN génomique des différentes populations a été réalisée au laboratoire de biologie moléculaire du Centre de Biologie des Populations Animales Sahélo-Soudaniennes (BIOPASS) de l'IRD de Bel-Air/ Dakar (Sénégal), suivant le code QIAGEN®. (Voir chapitre 2). Ensuite, les extraits d'ADN obtenus ont été conservés à -20°C jusqu'au génotypage.

#### **III.1.2.2.2 GENOTYPAGE DES LOCI DE *S. ZEAMIS***

Le génotypage des loci de *S. zeamais* a été effectué au plateau technique du CERAAS (Centre de Recherche d'Amélioration de l'Adaptation à la Sécheresse) de Thiès (Sénégal). Les amorces (Forward F et Reverse R) utilisés ont été développées et identifiées par Corrêa *et al.* (2012) (Tableau XIX). Nous avons adopté la méthode indirecte par une Queue-M13 fluorescente (Schuelke *et al.*, 2000) pour l'amplification par PCR. Pour les couples d'amorces sélectionnés, ce sont les amorces forwards qui sont flanquées à l'extrémité 5'- d'une amorce M13 composée de 23 pb (uni-43) (Tableau XIX) à savoir la séquence suivante : AGGGTTTTCCCAGTCACGACGTT (Venkatesan *et al.*, 2007). Pour un individu donné, le milieu réactionnel est ainsi constitué en plus des autres produits, 3 types d'amorces : de l'amorce Reverse (R), l'amorce Forward (F) flanquée à l'extrémité 5'- d'une amorce M13, d'un fluorochrome, généralement nommé, la Q-M13 ou Dye. Le volume réactionnel pour un individu et pour un marqueur donné est de 10 µl composés de 1 ou 2 µl de l'ADN-matrice avec 9 ou 10 µl du mix-PCR selon le marqueur. Le volume réactionnel du mix-PCR pour 90 individus est constitué de 501 µl d'eau Merq, 99 µl de tampon 10X, 20 µl de MgCl<sub>2</sub>, de 99 µl de dNTP, 8 µl de pour chacun des amorces Forward (F), flanquée à l'extrémité 5'- d'une amorce M13, Reverse(R), spécifiques à chaque locus, en plus du Dye (700 ou 800) et 50 µl de Taq polymerase. Les conditions d'amplification pour chacun des dix paires d'amorces sont : 4 min à 94°C pour une dénaturation initiale, suivi de 35 cycles de 94°C durant 30 s, 55°C pour 45 s et 72°C pendant 1 min et une élongation finale à 72°C pour 20 min, en utilisant un Vapo Protect Thermal Cycler. Les conditions de réaction sont : 95 °C for 5 min pour une dénaturation initiale ; 10 cycles de 94°C pendant 30 s, 55°C durant 1 min, et 72°C durant 30 s ; 27 cycles de 94 °C pendant 30 s, 55 °C durant 1 min et 72 °C pour 30 s, suivi par une extension de 10 min à 72°C (Venkatesan *et al.*, 2007). Les fiches PCR sont présentées à l'annexe 7.

**Tableau XIX** : Caractérisation des loci microsatellites nucléaires utilisés chez *S. zeamais* d'après Corrêa *et al.* (2012) adaptés avec l'amorce universelle M13 définie par Venkatesan *et al.* (2007)

N° de Loci	Loci	N° accession sur Genbank	Motif	Séquence des amorces (5'-3')	Taille de l'amplicon (Pb)
1	1E1	JX303735	(GT) <sub>7</sub>	F : CACCGATATACCTGAAAATA M13-F : AGGGTTTTCCAGTCACGACGTT-CACCGATATACCTGAAAATA R : AAAACGACAGATAGAACCTCGG	277-281
2	1A1	JX303736	(TG) <sub>6</sub>	F : CACCCTGTATATCGAGAGAACAT M13-F : AGGGTTTTCCAGTCACGACGTT-CACCCTGTATATCGAGAGAACAT R : TACAAAATTATCATCGCCCTC	235-253
3	1B1	JX303737	(GT) <sub>7</sub>	F : TAATGCGATTGTTGAGGATG M13-F : AGGGTTTTCCAGTCACGACGTT-TAATGCGATTGTTGAGGATG R : ACAACGATTTCTTTATCACCCG	149-161
4	1D10	JX303738	(ATC) <sub>6</sub>	F : ACATTCAGGAGTGGTCTTGGTC M13-F : AGGGTTTTCCAGTCACGACGTT-TAATGCGATTGTTGAGGATG R : CTCGGGATTAGCCAACAATAGA	249-261
5	1G7	JX303739	(GT) <sub>8</sub>	F : TTACCGAAATATCCAGTACCC M13-F : AGGGTTTTCCAGTCACGACGTT-TTACCGAAATATCCAGTACCC R : TCTATTGTGTCTATTACGACTT	105-109
6	3A11	JX303740	(TG) <sub>8</sub>	F : ATATTAGAGCGGGGCTGGTT M13-F : AGGGTTTTCCAGTCACGACGTT-ATATTAGAGCGGGGCTGGTT R : CTCCCTTTGGTGTGATGATG	241-313
7	3H6	JX303741	(TG) <sub>8</sub>	F : CAAGGAAGTCTGGATAAATGGT M13-F : AGGGTTTTCCAGTCACGACGTT-CAAGGAAGTCTGGATAAATGGT R : GTGACCTCGTGTCTTGGTAGT	284-318
8	3G1	JX303742	(CA) <sub>7</sub>	F : CCTGGTTTTGTTGATATGGAG M13-F : AGGGTTTTCCAGTCACGACGTT-CCTGGTTTTGTTGATATGGAG R : TTCTGGTGTGGATGTGAATGTA	140-152
9	3G7	JX303743	(AG) <sub>9</sub> (GA) <sub>6</sub>	F : GTTTCGTGATCCGGTTTTGT M13-F : AGGGTTTTCCAGTCACGACGTT-GTTTCGTGATCCGGTTTTGT R : GCGCTAGCCTGAGGAGAAT	187-193
10	1B10	JX303744		F : GGTCCGATCTCATTTGGAGTAG M13-F : AGGGTTTTCCAGTCACGACGTT-GGTCCGATCTCATTTGGAGTAG R : AGGCAAGATGGAACCTATCG	

### III.1.3 ANALYSES GENETIQUES ET STATISTIQUES

En génétique des populations, les approches statistiques ont deux objectifs distincts qui sont la description des données et la possibilité d'inférer les processus évolutifs qui ont généré les patrons observés (Blum, 2012). Dès lors, différents paramètres ont été étudiés dans les différents niveaux d'analyses au sein des sous-populations et à l'échelle de la population globale afin de quantifier le niveau de variabilité génétique au sein des populations et de différenciation génétique entre les populations, d'inférer sur la structure génétique des populations et la sur l'évolution démographique.

#### III.1.3.1 LECTURE, CORRECTION DES GENOTYPES ET CONSTITUTION DU JEU DE DONNEES

La lecture et la correction des génotypes ont été effectuées manuellement. La fréquence des allèles nuls a été estimée par Micro-checker ver. 2.2.3. (Van Oosterhout *et al.*, 2004).

#### III.1.3.2 VARIABILITE ET DIVERSITE GENETIQUES DES POPULATIONS

Six paramètres ont été utilisés pour quantifier et analyser la variabilité génétique pour chaque locus et chaque population.

##### III.1.3.2.1 NOMBRE MOYEN D'ALLELES PAR LOCUS (N<sub>A</sub>)

C'est en moyenne le nombre d'allèles présents chez chaque locus sur l'ensemble des individus.

$$N_A = \sum a / r$$

a : somme des allèles dans tout l'échantillon et r : nombre de loci étudiés.

### III.1.3.2.2 INDICE DE POLYMORPHISME (I.P.)

C'est le pourcentage de loci polymorphes dans l'échantillon étudié. En effet, c'est la probabilité d'observer au moins deux allèles au même locus et cette probabilité dépend des fréquences respectives des allèles et aussi de la taille de l'échantillon (Foulley et Ollivier, 2006a). Un locus est considéré comme polymorphe lorsque l'allèle le plus fréquent a une fréquence inférieure ou égale à 0,95.

$$\text{I.P.} = \text{Nombre de loci polymorphes} / \text{Nombre de loci étudiés}$$

### III.1.3.2.3 FREQUENCES ALLELIQUES ( $F_A$ )

Les fréquences alléliques sont calculées à partir des génotypes des individus étudiés dans une population. C'est le rapport entre le nombre de copies d'un allèle  $a$  dans la population sur le nombre total de copies de tous les allèles.

$$F_A = \frac{2 \times \text{nombre d'individus homozygotes pour l'allèle } i + \text{nombre d'individus hétérozygotes pour l'allèle } i}{2N}$$

$N$  : nombre total d'individus génotypés au locus K.

### III.1.3.2.4 TAUX D'HÉTÉROZYGOTIE

L'hétérozygotie ou diversité génétique de Nei (1973) caractérise le niveau de variabilité des gènes au sein d'une même population, c'est-à-dire qu'elle exprime la probabilité pour que deux variants (allèles) à un même locus tirés au hasard dans une population soient différents. D'une part, il y a l'hétérozygotie observée  $H_O$  (fréquence observée d'individus hétérozygotes), ou fréquence d'apparition d'hétérozygotes par locus dans l'échantillon, est le rapport du nombre d'individus hétérozygotes sur le nombre total d'individus génotypés pour un locus. C'est une mesure du polymorphisme. D'autre part, l'hétérozygotie attendue ou théorique ou indice de diversité ( $H_E$ ) correspond à l'hétérozygotie attendue sous les hypothèses de l'équilibre de Hardy-Weinberg (Nei, 1987) et l'hétérozygotie attendue non biaisée  $H_{nb}$  proposée par Nei (1978) lorsque le nombre d'individus testés est faible. Cette dernière est un bon estimateur de la variabilité génétique d'une population puisqu'elle est moins sensible aux aléas d'échantillonnage comme le nombre d'allèles observés par exemple (Nei, 1987). A partir de fréquences alléliques, l'hétérozygotie attendue non biaisée ( $H_{nb}$ ) et l'hétérozygotie observée ( $H_O$ ) ont été calculées pour chaque population et pour la population totale ; afin de tester si ces populations étaient dans les conditions du modèle d'Hardy-Weinberg ; c'est-à-dire  $H_E = H_O$  ou non et  $H_E <$  ou  $> H_O$ .

### III.1.3.2.5 CONTENU INFORMATIF DU POLYMORPHISME ET INDEX INFORMATIF DE SHANNON

Le contenu informatif du polymorphisme (PIC), aussi dénommé le potentiel en information du polymorphisme, permet d'évaluer la capacité informative (discriminante) d'un marqueur dans une population à partir des fréquences alléliques (Botstein *et al.*, 1980 ; Thiruvankadan *et al.*, 2014).

Effectivement, un locus codominant est très informatif si  $PIC > 0,5$  ; il est raisonnablement informatif si  $0,5 > PIC > 0,25$  ou n'est que légèrement informatifs si  $PIC < 0,25$  (Botstein *et al.*, 1980). Alors que l'index informatif de Shannon (I) offre des propriétés statistiques idéales pour mesurer l'information biologique à travers de multiple niveaux, des gènes aux paysages comme l'ont démontré Sherwin *et al.* (2006) et Rossetto *et al.* (2008) (Peakall et Smouse, 2012). Des valeurs élevées témoignent d'une forte diversité génétique.

### III.1.3.2.6 RICHESSE ALLELIQUE

La richesse allélique d'une population, définie comme étant le nombre d'allèles présents à un locus donné, dépend de la taille de l'échantillon, puisque les chances de découvrir un nouvel allèle augmentent chaque fois qu'un nouvel individu est observé (Foulley *et al.*, 2006b). Pour comparer la richesse allélique de 2 populations, il est donc nécessaire de corriger le nombre d'allèles observé en fonction de la taille de l'échantillon. Elle est calculée pour chaque population en utilisant une procédure de correction basée sur le principe de « raréfaction » (El Mousadik et Petit, 1996). La richesse allélique estimée est donc, au plus, égale au nombre d'allèles observés. La richesse allélique a été estimée à l'échelle de chaque population et de la population globale.

Ces paramètres fournissent une bonne estimation de la variabilité génétique à l'intérieur de chaque population et ils sont calculés grâce aux logiciels GENETIX ver. 4.05.2 (Belkhir *et al.*, 2004) et GenALEX ver. 6.5 (Peakall et Smouse, 2012). Les paramètres tels que l'index informatif de Shannon (I) et le potentiel en information du polymorphisme (PIC), les logiciels GenALEX et Cervus ver. 3.0.7 (Kalinowski *et al.*, 2007) ont été utilisés respectivement.

### III.1.3.3 EQUILIBRE GENETIQUE

#### III.1.3.3.1 TEST DE CONFORMITE A L'EQUILIBRE HARDY-WEINBERG

Une population dans laquelle les individus sont diploïdes et l'association de deux gamètes pour former les individus se fait au hasard par rapport aux génotypes de ces gamètes, est caractérisée par une structure génétique dénommée équilibre ou structure de Hardy-Weinberg (HWE, *Hardy-Weinberg Equilibrium*). Cette dernière est vérifiée par le test de HWE en utilisant l'indice de fixation  $F_{IS}$  qui caractérise l'écart à la panmixie (structure de Hardy-Weinberg) et mesure la réduction de l'hétérozygotie attendue sous hypothèse de la panmixie ( $F_{IS} = 0$ ). Il varie de -1 à +1 et permet de connaître le déficit en hétérozygotes par population, par locus et pour l'ensemble des loci.

$$F_{IS} = 1 - \frac{H_O}{H_E}$$

Lorsque la valeur du  $F_{IS}$  est positif ou négatif alors la population présente un déficit ou excès en hétérozygotes respectivement par rapport à l'équilibre panmictique. Autrement dit, les valeurs positives de  $F_{IS}$  indiquent que les individus d'une population sont plus liés que ce à quoi il fallait s'attendre avec un modèle HWE. Tandis que les valeurs de  $F_{IS}$  négatives indiquent que les individus d'une population sont moins liés qu'attendues.

L'écart à cette dernière peut être la résultante de la contribution de certains facteurs : consanguinité, dérive, sélection, effet Wahlund, *etc.*

### **NB : le coefficient de consanguinité individuel, $F_i$ (Coefficient inbreeding individual)**

Il est utilisé pour évaluer la force de la consanguinité. Nous avons déterminé le coefficient individuel de consanguinité ( $F_i$ ) qui est la probabilité d'identité par état (IBS, identical by state) entre les gènes au sein d'individus (Ritland 1996 ; Hardy 2003). Autrement dit,  $F_i$  est la valeur intra-individuelle du coefficient de parenté fourni par SPAGeDI (Vekemans & Hardy 2004). Par ailleurs, deux allèles sont définis comme étant identiques par descendance (IBD) s'ils sont identiques en raison d'un ancêtre commun au cours des dernières générations. En outre, il faut distinguer que deux allèles IBD sont IBS mais le réciproque n'est pas toujours vraie en raison des mêmes mutations chez des individus différents, par exemple. Ces paramètres ont été estimés par SPAGeDI, v. 1.4 (Hardy et Vekemans 2002).

### **III.1.3.3.2 TEST DE L'EQUILIBRE DE LIAISON**

Les allèles de différents loci s'associent de manière aléatoire sont dits en équilibre de liaison (*Linkage equilibrium*, LE) ; alors que ceux dont l'association n'est pas liée au hasard sont dits en déséquilibre de liaison (*Linkage Disequilibrium*, LD). Le déséquilibre de liaison (LD) est très fréquemment le résultat d'un événement de mutation ancestral et d'une succession de recombinaisons au cours des générations ; mais il est très rarement observé dans la nature. Le LD peut être la conséquence d'un contact récent entre populations, d'une sélection, d'une dérive génétique ou de mutations (Slatkin 1994 ; Cutter 2006). Il caractérise donc un écart à l'équilibre de Hardy-Weinberg. La présence du LD est cherchée par le test exact de Fisher dont l'hypothèse nulle est l'indépendance entre loci.

Le test sur l'équilibre Hardy-Weinberg pour chaque population et chaque locus, et sur l'équilibre de liaison entre paires de loci ont été examinés par des tests exacts de Fisher implémentés dans le logiciel GENEPOP ver. 4.7 (Rousset, 2008).

### **III.1.3.4 DIFFERENCIATION ET STRUCTURE GENETIQUES INTER-POPULATIONS**

D'après Yu *et al.* (2006), la structure d'un échantillon doit être abordée selon 2 aspects dont l'un fait appel au concept « structure des populations » dans l'idée que les individus échantillonnés

appartiennent les uns les autres à des populations différentes soumises à une sélection divergente ou par le fait que leur divergence est le fruit d'une adaptation localisée. Il y a par exemple des populations appartenant à des zones biogéographiques différentes. Alors que l'autre aspect se réfère au concept « apparemment ou *kinship* ». Dans ce dernier cas, les différents individus du spécimen ont des liens familiaux symbolisés par le partage de gènes (allèles ou haplotypes) hérités d'un ancêtre très ancien à récent. Le degré de différenciation entre populations dépend donc de manière étroite de l'ancienneté de la séparation des populations et des flux de gènes entre celle-ci.

### III.1.3.4.1 DIFFERENCIATION GENETIQUE ET FLUX GENIQUE

#### III.1.3.4.1.1 DIFFERENCIATION GENETIQUE DES POPULATIONS

Les  $F$ -statistiques (Wright, 1951 ; 1969 ; 1978 ; Weir et Cocherham, 1984 ; Holsinger et Weir, 2009) ou la théorie des indices de fixation (Nagylaki, 1998) permettent de décrire la répartition de la variabilité génétique entre les populations et à l'intérieur de ces populations ; de déterminer la structure génétique d'une population, d'une espèce. Ces  $F$ -statistiques (Wright, 1969) sont composées de 3 indices à savoir le  $F_{IS}$ , le  $F_{IT}$  et le  $F_{ST}$  correspondant respectivement aux  $F_{IS}$ ,  $F_{IT}$  et  $F_{ST}$  de Weir et Cockerham (1984). Ces indices mesurent l'écart de l'hétérozygotie par rapport à l'équilibre de Hardy-Weinberg à trois niveaux de complexité de l'hétérozygotie définis par Wright (1978) : les individus (I), les sous-populations (S) et la population globale (T). Ces indices permettent d'estimer dans quel compartiment la variabilité génétique se retrouve. Ils sont liés par la relation :

$$(1 - F_{IT}) = (1 - F_{IS}) * (1 - F_{ST})$$

Il faut noter aussi que si toutes les sous-populations constituant la population globale sont bien en équilibre panmictique, il y a :  $F_{IS} = 0$ , et donc  $F_{IT} = F_{ST}$ . Alors que si elles sont toutes en équilibre de Hardy-Weinberg et ont les mêmes fréquences alléliques, alors  $F_{IS} = F_{ST} = F_{IT} = 0$ . Cependant, si le  $F_{IS}$  et le  $F_{IT}$  sont différents, c'est qu'il existe une autre source de déficit en hétérozygotes, due à une différenciation entre les différents échantillons en sous-populations distinctes. Ce phénomène est connu sous le nom d'effet Wahlund (Wahlund, 1928).

En plus du  $F_{ST}$ , nous avons calculé, par paire de populations et pour la population globale, les autres indices de différenciation, à savoir le  $G_{ST}$ , le  $N_{ST}$  (Se référer au chapitre 2 au niveau du paragraphe II.1.3.3.3.4) et le  $R_{ST}$ .

En résumé, nous avons calculé les statistiques globales suivantes et sont basées sur :

- l'identité des allèles [ $F$ -statistiques (Weir et Cockerham, 1984),  $G$ -statistiques et  $D_s$  (distance génétique standard de Nei (1972) ou non biaisée de Nei (1978) où  $D_s = 2\mu t$  avec  $\mu$  comme étant le taux de mutation, et  $t$  est le nombre de générations depuis la divergence de la population)]. En plus de la

distance génétique de Nei (1972, 1978), il y a celle de Cavalli-Sforza et Edwards (1967). Celle de Nei (1972) suppose que les différences génétiques sont causées par la mutation et la dérive génétique alors que celle de Cavalli-Sforza et Edwards (1967) suppose que les différences génétiques surviennent en raison de la dérive génétique seulement.

- la taille des allèles microsatellites (*R*-statistiques ; Slatkin, 1995 ; Rousset, 1996 ; Michalakis et Excoffier, 1996).
- la phylogénie des allèles (*N<sub>ST</sub>*, basé sur une décomposition des indices de diversité suivant Pons et Petit (1996) où les populations ont un poids égal, quelle que soit la taille de l'échantillon).

Ces différents paramètres et tests ont été réalisés par les logiciels SPAGeDI et GENETIX.

Ainsi les indices de fixation de Wright sont calculés d'après les formules de Weir et Cockerham (1984) qui prennent en compte l'effet de la taille des échantillons utilisés pour estimer les fréquences alléliques et génotypiques.

Par ailleurs, les matrices de divergence et de différenciation génétiques ont été utilisées pour une analyse de *clustering* hiérarchique à l'aide des package *stats* et *pvclust* du logiciel R ver. 3.2.3 (R Core Team, 2015) afin de définir des *clusters* génétiques pour mettre en évidence la structure génétique des populations.

Par ailleurs, pour les analyses impliquant des tests multiples, la correction de Bonferroni (Rice, 1989) a été utilisée. Ainsi, les valeurs-seuils des *P-values* sont réévaluées et sont respectivement de 0,00833 [= 0,05/6 (nombre de pop.), pour l'analyse sur les populations] et 0,00625 [= 0,05/8 (nombre de locus), pour l'analyse sur les loci]. Alors que pour les tests par paire de populations comme le *F<sub>ST</sub>*, la valeur-seuil de la correction de Bonferroni est de : 0,0033 (= 0,05/ (6\*5/2)). Le paramètre *Nm* qui quantifie le nombre de migrants par génération par paire de populations a été calculé à l'aide du logiciel GENETIX et la valeur moyenne par GenALEX. D'après Wright (1978), il y a :

- une différenciation faible si :  $0 < F_{ST} < 0,05$  ;
- une différenciation modérée si :  $0,05 < F_{ST} < 0,15$  ;
- une différenciation importante si :  $0,15 < F_{ST} < 0,25$  ;
- une différenciation très importante si :  $F_{ST} > 0,25$ .

#### III.1.3.4.1.2 FLUX GENIQUE ET NOMBRE DE MIGRANTS PAR GENERATION

Le flux génique est calculé à partir des valeurs de *F<sub>ST</sub>* selon le modèle en île c'est-à-dire une population globale non soumise à la sélection subdivisée en sous-populations de tailles identiques échangeant un nombre de gènes identiques (Wright, 1951 et 1969) et à l'équilibre migration/dérive (Nei, 1987 ; Hartl et Clark, 1989) s'écrivant selon la formule suivante :

$$F_{ST} = \frac{1}{1 + \left(4N \left(\frac{s}{s-1}\right) (\mu + m)\right)} \quad \Rightarrow \quad F_{ST} = \frac{1}{1 + 4N\mu + 4Nm} \quad \Rightarrow \quad Nm = \left(\frac{1}{F_{ST}} - 1 - 4N\mu\right) / 4$$

Lorsque la mutation est négligeable devant la dérive génétique, on a :

$$F_{ST} = \frac{1}{1 + 4Nm} \quad \Rightarrow \quad Nm = ((1/F_{ST}) - 1)/4$$

Où  $N$  représente la taille efficace de la population ;  $\mu$ , le taux de mutation ;  $m$ , le taux de migration  $s$ , le nombre de populations considérées et enfin le  $Nm$  correspond au nombre de migrants par génération.

Ce paramètre a été estimé grâce à 2 méthodes :

- à partir du  $F_{ST}$  comme démontré précédemment  $Nm$  est calculé pour la population globale et pour chaque paire de populations selon l'approche de Weir et Cockerham (1984) grâce au logiciel GENETIX. Selon Whitlock et McCauley (1999), il faut être extrêmement prudent dans l'interprétation des niveaux inférés de flux de gènes.
- à l'aide de la méthode de l'allèle privé de Slatkin (1985) et de Barton et Slatkin (1986) dans le programme Genepop. En effet dans cette méthode, les fréquences moyennes des allèles privés sont utilisées pour estimer indirectement le flux génique. La logique derrière cette approche est que lorsque le flux de gènes est restreint, la probabilité qu'un migrant transporte un allèle rare entre les populations est faible. En conséquence, les fréquences des allèles privés rares vont augmenter dans les différentes sous-populations à des niveaux modérés. Inversement, lorsque le flux génétique est important, il y a une probabilité beaucoup plus élevée qu'un migrateur possède un allèle rare et, par conséquent, la fréquence des allèles privés dans les différentes sous-populations sera beaucoup plus faible. Cette méthode permet, en outre, restreint un ajustement par rapport à la taille des échantillons.

Par ailleurs, nous avons utilisé la fonction `divMigrate` selon la méthode de Sundqvist *et al.* (2016) du package `diveRsity` (Keenan *et al.*, 2013) du logiciel R ver. 3.2.3 (R Core Team, 2015) pour caractériser les réseaux de migration relative directionnelle avec les paramètres du  $Nm$  (Alcala *et al.*, 2014), du  $G_{ST}$  (Nei, 1973) et du  $D$  de Jost (Jost, 2008).

### III.1.3.4.2 FACTEURS DE DIFFERENCIATION ET DE STRUCTURATION GENETIQUES

#### III.1.3.4.2.1 ISOLEMENT PAR LA DISTANCE

L'isolement par la distance ou IBD (Isolement By Distance), est testée par une étude de la significativité de la pente de la régression de  $F_{ST}/(1-F_{ST})$  ou  $R_{ST}/(1-R_{ST})$  en fonction de la distance entre les populations (Rousset, 1997). L'hypothèse nulle postule que la pente de régression distance génétique/distance géographique est nulle. Une régression significative est le signal d'un isolement par la distance, c'est-à-dire que des populations géographiquement proches seraient plus proches génétiquement que les populations éloignées. Cette analyse a été réalisé à l'aide du logiciel R 3.4.3 (R Core Team, 2017) avec les packages `car` (Fox and Weisberg, 2011), `nlme` (Pinheiro *et al.*, 2017) et `ggplot2` (Wickham, 2016).



### III.1.3.4.2.2 SOURCES DE LA VARIANCE MOLECULAIRE

La structure et le niveau de différenciation des populations ont également été calculés à l'aide d'analyses hiérarchiques de Variance Moléculaire appelée AMOVA (Analysis Molecular VARIance) réalisées par le logiciel Arlequin. L'AMOVA décompose la variance totale en 4 composantes : qui sont attribuées aux différences entre les groupes définis ( $F_{CT}$ ), entre les populations à l'intérieur des groupes définis ( $F_{SC}$ ) et entre les individus au sein des populations ( $F_{IS}$ ) et au sein des individus dans la population globale ( $F_{IT}$ ). L'hypothèse nulle définit les individus comme appartenant à une même population génétique quelle que soit leur groupe ou leur population. La significativité ( $P < 0,05$ ) des hypothèses a été testée pour toutes les valeurs par 1023 permutations.

### III.1.3.4.2.3 IMPACT DES LOCI DANS LA DIFFERENCIATION GENETIQUE

Pour identifier les loci sous sélection dans la population globale, 3 approches ont été adoptées à savoir celle du serveur web de LOSITAN (Antao *et al.*, 2008) et des logiciels d'ARLEQUIN (Excoffier et Lischer, 2010), v. 3.5.2.2 et de BAYESCAN ver. 2.1 (Foll et Gaggiotti, 2008). Ces approches sont basées sur la méthode des  $F_{ST}$ -outliers (valeurs de  $F_{ST}$  atypiques, présentant des valeurs extrêmes) pour le model des loci neutres. Les analyses estiment la probabilité qu'un locus puisse être sous sélection par des valeurs de  $F_{ST}$  et d' $H_E$  observés comparées aux distributions neutres simulées (LOSITAN et ARLEQUIN) alors que BAYESCAN prend en compte les valeurs de  $F_{ST}$  en fonction du log (10) des probabilités postérieures. En effet, d'après Foll et Gaggiotti, (2008), BAYESCAN est robuste lorsqu'il examine des scénarii démographiques complexes pour la différenciation génétique neutre. Cette méthode bayésienne améliorée infère directement la probabilité *a posteriori* de chaque locus d'être sous l'effet de la sélection en définissant et en comparant deux modèles alternatifs, c'est-à-dire sélection (M1) et neutralité (M2) (Foll et Gaggiotti, 2008). Dans le contexte des tests multiples, nous devons également intégrer notre scepticisme quant à la chance que chaque locus soit sous sélection. Ceci est fait dans BayeScan en fixant les cotes antérieures pour le modèle neutre  $P(M1) / P(M2)$  et en utilisant les cotes postérieures (PO) au lieu des facteurs de Bayes pour prendre des décisions :

$$PO = P(M2 | N) / P(M1 | N) = BF * P(M2) / P(M1)$$

Les cotes postérieures sont simplement le rapport des probabilités a posteriori et indiquent à quel point le modèle avec sélection est plus susceptible d'être comparé au modèle neutre (l'échelle de preuves de Jeffrey peut aussi être utilisée avec des cotes postérieures).

Un gros avantage des probabilités postérieures est qu'elles permettent directement le contrôle du FDR (False Discovery Rate). Le FDR est défini comme la proportion attendue de faux positifs parmi les marqueurs de valeurs aberrantes. Jeffreys (1961) a proposé une échelle logarithmique pour le choix

du modèle définie comme : > 3 substantielle ( $\log_{10}PO > 0,5$ ) ; > 10 forts ( $\log_{10}PO > 1,0$ ) ; > 32 très fort ( $\log_{10}PO > 1,5$ ) ; et > 100 preuves décisives pour accepter un modèle ( $\log_{10}PO > 2.0$ ) (Foll, 2012).

Par suite, dans nos analyses sous BAYESCAN, un seuil pour  $\log_{10}PO > 10$  (fort) a été utilisé comme un indice. En effet, tout locus détecté sous ce seuil va être considéré comme un locus potentiellement candidat à la sélection positive, négative ou balancée. Alors que sous LOSITAN et ARLEQUIN, parmi les valeurs, une valeur « aberrante », est présentée comme une valeur  $F_{ST}$  inférieure à celle attendue dans la neutralité, suggérant que le locus a potentiellement fait l'objet d'une sélection balancée ; et celle se présentant avec une valeur de  $F_{ST}$  plus élevée que prévu sous la neutralité, correspond à un locus potentiellement influencé par la sélection directionnelle.

Pour les résultats obtenus avec Arlequin, chaque point représente un marqueur SSR. Les points rouges sont classés comme valeurs « *outliers* » potentiellement sous sélection. Pour ceux avec Bayescan, la ligne verticale indique le PO critique utilisé pour identifier les marqueurs de valeurs « *outliers* » (aberrantes). Le marqueur du côté droit de la ligne verticale est candidat à une sélection positive.

#### III.1.3.4.2.4 DETECTION DE SIGNAL PHYLOGEOGRAPHIQUE

Pour détecter du signal phylogéographique, en plus de la méthode de Burban *et al.* (1999) sur les valeurs du  $N_{ST}$  et du  $G_{ST}$  (cf II.1.3.3.3.4), des tests ont été faits à travers une série de permutations de l'identité des allèles suivis de la comparaison des valeurs du  $R_{ST}$  et du  $F_{ST}$  obtenues après permutations avec la valeur observée (Hardy *et al.*, 2003).

#### III.1.3.5 CLUSTERING, ORGANISATION SPATIALE ET INTROGRESSION DANS LES POPULATIONS

En général, les populations définies *a priori* pour estimer les paramètres génétiques sont identifiées en fonction de la proximité géographique des individus. Alors que des individus proches géographiquement peuvent correspondre à des populations génétiques différentes et une même population génétique peut être représentée par des individus à répartition géographique très large. Pour pallier à ce biais, plusieurs méthodes sont proposées. Ces méthodes dites de « *clustering* » qui sont des analyses de classification permettant de grouper des individus en populations homogènes définies selon leur génotype ; ces populations sont appelés *clusters* .

Nous avons adopté 2 types d'analyse, à savoir l'analyse multivariée et l'analyse bayésienne des génotypes multilocus des individus. Les allèles étant les variables. Il faut souligner que l'une des différences fondamentale entre ces 2 types d'analyse est que la deuxième repose sur un modèle génétique explicite de la population (HWE) alors que le premier est exploratoire et donc plus flexible car ne nécessitant pas d'hypothèse sur le modèle génétique de la population à étudiée.

##### III.1.3.5.1 ANALYSE MULTIVARIEE : AFC, ACP ET DAPC

Nous avons utilisé 3 types d'analyse multivariée : une analyse factorielle de correspondance (AFC), une Analyse en Composantes Principales (ACP) et une analyse discriminante des composantes

principales (DAPC pour Discriminant Analysis of Principal Components). Ces analyses nous ont permis d'évaluer le niveau de différenciation entre populations et de décrire les *clusters* génétiques afin de représenter graphiquement les relations génétiques entre individus/populations. L'exploration de ces relations est réalisée grâce aux analyses de l'AFC, de l'ACP et du DAPC à l'aide des logiciels GENETIX ver. 4.05.2 (Belkhir *et al.*, 2004), GenALEX ver. 6.5 (Peakall et Smouse, 2012) et le serveur du DAPC utilisant le package ADEGENET (Jombart, 2008 ; Jombart and Ahmed, 2011). L'AFC est effectué à partir des fréquences alléliques. Elle permet de mettre en exergue une structure génétique due à un effet Wahlund, d'identifier des éventuels d'admixture génétiques entre populations. Les trois premiers axes de l'AFC étant les plus informatifs, représentent un diagramme de dispersion tridimensionnel.

- Ensuite, une ACP (plus exactement un PCoA pour *Principal Coordinates Analysis* ; Jolliffe et Cadima, 2016) est construite à partir de la matrice de Covariance de la distance génétique non biaisée ( $D_S$  ; Nei, 1978), laquelle est déduite des fréquences alléliques. Elle permet de visualiser graphiquement les similarités ou dissemblances entre les individus/populations. Le PCoA est généralement bien adapté aux marqueurs génétiques car plusieurs distances génétiques peuvent être utilisées pour résumer la variabilité génétique (Jombart *et al.*, 2009).
- Enfin, nous avons adopté une nouvelle approche appelée analyse de discrimination des composantes principales (DAPC, pour *Discriminant Analysis of Principal Components*) (Jombart *et al.*, 2010). Cette méthode vise à identifier et à décrire les *clusters* génétiques et les relations entre ces *clusters*. Dans cette approche, la variance de l'échantillon est divisée en deux composantes (intergroupe et intragroupe) ; et le but est de maximiser la discrimination entre les *clusters*. Il y a d'abord la transformation des données (simplification des données) à l'aide d'une analyse en composantes principales (ACP ou PCA, *Principal Components Analysis*) comme une étape préalable, ensuite les *clusters* génétiques sont identifiés à l'aide d'une analyse discriminante (DA, pour *Discriminant Analysis*) sur les composants principaux conservés. Une des différences entre PCA (PCoA) et DAPC est que la PCoA vise à résumer la variabilité globale parmi les individus, qui inclut à la fois la divergence entre les *clusters* (c'est-à-dire, la variabilité génétique structurée), et la variation au sein des *clusters* (variabilité génétique "aléatoire") alors que la DAPC met l'accent sur la variabilité entre les *clusters*, tout en négligeant la variation intra-groupe (Jombart *et al.*, 2010). La méthode (DAPC) réalise donc la meilleure discrimination des individus en groupes prédéfinis et permet également une affectation probabiliste des individus à chaque groupe, comme dans les méthodes de *clustering* bayésien.

### III.1.3.5.2 ANALYSE BAYESIENNE

- **ALGORITHME STRUCTURE** (Pritchard *et al.*, 2000) : ce type d'analyse bayésienne sur la recherche de la structuration des populations a été réalisé grâce au logiciel STRUCTURE ver. 2.3.4 (Pritchard *et*

*al.*, 2000) qui considère que toutes les populations inférées sont considérées être à l'équilibre de Hardy-Weinberg et que les loci au sein des populations sont considérées être à l'équilibre de liaison. Il a pour but de caractériser la structure de populations à travers l'assignation d'individus à des populations où leur génotype a la plus forte probabilité de se trouver et par conséquent la détection du nombre réel de populations ; et l'identification de migrants, de zones et/ou d'individus hybrides ou « admixés » (estimation du degré d'introgression/hybridation). Le principe adopté est celui d'une admixture ou introgression parmi les quatre modèles (chaque individu tire une certaine fraction de son génome de chacune des K populations détectées ; c'est une méthode de *clustering* quantitative) et une corrélation des fréquences alléliques qui postule l'hypothèse d'une population ancestrale unique d'où auraient divergé les différentes populations actuellement observées. Ce modèle est bien cadré avec ceux des populations d'insectes dont la dispersion naturelle est intensifiée par les activités anthropiques. L'approche bayésienne permet d'estimer le nombre le plus probable (K représente le nombre de regroupements «à priori» à identifier avec  $2 \leq K \leq 10$ ) génotypes multiloci dans l'échantillonnage étudié et la probabilité (Q) d'assignation de chaque individu à chacune des K populations avec les 10 loci microsatellites étudiés. K allant de 1 à 10 avec 5 répétitions pour chacun afin de tester la reproductibilité des résultats. La méthode repose sur une simulation de Monte Carlo par chaîne de Markov (MCMC) pour déterminer le nombre de génotypes multilocus adéquat K qui maximise la vraisemblance des données génétiques via 2 procédures qui présentent chacune des avantages respectives pour inférer quant à la structure des populations. En effet, il y a la méthode du plateau : la meilleure structure est celle où la valeur de K est maximale avant le plateau de la courbe représentant  $\ln P(D)$  en fonction de K (Pritchard *et al.*, 2000). La deuxième méthode est basée sur la variation de second ordre du log-likelihood (Evanno *et al.*, 2005). Cette méthode assigne aussi par probabilité des individus génotypés à chacun des K génotypes multilocus identifiés. De ces analyses, des fichiers de résultats bruts vont être générés et importés sous le programme en ligne STRUCTURE Harvester ver. 0.6.94 (Earl et VonHoldt, 2012) : <http://taylor0.biology.ucla.edu/structureHarvester/>. Par ailleurs, un individu représenté par une barre unicolore correspondra donc à un individu assigné à un unique groupe, alors qu'un individu représenté par une barre multicolore présentera un génotype multilocus qui provient de plusieurs *clusters* dans les proportions représentées par les différentes couleurs et donc il indiquerait un processus d'admixture.

- **MIGRANTS DE PREMIERE GENERATION** : nous avons aussi utilisé le logiciel GeneClass ver. 2.0 (Piry *et al.*, 2004) pour détecter le nombre de migrants de première génération suivant d'une part la méthode des fréquences alléliques de Paetkau *et al.* (1995) avec un seuil d'affectation des scores fixé à 1%. D'autre part, les tests ont évalué la proportion de génotype de chaque individu qui est issu d'une des populations de l'échantillon. Les assignations des individus ont été faites par le calcul de la vraisemblance (Paetkau *et al.*, 1995) et des assignations bayésiennes partielles (Rannala et Mountain, 1997) (100000 répétitions, seuil de détection des migrants :  $P < 0,01$ ).

- **ASSIGNATION DES INDIVIDUS A LEUR POPULATION SOURCE** : l'assignation génétique des individus a été effectuée avec le programme GenALEX. L'option "Likelihoods Positive" permet la conversion du log-likelihoods en valeurs positives. Les valeurs les plus petites indiquent la population la plus vraisemblable. Le seuil de la valeur de P est fixé à 0,01 suivant la recommandation de Paetkau *et al.* (2004). Si la probabilité de l'individu est inférieure au seuil établi pour cette population, alors l'individu n'est pas originaire de cette population. Donc, un individu est considéré comme étant correctement assigné à une population lorsqu'il est exclu de toutes les autres avec une probabilité hautement significative ( $p < 0,001$ ), sauf celle dont il est originaire.

### III.1.3.5.3 IMPACT DE LA MUTATION/FLUX GENIQUE SUR LA DIFFERENCIATION GENETIQUE

Hardy *et al.* (2003) ont étudié l'influence de la mutation sur la différenciation génétique afin de déterminer l'importance relative des mutations par rapport aux flux de gènes dans la mesure de la différenciation génétique des populations en comparant les valeurs de  $F_{ST}$  et  $R_{ST}$ . En effet, l'influence de la mutation est généralement considérée comme négligeable dans la mesure de la structuration génétique à l'exception des marqueurs microsatellites qui présentent de forts taux de mutation. Cette comparaison s'effectue en excluant les individus présentant des allèles dont les tailles étaient incompatibles avec un modèle de mutation en SSM (Stepwise Mutation Model) qui est le modèle de mutation sur lequel repose le test qui a été réalisé grâce au logiciel SPAGeDI (Spatial Pattern Analysis of Genetic Diversity) ver. 1.4. (Hardy et Vekemans 2002). Ce dernier met en œuvre une procédure de randomisation qui permet de tester la contribution de la mutation génétique à la différenciation. Le  $R_{ST}$ , dépendant du taux de mutation des loci, est une fraction de la variance totale existant entre populations et qui prend en compte la différence de taille entre les allèles ; alors que le  $F_{ST}$  est un estimateur de la différenciation génétique qui renseigne sur le niveau de flux de gènes. Par conséquent, une valeur de  $F_{ST}$  égale au  $R_{ST}$  signifie que la mutation n'a pas joué de rôle dans la différenciation génétique des populations. A l'inverse, une valeur de  $R_{ST}$  significativement supérieure au  $F_{ST}$  est indicatrice d'une influence de la mutation dans la différenciation génétique qui n'est pas seulement due aux flux de gènes ; la taille des allèles contribue de manière significative à la différenciation de la population. Ce qui suggère un processus de mutation par étape (Girod, 2010).

### III.1.3.6 ANALYSES PHYLOGENETIQUES

L'arbre phylogénétique est un outil puissant, permettant de synthétiser une très grande quantité d'information en une seule image (Mahé, 2009). Les arbres phylogénétiques ont été explorés grâce à 2 méthodes : Par agglomération des plus proches voisins du Neighbor-Joining (NJ) (Saitou and Nei, 1987) qui a généré les arbres des individus et des populations construits à partir de deux matrices : celle de la distance génétique angulaire (Nei *et al.*, 1983) ou  $D_A$  issue de la modification de Cavalli-Sforza et Edwards (1967) et de la distance des allèles partagés ou *DAS* (*model-free shared allele*

*distances*) de Chakraborty et Jin (1993) au moyen des logiciels POPULATIONS.exe (Langella, 2002) et Figtree (Rambaut, 2014). La distance des allèles partagés *DAS* est recommandée pour des populations fortement apparentées ainsi que pour l'analyse des microsatellites de l'ADN nucléaire d'une manière générale (Nei *et al.*, 1983). La distance génétique angulaire  $D_A$  est une mesure appropriée pour l'analyse des arbres phylogénétiques indépendamment de la présence de *bottleneck* démographique (Choi *et al.*, 2012) et lorsque la dérive génétique est le premier facteur de la différenciation évolutive (Maudet *et al.*, 2002).  $D_A$  fait partie des mesures de distance les plus largement utilisées car il a été rapporté qu'elle augmente lentement avec le temps et maintient la relation linéaire pour de longues périodes (Nei, 1987). La robustesse des nœuds est estimée par le pourcentage de 1000 bootstraps (ré-échantillonnage aléatoire des loci à plusieurs reprises).

### III.1.3.7 EVOLUTION DEMOGRAPHIQUE

Nous allons utiliser deux méthodes pour caractériser l'évolution démographique. La première concerne les statistiques résumées pour déterminer l'évolution démographique des populations échantillonnées. Cette méthodologie dont le fondement est le calcul d'une statistique, résumant l'information disponible dans les données génétiques, qui pourra faire l'objet d'un test permettant de détecter le changement démographique (Girod, 2010). Il y a la méthode utilisant le logiciel BOTTLENECK ver 1.2.02 (Cornuet et Luikart, 1997) qui permet la détection d'une réduction massive et récente de l'effectif d'une population (ou goulot d'étranglement). Ainsi, dans une population ayant subi une réduction d'effectifs récente, la réduction de la diversité allélique sera plus rapide que la réduction de l'hétérozygotie attendue sous l'équilibre de Hardy-Weinberg ( $H_o > H_e$ ) sous 3 modèles mutationnels : l'IAM, le SMM et le TPM (Cornuet et Luikart, 1996). En plus, d'après ces derniers, le test le plus puissant et robuste pour tester un goulot d'étranglement est le Wilcoxon ; ce test est le plus adapté et le plus précis dans le cas où le nombre de loci polymorphes est faible (Luikart et Cornuet, 1998). Ensuite, nous avons utilisé la méthode de la représentation graphique de l'indicateur « Mode-shift » de Luikart *et al.* 1998. Cette méthode implémente un descripteur qualitatif de la distribution des fréquences alléliques ("mode-shift" Indicator) permettant de discriminer les populations en bottleneck des populations stables.

## III.2 RESULTATS

### III.2.1 SCREENING DES LOCI MICROSATELLITES

Des problèmes d'amplification ont été constatés pour le marqueur 1G7. Il a présenté un patron d'amplification trop complexe pour être lu sans ambiguïté. C'est probablement lié à l'amplification simultanée de 2 loci aux tailles alléliques très proches. Le locus 1B10 est monoallélique comme trouvé par Corrêa *et al.* (2012) a été confirmée. Ces 2 marqueurs ont donc été exclus dans l'ensemble des analyses. Notre jeu de données a donc été constitué de 90 individus analysés avec 8 marqueurs microsatellites. La taille des allèles est donnée dans le Tableau XX.

Les résultats de l'analyse de Micro-checker ont identifié 3 loci 3G7 (Oosterhout = 0,3444, Burkina Faso ; Oosterhout = 0,415 ; Côte d'Ivoire) 1E1 (Oosterhout = 0,193 ; Ghana) et 3H6 (Oosterhout = 0,375 ; Ghana) comme potentiellement affectés par la présence d'allèles nuls (Tableau XX). La détection d'allèles nuls potentiels chez ces deux derniers s'est faite uniquement dans une population. En plus, pour ces 3 loci, les écarts à l'équilibre d'Hardy-Weinberg à l'échelle globale ont été non significatifs.

**Tableau XX** : Détection des allèles nuls, de la taille et du nombre d'allèles par locus.

Populations Locus	Burkina Faso	Mali	Niger	Sénégal	Côte d'Ivoire	Ghana	$F_{IS}$	Taille des allèles N (nombre d'allèles)
1A1	-	-	-	-	-	-	***	N (7) : 147-247-250-255-258-262-265
1E1	-	-	-	-	-	nul	NS	N (4) : 292-295-297-299
1B1	-	-	-	-	-	-	**	N (4) : 148-152-168-172
3G7	nul	-	-	-	nul	-	NS	N (4) : 185-200-208-212
3G1	-	-	-	-	-	-	NS	N (4) : 162-166-168-171
3H6	-	-	-	-	-	nul	NS	N (3) : 318-321-325
3A11	-	-	-	-	-	-	**	N (4) : 260-261-319-321
1D10	-	-	-	-	-	-	NS	N (3) : 265-269-274

### III.2.2 DIVERSITE ET EQUILIBRE GENETIQUES DES POPULATIONS

L'analyse des tests d'équilibre de Hardy-Weinberg ont révélé que seulement 3 loci (1E1, 3G1 et 1D10) sont en équilibre génétique contre 5 qui ont dévié l'équilibre de Hardy-Weinberg (1B1 et 3A11, 1A1, 3G7 et 3H6) (Tableau XXI). Cependant, les loci ont globalement dévié l'équilibre de Hardy-Weinberg avec un taux d'hétérozygotie attendue non biaisé qui est supérieure à celui observé ( $H_{nb}/H_O=0,526/0,555$ ) mais non significatif ( $P=0,5823$ ). A cela s'ajoute une valeur de  $F_{IS}$ , permettant de mesurer le degré de consanguinité des individus issus au sein des populations, qui est significativement négative ( $F_{IS} = -0,0626$ ). Toutes les estimations multilocus ont indiqué un écart significatif à la loi de Hardy-Weinberg, reflétant que toutes les populations, en plus de la population globale, ont dévié la loi de Hardy-Weinberg (Tableau XXI). En outre, les  $P$ -values de l'équilibre de liaison génotypique entre paires de loci sont supérieures au niveau de significativité (Tableau XXII).

**Tableau XXI : Indices des caractéristiques génétiques par locus microsatellite et par population**

Loci		1A1	1E1	1B1	3G7	3G1	3H6	3A11	1D10	Multilocus
Variation de la taille allélique	Indices de variabilité	147-265	292-299	148-172	185-212	162-171	318-325	260-321	265-274	-----
Burkina Faso	N	15	15	15	15	15	15	15	15	15,000
	N <sub>A</sub>	6	4	4	4	3	2	4	3	3,750
	PIC	0,731	0,676	0,615	0,411	0,393	0,365	0,613	0,572	0,5469
	I	1,597	1,336	1,227	0,869	0,770	0,673	1,215	1,068	1,094
	R <sub>A</sub>	5,762	3,992	3,963	3,837	2,896	2,000	3,896	3,000	---
	H <sub>ab</sub>	0,791	0,752	0,697	0,453	0,467	0,497	0,692	0,669	0,627
	H <sub>O</sub>	0,667	0,733	0,933	0,067	0,267	0,800	0,933	0,600	0,625
	F <sub>IS</sub>	0,1617 *	0,0253 NS	-0,3564 *	0,8571 *	0,4372 NS	-0,6471 NS	-0,3659 NS	0,1064 NS	0,003 NS
Test d'EHW/Chi2										**
Niger	N	10	15	15	13	15	10	15	12	13,125
	N <sub>A</sub>	4	4	4	2	2	3	3	3	3,125
	PIC	0,389	0,549	0,465	0,132	0,269	0,482	0,460	0,533	0,4100
	I	0,826	1,107	0,938	0,271	0,500	0,927	0,889	1,005	0,808
	R <sub>A</sub>	4,000	3,867	3,333	1,954	2,000	3,000	2,897	3,000	---
	H <sub>ab</sub>	0,437	0,632	0,582	0,148	0,331	0,595	0,577	0,627	0,491
	H <sub>O</sub>	0,500	0,733	1,000	0,000	0,267	0,400	1,000	0,500	0,550
	F <sub>IS</sub>	-0,1538 NS	-0,1667 NS	-0,7647 *	1,0000 NS	0,2000 NS	0,3394 NS	-0,7797 *	0,2096 NS	-0,124 NS
Test d'EHW/Chi2										*
Mali	N	15	15	14	15	15	12	15	15	14,500
	N <sub>A</sub>	3	3	4	2	3	1	3	3	2,750
	PIC	0,578	0,414	0,595	0,375	0,478	0,000	0,421	0,562	0,4277
	I	1,076	0,802	1,188	0,693	0,918	0,000	0,816	1,052	0,818
	R <sub>A</sub>	3,000	2,667	3,923	2,000	2,970	1,000	2,667	3,000	---
	H <sub>ab</sub>	0,674	0,536	0,680	0,517	0,577	0,000	0,549	0,660	0,524
	H <sub>O</sub>	0,733	0,533	0,786	1,000	0,400	0,000	1,000	0,533	0,623
	F <sub>IS</sub>	-0,0922 *	0,0044 NS	-0,1626 NS	-1,0000 *	0,3143 NS	NA	-0,8750 *	0,1971 NS	-0,197 *
Test d'EHW/Chi2										*
Sénégal	N	15	15	15	15	15	14	15	15	14,875
	N <sub>A</sub>	4	4	3	2	2	3	2	3	2,875
	PIC	0,659	0,497	0,309	0,375	0,365	0,561	0,375	0,175	0,4146
	I	1,302	1,003	0,628	0,693	0,673	1,051	0,693	0,389	0,804
	R <sub>A</sub>	3,970	3,637	2,889	2,000	2,000	3,000	2,000	2,563	---
	H <sub>ab</sub>	0,738	0,577	0,349	0,517	0,497	0,659	0,517	0,191	0,506
	H <sub>O</sub>	0,600	0,667	0,400	1,000	0,533	0,429	1,000	0,133	0,595
	F <sub>IS</sub>	0,1923 NS	-0,1618 NS	-0,1507 NS	-1,0000 *	-0,0769 NS	0,3580 NS	-1,0000 *	0,3086 NS	-0,185 *
Test d'EHW/Chi2										*
Côte d'Ivoire	N	13	15	15	15	15	10	15	15	14,000
	N <sub>A</sub>	5	2	4	3	2	2	2	3	2,750
	PIC	0,574	0,346	0,699	0,440	0,365	0,365	0,375	0,486	0,4563
	I	1,216	0,637	1,379	0,853	0,673	0,673	0,693	0,933	0,882
	R <sub>A</sub>	4,721	4,721	4,721	4,721	4,721	4,721	4,721	4,721	--
	H <sub>ab</sub>	0,543	0,431	0,524	0,000	0,480	0,416	0,517	0,638	0,444
	H <sub>O</sub>	0,615	0,400	0,933	0,000	0,267	0,200	1,000	0,667	0,427
	F <sub>IS</sub>	0,0352 NS	0,1340 NS	-0,2174 *	1,0000 *	0,4717 NS	0,6170 NS	-1,0000 *	-0,1336 NS	0,102 NS
Test d'EHW/Chi2										*
Ghana	N	15	14	15	15	15	11	15	12	14,125
	N <sub>A</sub>	3	4	4	1	3	2	2	3	2,875
	PIC	0,440	0,391	0,432	0,000	0,383	0,318	0,375	0,535	0,3594
	I	0,853	0,832	0,877	0,000	0,745	0,586	0,693	1,011	0,700
	R <sub>A</sub>	2,897	3,835	3,333	1,000	2,667	2,000	2,000	3,000	---
	H <sub>ab</sub>	0,637	0,460	0,772	0,543	0,497	0,505	0,517	0,591	0,565
	H <sub>O</sub>	0,400	0,214	0,667	0,000	0,467	0,000	1,000	0,667	0,510
	F <sub>IS</sub>	0,2696 NS	0,5125 NS	-0,2844 NS	NA	0,0297 NS	1,0000 *	-1,0000 *	-0,0476 *	0,041 NS
Test d'EHW/Chi2										*
Population globale	N	13,833	14,833	14,833	14,667	15,000	12,000	15,000	14,000	14,271
	N <sub>A</sub>	7	4	4	4	4	3	4	3	3,021
	PIC	0,627	0,565	0,588	0,677	0,403	0,421	0,465	0,557	0,538
	I	1,357	1,138	1,179	1,341	0,802	0,819	0,937	1,044	0,851
	R <sub>A</sub>	4,614	3,720	3,804	3,968	2,598	2,707	3,017	2,994	3,4278
	H <sub>ab</sub>	0,636	0,565	0,601	0,363	0,475	0,445	0,562	0,562	0,526
	H <sub>O</sub>	0,586	0,547	0,787	0,344	0,367	0,305	0,989	0,517	0,555
	F <sub>IS</sub>	0,127 NS	0,127 NS	-0,220 **	0,519 **	0,267 *	0,345 *	-0,757 **	0,193 *	0,084 **
Test d'EHW/Chi2	**	NS	*	*	NS	*	*	NS	**	

N (nombre d'individus typés par locus) ; N<sub>A</sub> (nombre d'allèles observés par locus) ; PIC (Contenu Informatif du Polymorphisme) ; I (Index informatif de Shannon) R<sub>A</sub> (richesse allélique par locus) ; H<sub>ab</sub> (hétérozygotie attendue non biaisée) ; H<sub>O</sub> (hétérozygotie observée) ; F<sub>IS</sub> (f) (coefficient de consanguinité de Weir and Cockerham, 1984) ; NS: F<sub>IS</sub> non significatif ; \*: F<sub>IS</sub> significatif ; PIC > 0,50 : hautement informatif ; 25 < PIC < 0,50 : modérément informatif ; PIC < 0,25 : peu informatif ; PIC = 0 : non informatif (selon Botstein, 1980) Test d'EHW/Chi2 : Test d'Hady-Weinberg basé sur la méthode du Chi 2.



**Tableau XXII:** Analyse de l'équilibre de liaison entre génotypes par paire de loci

Paires de locus		Chi2	Degrès de liberté	<i>P-value</i>
1A1	1E1	10,584	12	0,565
1A1	1B1	15,648	12	0,208
1E1	1B1	18,448	12	0,103
1A1	3G7	6,905	6	0,330
1E1	3G7	4,371	6	0,627
1B1	3G7	5,897	6	0,435
1A1	3G1	5,109	12	0,954
1E1	3G1	19,535	12	0,076
1B1	3G1	10,916	12	0,536
3G7	3G1	0,877	6	0,990
1A1	3H6	10,498	10	0,400
1E1	3H6	8,043	10	0,625
1B1	3H6	6,245	10	0,794
3G7	3H6	4,392	4	0,356
3G1	3H6	13,245	10	0,210
1A1	3A11	9,545	6	0,145
1E1	3A11	8,816	6	0,184
1B1	3A11	4,735	6	0,578
3G7	3A11	2,886	4	0,577
3G1	3A11	12,446	6	0,053
3H6	3A11	1,024	4	0,906
1A1	1D10	12,062	12	0,441
1E1	1D10	10,537	12	0,569
1B1	1D10	10,810	12	0,545
3G7	1D10	8,402	6	0,210
3G1	1D10	6,227	12	0,904
3H6	1D10	10,639	10	0,386
3A11	1D10	2,640	6	0,853

### III.2.3 VARIABILITE ALLELIQUE ET DIVERSITE GENETIQUE DES POPULATIONS

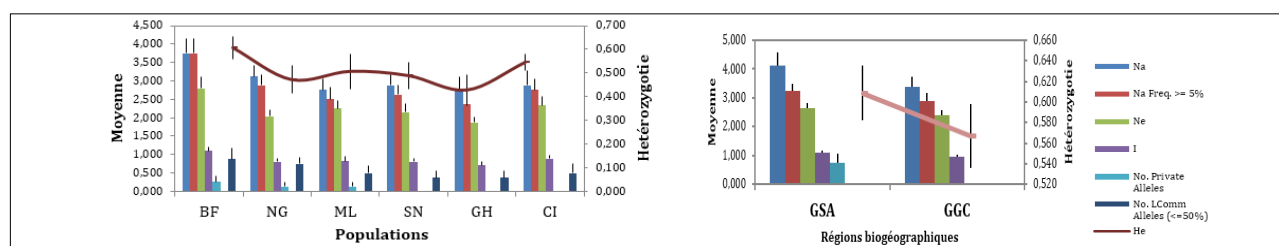
A l'échelle globale, il existe une grande variabilité allélique. En effet, le nombre moyen d'allèles par locus ( $N_a$ ) va de 3 (3H6 et 1D10) à 7 allèles (1A1) avec une valeur moyenne de 4,125 tandis que l'indice de la richesse allélique ( $R_A$ ) varie de 2,60 (3G1) à 4,61 (1A1) avec des valeurs de PIC variant entre 0,68 pour le 3G7 à 0,40 pour le 3G1 avec une valeur moyenne de 0,54. (Tableau XXI). L'analyse de la variabilité allélique a ainsi indiqué l'existence de trente trois (33) allèles sur l'ensemble des huit loci dans la population globale. Il y a eu 30 détectés au Burkina Faso, 25 au Niger, 23 au Sénégal et en Côte d'Ivoire, et 22 au Mali et au Ghana. Huit (8), 12,12%, de ces allèles ont été considérés comme rares (avec une fréquence d'allèle <1% selon la définition basée sur la faiblesse de leurs fréquences (Kimura, 1983b ; Joyce and Tavaré, 1993)) (Tableau XXIII). Trois (3) loci ont fourni 4 allèles privés répartis dans la population du Burkina Faso (2), du Mali (1) et du Niger (1) avec une fréquence de 12,08%. Le pourcentage de loci polymorphes est de 100% dans les populations du Burkina Faso, du Ghana, Niger et Sénégal contre 87,50% pour les populations du Mali et de la Côte d'Ivoire avec une moyenne de 95,83%.

Par ailleurs, le screening des populations indique que le nombre d'allèles par population va de 2,75 à 3,75, avec une moyenne de 3,75 (Burkina Faso), 2,75 (Côte d'Ivoire), 2,875 (Ghana), 2,75 (Mali), 3,125 (Niger), 2,785 (Sénégal) (Tableau XXI). Avec une moyenne de  $I=0,85$ , la plus forte valeur de l'indice de Shannon est observée au Burkina Faso ( $I=1,094$ ) et la plus faible au Ghana ( $I=0,700$ ). Les indices de diversité génétique sont plus importants dans la zone sahélo-soudanienne que dans la région guinéo-congolaise (Figure 28).

**Tableau XXIII : Liste des allèles avec leurs fréquences par population**

Locus	Allèles	Burkina Faso	Niger	Mali	Sénégal	Côte d'Ivoire	Ghana	Population globale
<b>1A1</b>	<u>147</u>	0,000	<u>0,050</u>	0,000	0,000	0,000	0,000	0,006*
	<u>247</u>	<u>0,100</u>	0,000	0,000	0,000	0,000	0,000	0,018 *
	<b>250</b>	0,233	0,000	0,300	0,233	0,154	0,333	0,223
	<b>255</b>	0,067	0,000	0,000	0,100	0,038	0,000	0,036
	<b>258</b>	0,067	0,100	0,000	0,000	0,154	0,000	0,048
	<b>262</b>	0,167	0,100	0,267	0,333	0,077	0,067	0,175
	<b>265</b>	0,367	0,750	0,433	0,333	0,577	0,600	0,494
<b>1 E1</b>	<b>292</b>	0,300	0,533	0,567	0,600	0,333	0,750	0,511
	<b>295</b>	0,133	0,300	0,400	0,267	0,667	0,071	0,309
	<b>297</b>	0,333	0,100	0,033	0,033	0,000	0,107	0,101
	<b>299</b>	0,233	0,067	0,000	0,100	0,000	0,071	0,079
<b>1B1</b>	<b>148</b>	0,433	0,467	0,464	0,800	0,267	0,633	0,511
	<b>152</b>	0,333	0,467	0,143	0,133	0,200	0,300	0,264
	<b>168</b>	0,133	0,033	0,321	0,067	0,267	0,033	0,140
	<b>172</b>	0,100	0,033	0,071	0,000	0,267	0,033	0,084
<b>3G7</b>	<b>185</b>	0,100	0,000	0,500	0,000	0,000	1,000	0,273
	<b>200</b>	0,100	0,000	0,500	0,500	0,067	0,000	0,199
	<b>208</b>	0,067	0,077	0,000	0,500	0,333	0,000	0,165
	<b>212</b>	0,733	0,923	0,000	0,000	0,600	0,000	0,364
<b>3G1</b>	<b>162</b>	0,067	0,000	0,000	0,000	0,000	0,033	0,017 *
	<b>166</b>	0,700	0,800	0,333	0,600	0,600	0,667	0,617
	<b>168</b>	0,233	0,200	0,567	0,400	0,400	0,300	0,350
	<u>171</u>	0,000	0,000	<u>0,100</u>	0,000	0,000	0,000	0,017*
<b>3H6</b>	<b>318</b>	0,000	0,100	0,000	0,214	0,000	0,000	0,056
	<b>321</b>	0,600	0,350	1,000	0,464	0,600	0,727	0,625
	<b>325</b>	0,400	0,550	0,000	0,321	0,400	0,273	0,319
<b>3A11</b>	<b>260</b>	0,067	0,067	0,033	0,000	0,000	0,000	0,028
	<b>261</b>	0,467	0,433	0,467	0,500	0,500	0,500	0,478
	<b>319</b>	0,233	0,500	0,500	0,500	0,500	0,500	0,456
	<u>321</u>	<u>0,233</u>	0,000	0,000	0,000	0,000	0,000	0,039
<b>1D10</b>	<b>265</b>	0,233	0,250	0,200	0,067	0,100	0,500	0,214
	<b>269</b>	0,433	0,542	0,433	0,900	0,367	0,167	0,482
	<b>274</b>	0,333	0,208	0,367	0,033	0,533	0,333	0,304

\* : Allèles rares avec une fréquence ≤ à 1% ; Allèles privés surlignés



**Figure 28 : Modèles alléliques dans les populations**

Moyenne (±SE) des paramètres génétiques estimés par population ; Na (nombre moyen d'allèles observés par population) ; Na Freq. >=5% (nombre moyen d'allèles observés avec une fréquence >=5%) ; Ne (nombre effectif moyen d'allèles par population) ; I (Index informatif de Shannon) ; No. Private Alleles (nombre moyen d'allèles privés par population) ; Nombre de loci commun trouvés dans moins 50% des populations ; He (hétérozygotie attendue) ; BF (Burkina Faso) ; NG (Niger) ; ML (Mali) ; SN (Sénégal) ; GH (Ghana) ; CI (Côte d'Ivoire).

### III.2.4 DIFFERENCIATION ET STRUCTURE GENETIQUES DES POPULATIONS

#### III.2.4.1 DIFFERENCIATION GENETIQUE ET FLUX GENIQUE

Globalement, l'analyse multilocus des valeurs de F-statistiques et des autres indices de différenciation a montré un niveau de significativité substantiel. Ce qui indique une forte structuration génétique. Toutefois, cette dernière est fortement impactée par le locus 3G7 qui est sous sélection

directionnelle. L'indice de consanguinité globale,  $F_{IT}$ , a indiqué un degré de consanguinité des individus au sein des populations de 10,7%. Cependant, il est devenu non significatif et a diminué jusqu'à 2,1% lorsque le locus 3G7 a été éliminé de l'analyse. Tandis que, l'analyse des valeurs du  $F_{IS}$  global sont significativement négatives montrant ainsi un excès hautement significative d'hétérozygotes. Concernant la différenciation génétique entre populations, la valeur du  $F_{ST}$  est hautement significative et importante ( $F_{ST}=16\%$ ) à modérée ( $F_{ST}=8,73\%$  lorsque le locus 3G7 est éliminé des analyses). La valeur du  $R_{ST}$  globale entre les populations étudiées corrobore ce résultat avec différenciation génétique hautement significative et importante  $R_{ST}=10,4\%$ . Cependant, elle diminue jusqu'à 1,1% sans influence du locus 3G7. Les analyses ont aussi indiqué qu'au locus 3G7, les valeurs du  $R_{ST}$  et du  $N_{ST}$  sont significativement supérieures à leurs valeurs permutées respectivement ( $pR_{ST}=0,044$  et  $pN_{ST}=0,035$ ). En outre, le flux génique moyen est de 2,547 et 2,878 migrants par génération respectivement avec les 8 et 7 loci selon la méthode traditionnelle (Tableau XXIV) alors qu'il est de 0,599 et 0,602 respectivement avec les 8 et 7 loci selon la méthode de l'allèle privé.

**Tableau XXIV** : F-statistiques et indices de différenciation par locus

Locus	$F_{IT}$	$F_{IS}$	$F_{ST}$	$R_{ST}$	$G_{ST}$	$N_{ST}$	$N_m$
<b>1A1</b>	0,133 <sup>NS</sup>	0,093 <sup>NS</sup>	0,044 <sup>NS</sup>	-0,032 <sup>NS</sup>	0,055 <sup>*</sup>	0,031 <sup>NS</sup>	3,067
<b>1E1</b>	0,143 <sup>NS</sup>	0,029 <sup>NS</sup>	0,118 <sup>***</sup>	0,116 <sup>NS</sup>	0,120 <sup>**</sup>	0,121 <sup>**</sup>	1,704
<b>1B1</b>	-0,201 <sup>NS</sup>	-0,326 <sup>**</sup>	0,094 <sup>**</sup>	0,180 <sup>NS</sup>	0,083 <sup>**</sup>	0,162 <sup>NS</sup>	2,325
<b>3G7</b>	0,560 <sup>**</sup>	0,044 <sup>NS</sup>	0,540 <sup>**</sup>	0,806 <sup>**</sup>	0,545 <sup>**</sup>	0,767 <sup>**</sup>	0,235
<b>3G1</b>	0,273 <sup>*</sup>	0,234 <sup>NS</sup>	0,051 <sup>NS</sup>	0,129 <sup>**</sup>	0,059 <sup>NS</sup>	0,080 <sup>NS</sup>	3,037
<b>3H6</b>	0,359 <sup>**</sup>	0,264 <sup>NS</sup>	0,128 <sup>NS</sup>	0,068 <sup>NS</sup>	0,142 <sup>NS</sup>	0,114 <sup>NS</sup>	1,379
<b>3A11</b>	-0,748 <sup>**</sup>	-0,810 <sup>**</sup>	0,034 <sup>**</sup>	-0,000 <sup>**</sup>	0,007 <sup>**</sup>	-0,032 <sup>**</sup>	7,163
<b>1D10</b>	0,210 <sup>NS</sup>	0,084 <sup>NS</sup>	0,137 <sup>***</sup>	0,061 <sup>NS</sup>	0,138 <sup>**</sup>	0,098 <sup>**</sup>	1,469
<b>Population globale</b>	0,107 <sup>***</sup>	-0,063 <sup>*</sup>	0,160 <sup>***</sup>	0,104 <sup>***</sup>	0,160 <sup>***</sup>	0,164 <sup>***</sup>	2,547
	<b>0,021<sup>NS</sup></b>	<b>-0,073<sup>*</sup></b>	<b>0,087<sup>***</sup></b>	<b>0,011<sup>*</sup></b>	<b>0,087<sup>***</sup></b>	<b>0,030<sup>***</sup></b>	<b>2,878</b>

La valeur moyenne des indices si le locus 3G7 est éliminé de l'analyse est en gras.

Dans l'ensemble, le coefficient de consanguinité individuelle au sein des populations est significatif et est de 8,4% sous l'analyse des 8 loci. Alors qu'elle est devenue non significative et qu'elle a diminuée jusqu'à 0,7%. Toutes les valeurs intra-individuelles du coefficient de parenté  $F_I$  sont négatives et non significatives avec l'analyse des 7 loci. Ceci a révélé qu'il n'y a d'écart significatif par rapport à l'équilibre de Hardy-Weinberg. Cette non-significativité a été aussi observée lorsque le locus 3G7 est inclus dans les analyses sauf pour les populations du Sénégal et du Mali où les valeurs du  $F_I$  sont significativement négatives. Il faut noter que dans la population du Ghana, la valeur du  $F_I=0,041$  y est restée la même quelque soit le nombre de loci considéré (8 et 7 loci) (locus 3G7 y est monomorphe). De même, dans la population de la Côte d'Ivoire, même si elle n'est significative, la valeur du  $F_I$  est de 10,2% sous l'impact du locus 3G7 qui est potentiellement sous sélection directionnelle.

**Tableau XXV** : Coefficient de consanguinité individuelle au sein des populations

Populations	Coefficient de consanguinité, <i>F<sub>i</sub></i>	
	Résultats avec les 8 loci	Résultats avec les 7 loci
Burkina Faso	0,003 NS	-0,084 NS
Niger	-0,124 NS	-0,170 NS
Mali	-0,197 *	-0,088 NS
Sénégal	-0,185 *	-0,069 NS
Ghana	0,041 NS	0,041 NS
Côte d'Ivoire	0,102 NS	-0,025 NS
Population globale	<b>0,084 **</b>	<b>0,007 NS</b>

Niveau de significativité : NS:  $p > 0,05$ = non significatif ; \* :  $P < 0,05$ = significatif ; \*\* :  $P < 0,01$ = très significatif.

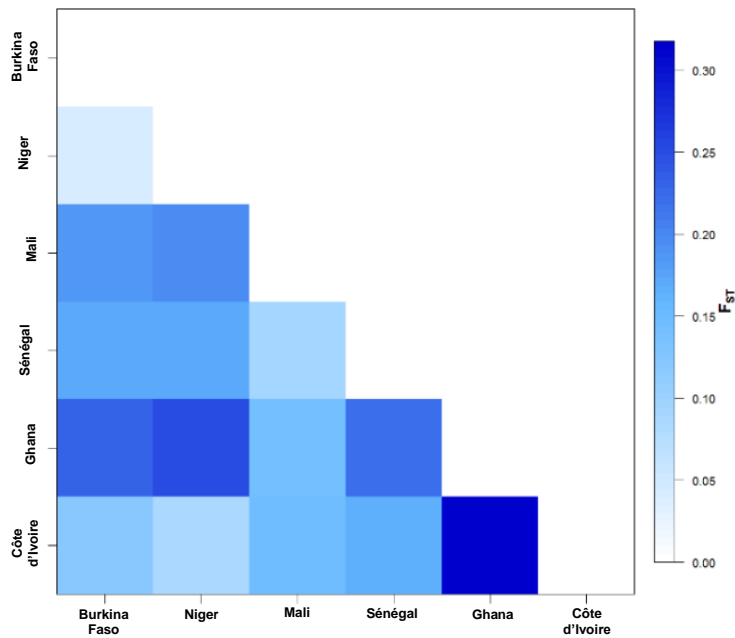
**Tableau XXVI** : Indices de différenciation génétique entre paires de populations

Population 1	Population 2	<i>G<sub>ST</sub></i>	<i>N<sub>ST</sub></i>	<i>R<sub>ST</sub></i>	<i>F<sub>ST</sub></i>	Nm	<i>D<sub>s</sub></i>
Burkina Faso	Niger	0,045 NS	0,029 NS	-0,005 NS	0,046 NS	6,05	0,050 NS
		<b>0,044*</b>	<b>0,021 NS</b>	<b>-0,009 NS</b>	<b>0,048*</b>	<b>4,99</b>	<b>0,061 NS</b>
Burkina Faso	Mali	0,153**	0,165**	0,010**	0,154**	2,04	0,273**
		<b>0,092**</b>	<b>0,015 NS</b>	<b>0,001 NS</b>	<b>0,094**</b>	<b>2,42</b>	<b>0,144 **</b>
Burkina Faso	Sénégal	0,152**	0,058**	0,013 NS	0,154**	1,32	0,256**
		<b>0,086**</b>	<b>0,011 NS</b>	<b>0,005 NS</b>	<b>0,055**</b>	<b>2,58</b>	<b>0,121 **</b>
Burkina Faso	Ghana	0,174**	0,249**	0,204**	0,174**	2,33	0,259**
		<b>0,054 NS</b>	<b>0,021 NS</b>	<b>0,008 NS</b>	<b>0,088 NS</b>	<b>4,30</b>	<b>0,067 NS</b>
Burkina Faso	Côte d'Ivoire	0,062*	0,028 NS	0,015 NS	0,060*	1,91	0,099**
		<b>0,061*</b>	<b>0,027 NS</b>	<b>0,016 NS</b>	<b>0,064*</b>	<b>3,68</b>	<b>0,102 *</b>
Niger	Mali	0,245**	0,263**	0,129**	0,249**	2,29	0,408**
		<b>0,149**</b>	<b>0,047**</b>	<b>0,002 NS</b>	<b>0,153**</b>	<b>1,38</b>	<b>0,222 **</b>
Niger	Sénégal	0,192**	0,097**	0,015 NS	0,194**	5,17	0,269**
		<b>0,080**</b>	<b>0,010 NS</b>	<b>-0,010 NS</b>	<b>0,082**</b>	<b>2,80</b>	<b>0,099 *</b>
Niger	Ghana	0,240**	0,329**	0,222**	0,240**	4,07	0,324**
		<b>0,065 NS</b>	<b>0,002 NS</b>	<b>-0,011 NS</b>	<b>0,065 NS</b>	<b>3,59</b>	<b>0,079 NS</b>
Niger	Côte d'Ivoire	0,074 NS	0,062**	0,024 NS	0,071 NS	2,40	0,090 NS
		<b>0,063 NS</b>	<b>0,058**</b>	<b>0,024 NS</b>	<b>0,064 NS</b>	<b>3,63</b>	<b>0,086 NS</b>
Mali	Sénégal	0,141**	0,123**	0,079**	0,146**	2,16	0,192**
		<b>0,110*</b>	<b>0,037*</b>	<b>0,020 NS</b>	<b>0,111*</b>	<b>2,00</b>	<b>0,140 **</b>
Mali	Ghana	0,124**	0,044 NS	0,041**	0,126**	1,96	0,139 **
		<b>0,078**</b>	<b>0,023 NS</b>	<b>0,014 NS</b>	<b>0,079**</b>	<b>2,93</b>	<b>0,094 **</b>
Mali	Côte d'Ivoire	0,138**	0,200**	0,128**	0,140**	1,385	0,211**
		<b>0,063*</b>	<b>0,002 NS</b>	<b>0,002 NS</b>	<b>0,066*</b>	<b>3,55</b>	<b>0,084 *</b>
Sénégal	Ghana	0,245**	0,192**	0,158**	0,247**	2,25	0,346**
		<b>0,122**</b>	<b>0,003 NS</b>	<b>-0,004 NS</b>	<b>0,123**</b>	<b>1,79</b>	<b>0,154 **</b>
Sénégal	Côte d'Ivoire	0,164**	0,127**	0,069*	0,164**	1,52	0,255**
		<b>0,132**</b>	<b>0,092**</b>	<b>0,054*</b>	<b>0,132**</b>	<b>1,64</b>	<b>0,191 **</b>
Ghana	Côte d'Ivoire	0,224**	0,331**	0,265**	0,221**	1,64	0,340**
		<b>0,104**</b>	<b>0,073*</b>	<b>0,045*</b>	<b>0,102**</b>	<b>2,20</b>	<b>0,142 *</b>

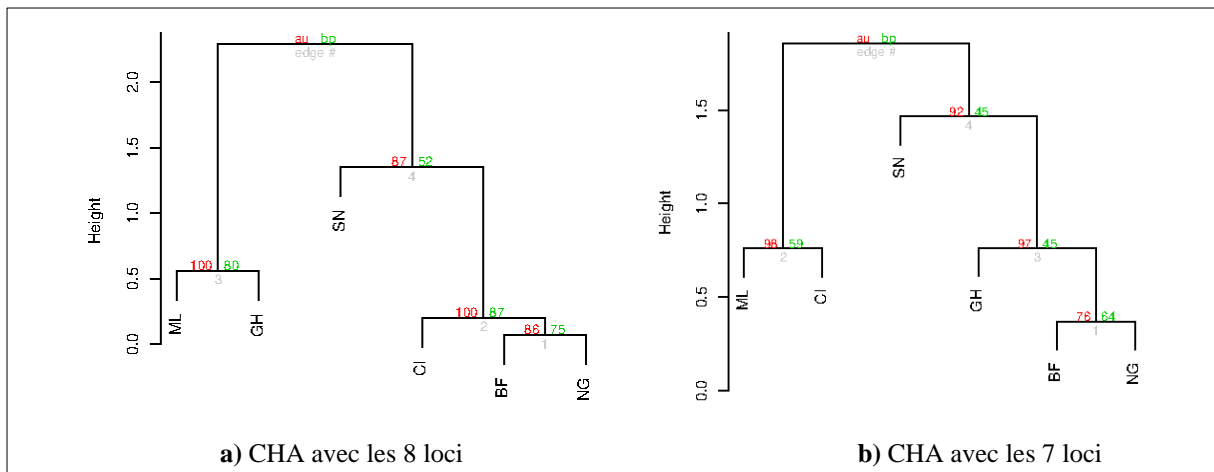
Valeurs moyennes des indices si le locus 3G7 est éliminé de l'analyse (en gras) ; *D<sub>s</sub>* : Distance génétique de Nei (1978).

L'analyse comparative du niveau de différenciation entre paire de populations permet de regrouper les populations en 3 ensembles globaux : le groupe à différenciation faible voire assez modérée (Burkina Faso, Côte d'Ivoire et Niger), celui à différenciation modérée (le Burkina Faso vis-à-vis du Sénégal, du Ghana et du Mali ainsi que le Mali envers le Côte d'Ivoire, le Ghana et le Sénégal) et enfin le groupe à différenciation importante à très importante (le Ghana en comparaison avec le Sénégal, la Côte d'Ivoire, le Niger ; le Sénégal par rapport au Niger et à la Côte d'Ivoire ; et finalement entre le Niger et le Mali).

L'analyse des matrices de la divergence et de la différenciation génétique (Figures 29 et 30 et Tableau XXVI) par une classification hiérarchique ascendante confirme la grande affinité entre le Burkina Faso, le Niger et la Côte d'Ivoire d'une part ; et d'autre part, elle a indiqué que le locus 3G7 a entraîné un rapprochement du Ghana et un isolement de la Côte d'Ivoire du Burkina Faso et du Niger.



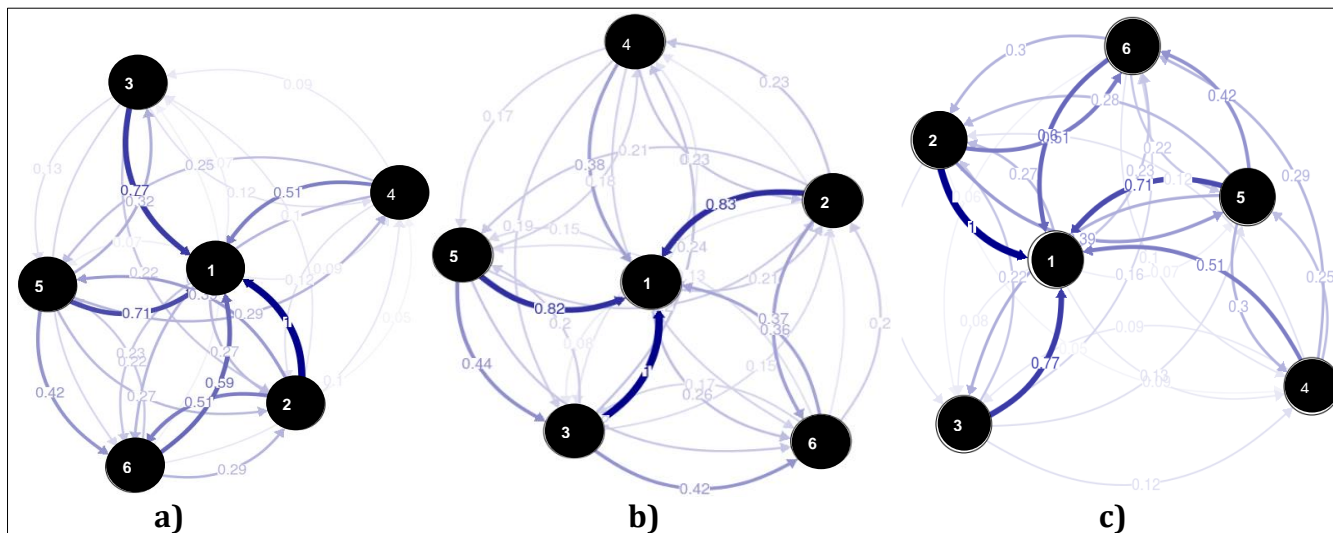
**Figure 29 :** Différenciation génétique entre populations



**Figure 30 :** Classification hiérarchique ascendante des populations

Les réseaux de migration relative directionnelle représentent les taux de migration relative entre les 6 populations (Figure 31). Suivant les trois estimateurs ( $D$  de Jost,  $G_{ST}$  et  $N_m$ ), les populations sont regroupées dans un modèle similaire aux analyses des  $F_{ST}$  par paire de populations. Le flux génétique bidirectionnel entre les populations est asymétrique particulièrement celui entre le Burkina Faso et les autres populations. En effet, le flux des autres populations vers la population du Burkina Faso est plus élevé que l'inverse selon les trois estimateurs (Jost's  $D$ ,  $G_{ST}$ ,  $N_m$ ) bien qu'il ait une

légère variation dans l'analyse du D de Jost. Cependant, cette asymétrie n'est significative qu'avec la population du Mali (analyse avec  $N_m$  et  $G_{ST}$ ) et celle du Ghana (analyse du D de Jost). En outre, il existe une asymétrie significative du Mali vers la Côte d'Ivoire et du Sénégal vers le Niger (analyses du  $N_m$  et du  $G_{ST}$  uniquement).



**Figure 31 :** Les réseaux standards de migration relative directionnelle des populations  
 Les valeurs de migrations relatives du  $N_m$  (a); du D de Jost (b) et du  $G_{ST}$  (c). Les populations sont indiquées par des numéros : 1 : Burkina Faso ; 2 : Niger ; 3 : Mali ; 4 : Sénégal ; 5 : Ghana ; 6 : Côte d'Ivoire.

### III.2.4.2 DIFFERENCIATION GENETIQUE ENTRE POPULATIONS DU NORD ET CELLE DU SUD

L'analyse par comparaison des moyennes entre la zone sahélo-soudanienne et la région guinéo-congolaise (Olson *et al.*, 2001) met en évidence des différences significatives de la diversité génétique en ce qui concerne l'hétérozygotie observée et la richesse allélique en faveur du Nord (Tableau XXVII). La région guinéo-congolaise montre un taux de consanguinité de  $F_{IS}=7,5\%$  et une valeur du  $F_{ST}$  significative et très importante (22,1%) contrairement à la zone Sahélo-soudanienne où la valeur du  $F_{IS}$  est significativement négative et celle (-0,12%) du  $F_{ST}$  significativement importante (15,6%). En outre, le taux de consanguinité globale est de 5% pour la zone sahélo-soudanienne contre 28% pour la région guinéo-congolaise. Cependant, la différenciation génétique entre les deux régions biogéographiques est faible mais significative (3,1%). Les indices de différenciation sont approximativement multipliés par 2 sous l'influence du locus 3G7, potentiellement sous sélection directionnelle (Tableau XXVII).

**Tableau XXVII :** Indices de variabilité et de différenciation génétiques

Populations	N	$R_A$	$N_A$	$H_E$	$H_O$	$F_I$	$F_{IT}$	$F_{ST-intra}$	$F_{ST-inter}$
Groupe des Pays de la zone Sahélo-soudanienne	60	3,519	4,125	0,614	0,605	0,015 <sup>NS</sup>	0,052 *	0,156 **	0,031 **
Groupe des Pays de la région guinéo-congolaise	30	3,023	3,375	0,577	0,467	0,194**	0,280 *	0,221 **	

N : Nombre de populations utilisées dans l'analyse ;  $R_A$  : Richesse allélique ;  $N_A$  : Nombre moyen d'allèles par locus ;  $H_O$  : Hétérozygotie attendue non biaisé,  $H_E$  : Hétérozygotie observée ;  $F_I$  : Coefficient de consanguinité.

### III.2.4.3 FACTEURS DE DIFFERENCIATION ET DE STRUCTURATION GENETIQUES

#### III.2.4.3.1 SOURCES DE LA VARIANCE MOLECULAIRE

La répartition hiérarchique de la variance moléculaire par AMOVA a révélé qu'il n'y a aucune différenciation génétique due entre populations des régions biogéographiques (zone sahélo-soudanienne et région guinéo-congolaise). Elle a aussi mis en évidence que 89 à 90% de la variation totale sont dues aux caractéristiques individuelles, alors que 15 à 17% de la variabilité génétique est expliquée par une différenciation entre populations avec l'analyse des 8 loci (Tableau XXVIII). Cependant, lorsqu'on élimine le locus potentiellement sous sélection directionnelle, les caractéristiques individuelles causent 97 à 98% de la variation génétique contre 8 à 9% qui sont liées aux divergences entre populations.

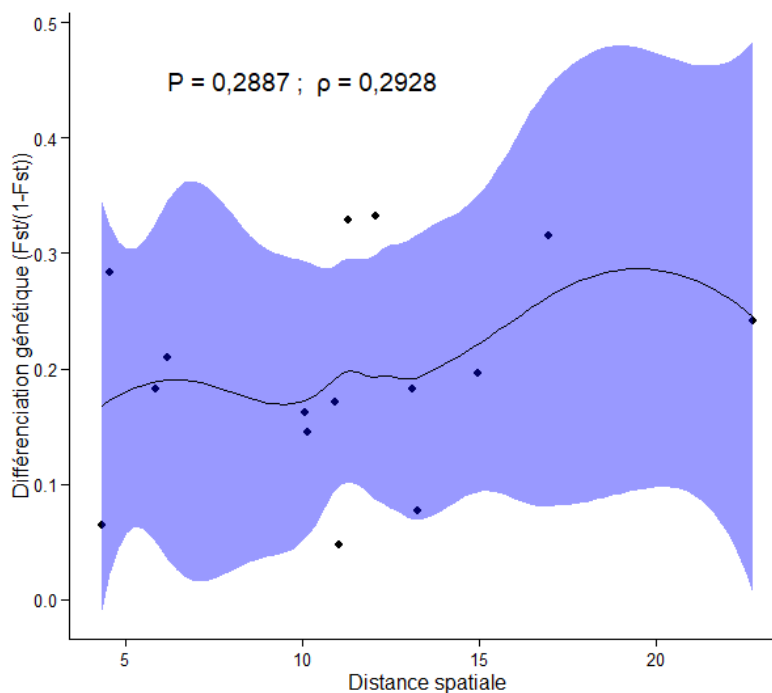
**Tableau XXVIII** : Source et répartition de la variance moléculaire

Source de la variance	Somme des carrés de déviation	Composantes de la variance	Pourcentage de la variance	Indices de fixation
<b>Répartition de la variance moléculaire entre zones biogéographiques</b>				
Entre Zone	8,503 / <b>4,720</b>	-0,08216 / <b>-0,03725</b>	-3,32649 / <b>-1,78241</b>	$F_{CT} = -0,03326 / -0,01782$
Entre populations au sein des zones	58,912 / <b>30,025</b>	0,44384 / <b>0,20331</b>	17,96930 / <b>9,72744</b>	$F_{SC} = 0,17391^{***} / 0,09557^{***}$
Entre individus au sein des populations	155,673 / <b>139,893</b>	-0,13194 / <b>-0,14009</b>	-5,34171 / <b>-6,70249</b>	$F_{IS} = -0,06258 / -0,07281$
Au sein des individus dans la population globale	194,000 / <b>178,500</b>	2,24025 / <b>2,06412</b>	90,69890 / <b>98,75746</b>	$F_{IT} = 0,09301^{**} / 0,01243$
Total	353,139 / <b>353,139</b>	2,09009 / <b>2,09009</b>	100	
<b>Répartition de la variance moléculaire entre populations</b>				
Entre populations	67,415 / 34,746	0,40029 / <b>0,18400</b>	15,95679 / <b>8,72835</b>	$F_{ST} = 0,15957^{***} / 0,08728^{***}$
Entre individus au sein des populations	155,673 / 139,893	-0,13194 / <b>-0,14009</b>	-5,25948 / <b>-6,64545</b>	$F_{IS} = -0,06258 / -0,07281$
Au sein des individus	194,000 / 178,500	2,24025 / <b>2,06412</b>	89,30269 / <b>97,91711</b>	$F_{IT} = 0,10697^{***} / 0,02083$
Total	417,088 / <b>353,139</b>	2,0861 / <b>2,10802</b>	100	

La valeur moyenne des indices si le locus 3G7 est éliminé de l'analyse est en gras.

#### III.2.4.3.2 ISOLEMENT PAR LA DISTANCE GENETIQUE

Les hypothèses de linéarité, de normalité et d'homogénéité et d'indépendance des résidus est rejetée (Annexes 7, 8, 10 et 11) et seule l'hypothèse d'indépendance des résidus a été acceptée (Annexes 9). Alors, les résultats de la régression linéaire ne sont pas valides. Par conséquent, une corrélation de Spearman a été effectuée, vu la validité de l'indépendance des résidus. Dès lors, l'analyse de la relation de  $F_{ST} / (1 - F_{ST}) = f(\text{distance spatiale})$  montre une faible association positive ( $\rho \leq 0,5$ ) entre la différenciation génétique des populations de *Sitophilus zeamais* en Afrique de l'Ouest. Cependant, cette liaison n'est pas statistiquement significative selon la corrélation de Spearman ( $P = 0,2887$  ;  $\rho = 0,2928$ ) (Figure 32).



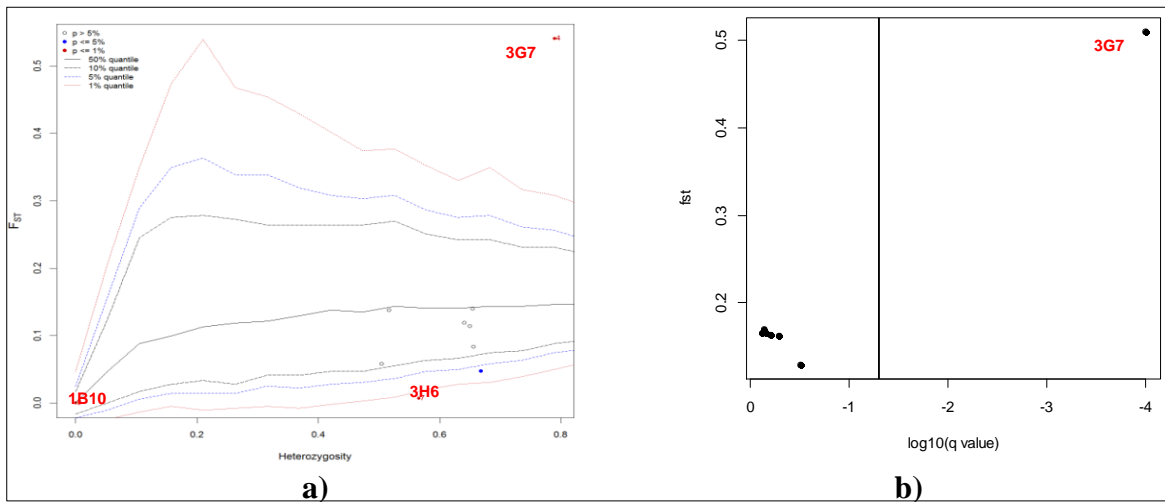
**Figure 32** : Relation de regression avec corrélation de Spearman

### III.2.4.3.3 IMPACT DES LOCI DANS LA DIFFERENCIATION GENETIQUE DES POPULATIONS

Après les analyses des tests de sélection, 3 loci sont identifiés comme étant sous sélection. Sous les modèles IAM (Infinite Alleles Model) et SMM (Stepwise Mutation Model), le logiciel LOSITAN a permis de détecter des signatures de sélection positive à un locus (3G7) et de sélection balancée à deux loci (3H6, 3A11) dans les six populations. Le locus (3G7) a été détecté comme étant sous sélection positive basé sur les analyses de Bayescan alors que les trois ont été détectés par ARLEQUIN (Figure 33). Cependant, il y'a plusieurs paramètres à prendre en compte avec les approches *outliers* (valeurs de  $F_{ST}$  atypiques, présentant des valeurs extrêmes), à savoir les erreurs de génotypage, la structure de la population et les faux positifs, la variation du taux de mutation et la limite de la sensibilité (faux négatifs) (Narun et Hess, 2011).

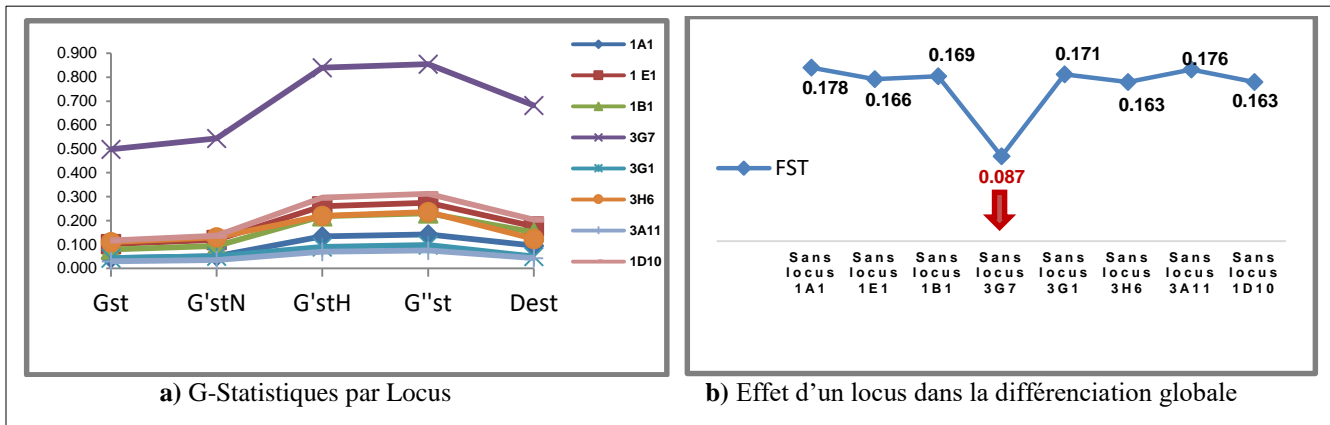
D'après Pérez-Figueroa *et al.* (2010), les faux positifs sont communs, même avec une combinaison de méthodes et de correction multi-tests, suggérant que les conclusions obtenues à partir de cette approche doivent être prises avec précaution. Le programme Bayescan est moins sensible aux faux positifs et négatifs (Pérez-Figueroa *et al.*, 2010 ; Narun et Hess, 2011 ; Feng *et al.*, 2015) contrairement à ARLEQUIN (Narun et Hess, 2011). Dès lors, le locus (3G7) a été systématiquement identifié comme étant sous sélection avec les trois méthodes. Ce locus est à l'origine d'une sélection positive directionnelle entre les populations ; ce qui est confirmé par les valeurs des  $F_{ST}$  et celles des  $G$ -statistiques (Figure 34).





**Figure 33 :** Analyses des loci sous sélection

Distribution des valeurs de  $F_{ST}$  en fonction de l'hétérozygotie par comparaisons inter-populationnelles sous ARLEQUIN (a) ; tracé des valeurs du  $F_{ST}$  en fonction au  $\log_{10}$  des chances postérieures (PO, Posterior Odds) sous BAYESCAN (b)



**Figure 34 :** Modèles alléliques dans la population globale

**III.2.4.3.4 DETECTION DE SIGNAL PHYLOGEOGRAPHIQUE**

Les tests de permutation des indices de différenciation basés sur la taille des allèles des microsatellites ( $R_{ST}$ , Hardy *et al.*, 2003) ou sur les distances phylogénétiques ( $N_{ST}$ , Burban *et al.*, 1999) entre allèles n'ont pas été significative aussi bien pour les 8 loci et pour les 7 loci (Tableau XXIX). Par ailleurs, à l'échelle globale les indices de Nei (1987) montrent que la contribution de la diversité génétique moyenne intra-population ( $H_S$ ) par rapport à la diversité génétique totale ( $H_T$ ) est supérieure à celle de la diversité génétique inter-population ( $D_{ST}$ ) (Tableau XXIX).

**Tableau XXIX :** Indices de signal phylogéographique et de Nei (1973 ; 1987)

Zones	Marqueur	Test de permutations		$H_T$	$H_S$	$D_{ST}$
		$N_{ST}$	$R_{ST}$			
Population globale	Multiloci	P = 0,371	P = 0,395	0,609	0,525	0,084

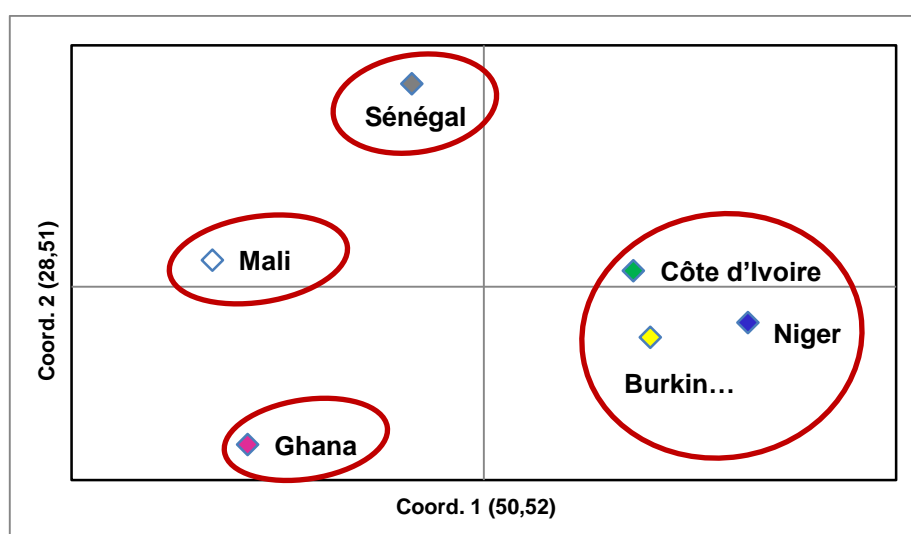
Indices de différenciation génétique ( $N_{ST}$ ,  $R_{ST}$ ), Indices de Nei ( $H_T$ ,  $H_S$  et  $D_{ST}$ , 1973 ; 1987).

### III.2.3 CLUSTERING, ORGANISATION SPATIALE ET INTROGRESSION DES POPULATIONS

#### III.2.3.1 ANALYSE MULTIVARIEE

##### III.2.3.1.1 ANALYSE PRINCIPALE DES CORRESPONDANCES (ACP)

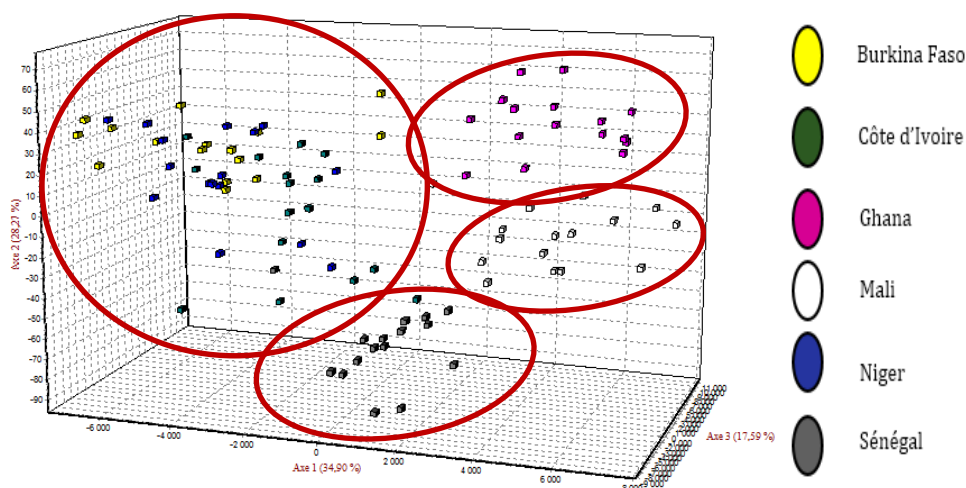
L'analyse en composantes principales (ACP) explique 79,03% de la variabilité génétique totale à travers les deux premiers axes (Figure 35). La première dimension qui explique 50,52% de la variabilité génétique révèle globalement deux groupes génétiques : le groupe des populations du Burkina Faso, du Niger et de la Côte d'Ivoire qui se distinguent du Ghana, du Mali et du Sénégal. La deuxième dimension, avec 28,51% de la variabilité génétique, a montré un isolement modérée à très importante de la population du Ghana. Ainsi, les résultats de l'ACP, les populations du Burkina Faso, du Niger et de la Côte d'Ivoire forment un groupe génétique homogène.



**Figure 35 :** Diagramme de l'Analyse en Composantes Principales (ACP) inférée de la matrice de covariance de la distance génétique non biaisée (DS) de Nei (1978)

##### III.2.3.1.2 ANALYSE FACTORIELLE DE CORRESPONDANCE (AFC)

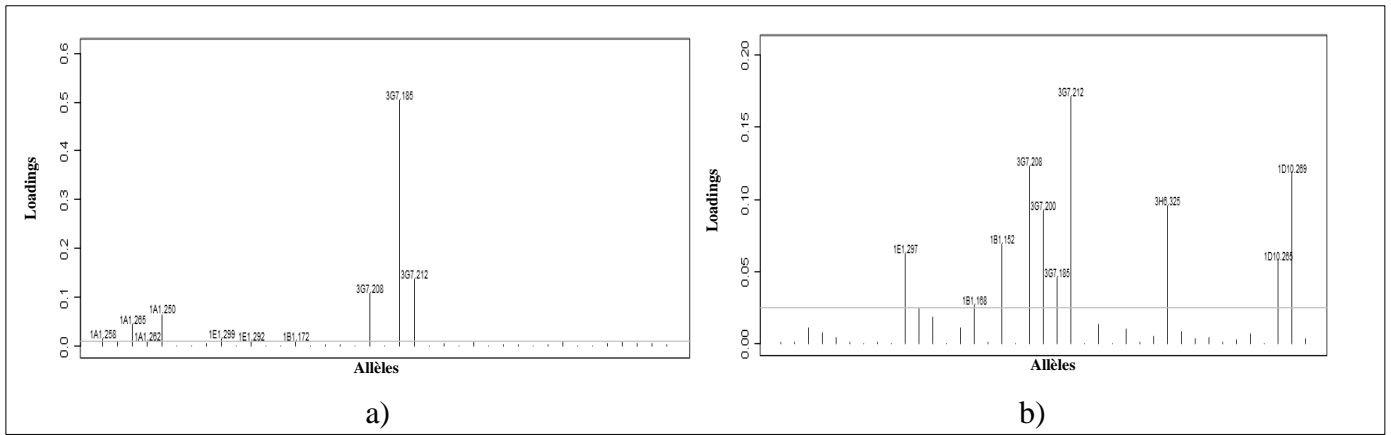
L'AFC des populations, réalisée à partir des matrices de distances génétiques, suggère une forte introgression entre les populations (Figure 36). En effet, les 3 premiers axes expliquent l'essentiel de la variance (80,76%). Le premier axe de l'AFC révèle 34,90% de la variabilité génétique totale et regroupe les populations du Mali et du Ghana d'une part et d'autre part les populations du Burkina Faso, de la Côte d'Ivoire et du Niger ; il existe un groupe intermédiaire formé par les individus du Sénégal qui introgressent avec le second groupe génétique. Alors que le second axe, qui assigne 28,27% de la variabilité totale, isole la population du Sénégal du reste des autres populations. Suivant le troisième axe factoriel qui explique 17,59% de la variabilité génétique totale, 2 regroupements se sont révélés : le Burkina Faso, le Niger et la Côte d'Ivoire d'une part et d'autre part le Mali et le Ghana ; alors que le Sénégal constitue un groupe intermédiaire.



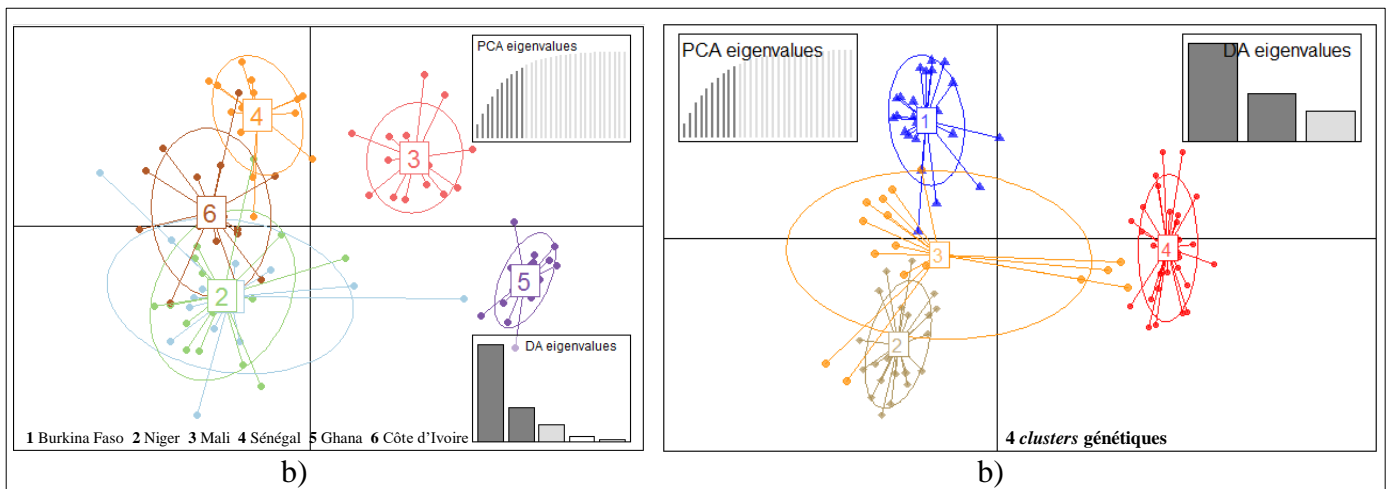
**Figure 36 :** Mise en évidence des relations génétiques entre populations par AFC

### III.2.3.1.3 ANALYSE DISCRIMINANTE DES COMPOSANTES PRINCIPALES (DAPC)

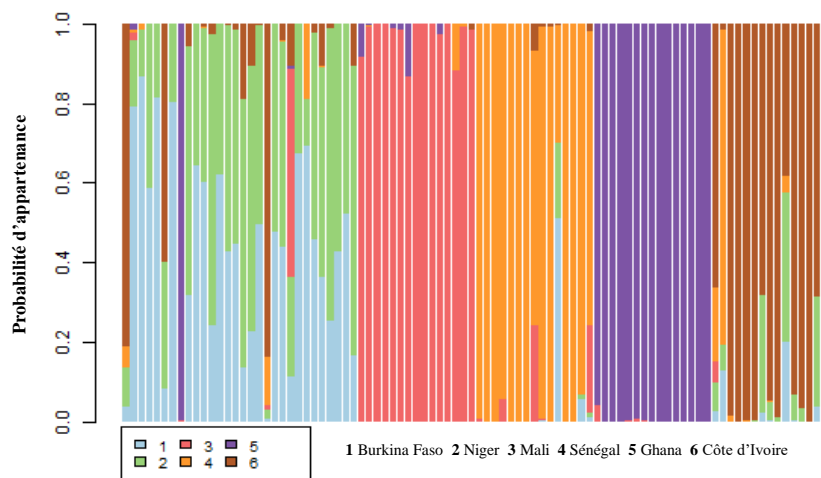
L'assignation correcte globale est de 87,78%. Les proportions de réassignations réussies (basées sur les fonctions discriminantes, DA) des individus à leurs populations d'origine sont de 53, 67, 100, 93, 100 et 87% pour respectivement le Burkina Faso, le Niger, le Mali, le Sénégal, le Ghana et la Côte d'Ivoire. Les grandes valeurs indiquent des populations avec des caractéristiques propres, tandis que les faibles valeurs suggèrent des populations issues d'une admixture. Les analyses ont aussi montré que parmi les 10 allèles sélectionnés pour contribuer au DAPC dont principalement les allèles 185, 212 et 208 du locus 3G7, qui est identifié comme étant potentiellement sous sélection directionnelle, ont eu les plus fortes contributions à la construction du premier axe discriminant avec 0,50 ; 0,14 et 0,11 respectivement. Alors que la construction du deuxième axe discriminant a nécessité, en plus du locus 3G7, les loci 1B1, 1E1, 3H6 et 1D10 (Figure 37). De plus, la première fonction discriminante a montré l'évolution spatiale de *S. zeamais* indiquant une forte introgression entre les populations du Burkina Faso, de la Côte d'Ivoire, du Niger et du Sénégal, ces dernières étant isolées des populations du Ghana et du Mali dont l'introgression est moins importante. Tandis que la seconde a indiqué que la population de la Côte d'Ivoire est intermédiaire entre deux blocs formés d'une part, par les populations du Burkina Faso, du Niger et du Ghana ; et d'autre part, par celles du Mali et du Sénégal. En outre, quatre *clusters* génétiques ont été identifiés. Le cluster C1 « bleu » est composé des individus originaires de la Côte d'Ivoire (3), du Niger (3), du Mali (3), du Ghana (1) et du Sénégal (14). Puis, le C2 « marron » est formé d'individus échantillonnés au Sénégal (1), au Burkina Faso (3), au Ghana (13) et au Mali (5). Ensuite, le C3 « orange » rassemble des individus du Burkina Faso (2) du Mali (7), du Sénégal (1) et de la Côte d'Ivoire (6). Enfin, le C4 « rouge » regroupe des individus issus des populations du Burkina Faso (10) du Niger (12) et de la Côte d'Ivoire (6) (Figure 38). Il faut noter qu'il existe une congruence entre les analyses multivariées des populations.



**Figure 37 :** Contributions à la fonction discriminante 1 (a) et 2 (b) des allèles



**Figure 38 :** Diagramme de dispersion de l'analyse du DAPC  
Structure des populations (a) et des *clusters* génétiques (b)

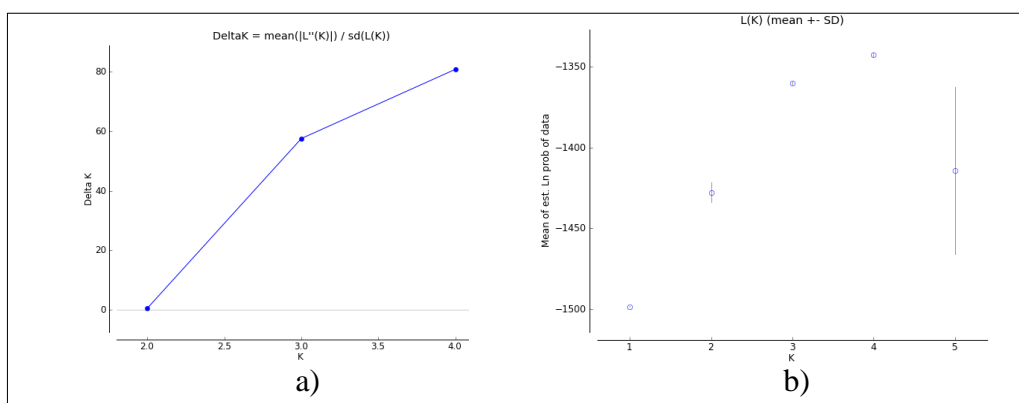


**Figure 39 :** Assignment des individus selon le DAPC

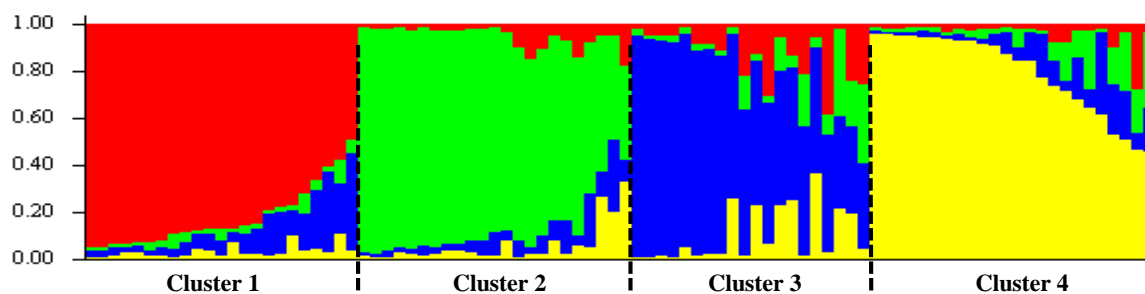
### III.2.3.2 ANALYSE BAYESIENNE

#### III.2.3.2.1 STRUCTURE BASEE SUR LE MODELE DE HARDY-WEINBERG

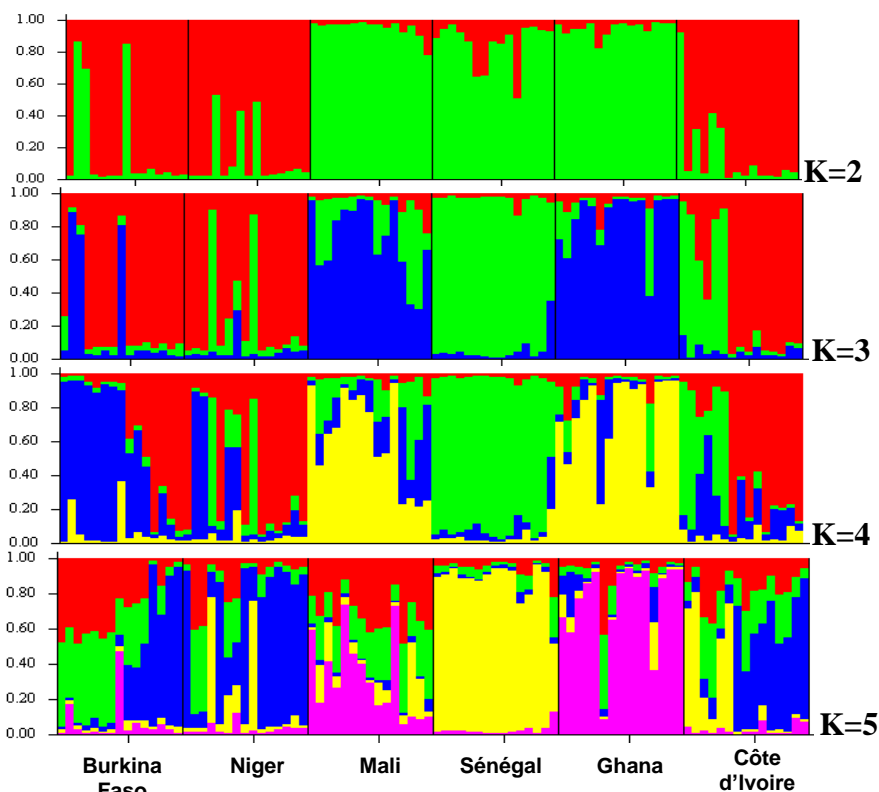
Le *clustering* par approche bayésienne, à l'aide de la méthode du  $\Delta K$  et celle de la vraisemblance en fonction du nombre de *clusters*  $K$  pour déterminer le nombre optimal de cluster  $K$  (Figures 40). L'analyse donne la vraisemblance la plus importante avec  $K = 4$ . Ce qui indique que la population entière peut être regroupée en 4 *clusters* (Figures 41). Pour chacune des valeurs de  $K=4$ , nous avons sélectionné parmi les 10 chaînes lancées la chaîne pour laquelle la vraisemblance a été la plus élevée, c'est-à-dire la valeur avec laquelle la vraisemblance est maximale, à savoir la probabilité d'observer les données comme réalisation d'un échantillon de la loi. L'évolution de la structure génétique en fonction du nombre de cluster est représentée dans la figure 41. Dans cette figure, chaque individu est représenté par une barre verticale, chaque couleur représentant la proportion du génotype multilocus de cet individu venant d'un groupe donné. L'analyse de la proportion du génotype multilocus confirme le groupe déjà identifié par l'approche multivariée mais aussi l'isolement du Sénégal du reste des autres populations et enfin, l'affinité entre le Mali et le Ghana (Figures 42).



**Figure 40 :** Détermination de  $K$  le plus probable avec (a) : la méthode du  $\Delta K$  d'Evanno *et al.* (2005) et (b) : celle de la vraisemblance  $L(K)$

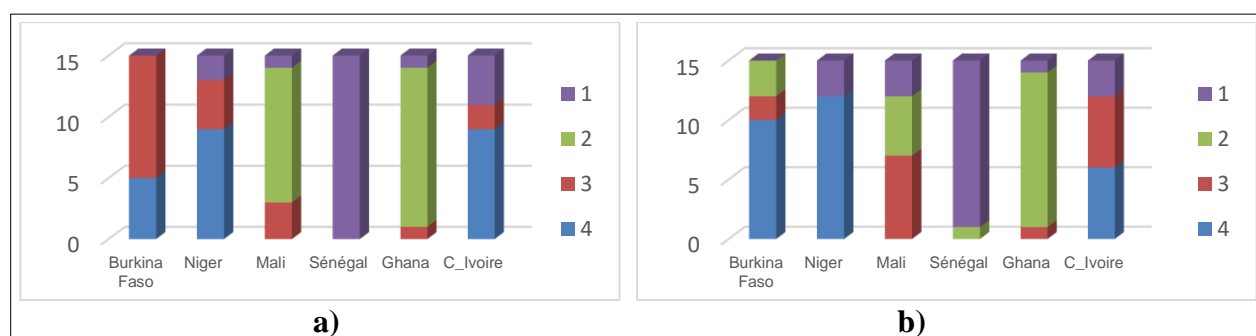


**Figure 41 :** Assignation des individus aux différents *clusters* génétiques inférée à l'aide de l'indice d'assignation ( $Q$ )



**Figure 42 :** Evolution de la structuration génétique des populations pour un nombre de *clusters* génétiques K assignés par une couleur et variant de 2 à 5

Les 4 *clusters* génétiques ont été aussi bien identifiés par l’approche bayésienne et multivariée avec des contenus plus ou moins différents (Figure 43). En effet, l’analyse des résultats des deux approches a montré que le groupe G1 identifié par la méthode bayésienne, majoritairement composé du cluster rouge, correspond au groupe 4 « rouge » de l’analyse DAPC et regroupant des individus échantillonnés au Burkina Faso, au Niger et en Côte d’Ivoire (Tableau XXX). Le groupe G2 formé principalement du cluster vert (algorithme structure) équivaut au groupe 1 « bleu » (analyse de DAPC) qui rassemble majoritairement des individus venus du Sénégal. Le G3 composé généralement du cluster jaune concorde au résultat du groupe 2 « marron ». Cependant, le G4 majoritairement composé du cluster bleu ne présente pas une nette concordance avec le groupe 3 « orange » de l’analyse du DAPC.

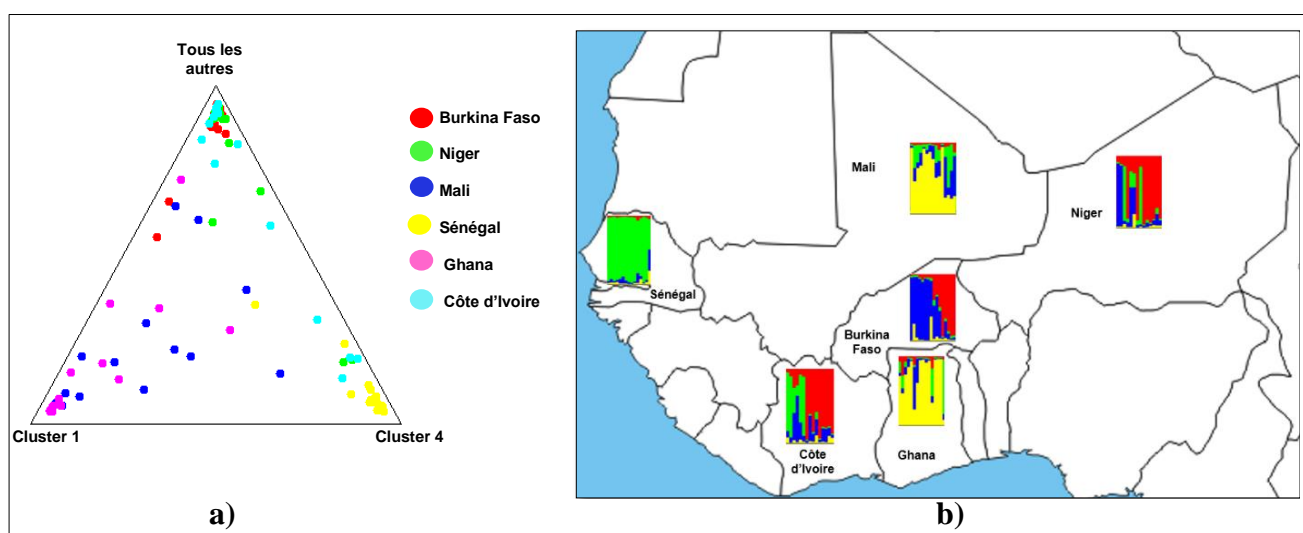


**Figure 43 :** Contenu des populations en fonction des 4 *clusters* (à) par inférence bayésienne et (b) par inférence du DAPC.

**Tableau XXX :** Proportion d'appartenance de chaque population pré-définie dans les 4 *clusters*

Populations de référence	Cluster 1 (rouge)	Cluster 2 (vert)	Cluster 3 (bleu)	Cluster 4 (jaune)
<b>Burkina Faso</b>	0,325	0,033	0,574	0,067
<b>Niger</b>	0,588	0,157	0,216	0,039
<b>Mali</b>	0,037	0,162	0,194	0,607
<b>Sénégal</b>	0,024	0,874	0,062	0,040
<b>Ghana</b>	0,061	0,067	0,104	0,767
<b>Côte d'Ivoire</b>	0,526	0,251	0,179	0,043

L'analyse de la variance moléculaire indique une différenciation très importante et très significative entre les *clusters* ( $F_{ST}=0,6075$  et  $F_{ST}=0,4555$  résultant de l'analyse avec et sans le locus 3G7 respectivement) (Tableau XXXI). Dès lors, le locus 3G7 qui est potentiellement sous sélection directionnelle influe fortement l'hétérogénéité du pool génétique des populations mais aussi de la structure génétique.



**Figure 44 :** Inférence de la structure génétique des *clusters*

**Tableau XXXI :** Source et répartition de la variance moléculaire des *clusters* avec 8 et 7 loci.

Source de la variance	Somme des carrés de déviation	Composantes de la variance	Pourcentage de la variance	Indices de fixation
Entre <i>clusters</i>	156,456 / 99,278	1,16065 / 0,73648	60,75 / 49,545	$F_{ST} = 0,60745^{***} / 0,49545^{***}$
Au sein des <i>clusters</i>	0,000 / 0,000	-0,75000 / -0,75000	-39,25 / -50,4547	$F_{IS} = -1,000 / -1,000$
Au sein des individus	135,000 / 135,000	1,50000 / 1,50000	78,51 / 100,909	$F_{IT} = 0,21493 / -0,00909$
Total	291,456 /	1,91065 /		

La valeur moyenne des indices si le locus 3G7 est éliminé de l'analyse est en gras

### III.2.3.2.2 DETECTION DE MIGRANTS DE PREMIERE GENERATION

Nous avons détecté un nombre de quatre migrants de première génération sur l'ensemble du jeu de données selon la méthode des fréquences alléliques de Paetkau *et al.* (1995). Deux individus en provenance de la population du Burkina Faso dont l'un a été détecté dans la population du Niger

( $P=0,0000$ ) et l'autre dans la population du Ghana ( $P=0,007$ ). Ensuite, un individu en provenance de la population du Ghana a été détecté dans la population du Mali ( $P=0,0000$ ) et un dernier en provenance de la population du Sénégal a été détecté dans la population de la Côte d'Ivoire ( $P=0,0000$ ). Les origines probables de ces deux derniers ont été confirmées par les assignations bayésiennes partielles de Rannala et Mountain (1997).

L'assignation par la méthode des fréquences alléliques indique que 65,6% des individus de la population totale ont été correctement assignés à leur population d'origine contre 34,4% mal assignés. Chez la population du Mali et du Sénégal, le pourcentage des individus correctement assignés est de 100% contre 66,67% du Burkina Faso et du Niger et 80% de la Côte d'Ivoire. Cependant, dans la population du Ghana, aucun individu n'est correctement assigné.

### III.2.3.3 INFLUENCE DE LA MUTATION ET DU FLUX GENIQUE SUR LA DIFFERENCIATION GENETIQUE

Les valeurs des indices de différenciation ( $R_{ST}$  et  $F_{ST}$ ) ont été comparées afin de caractériser l'importance relative de la mutation et du niveau de flux génique dans la structure de la diversité observée. Les analyses ont indiqué une influence significative du flux de gènes à l'échelle globale te suivant 2 loci : 1A1 et 3A11 (Tableau XXXII).

**Tableau XXXII** : Influence de la mutation dans la différenciation génétique des SSR

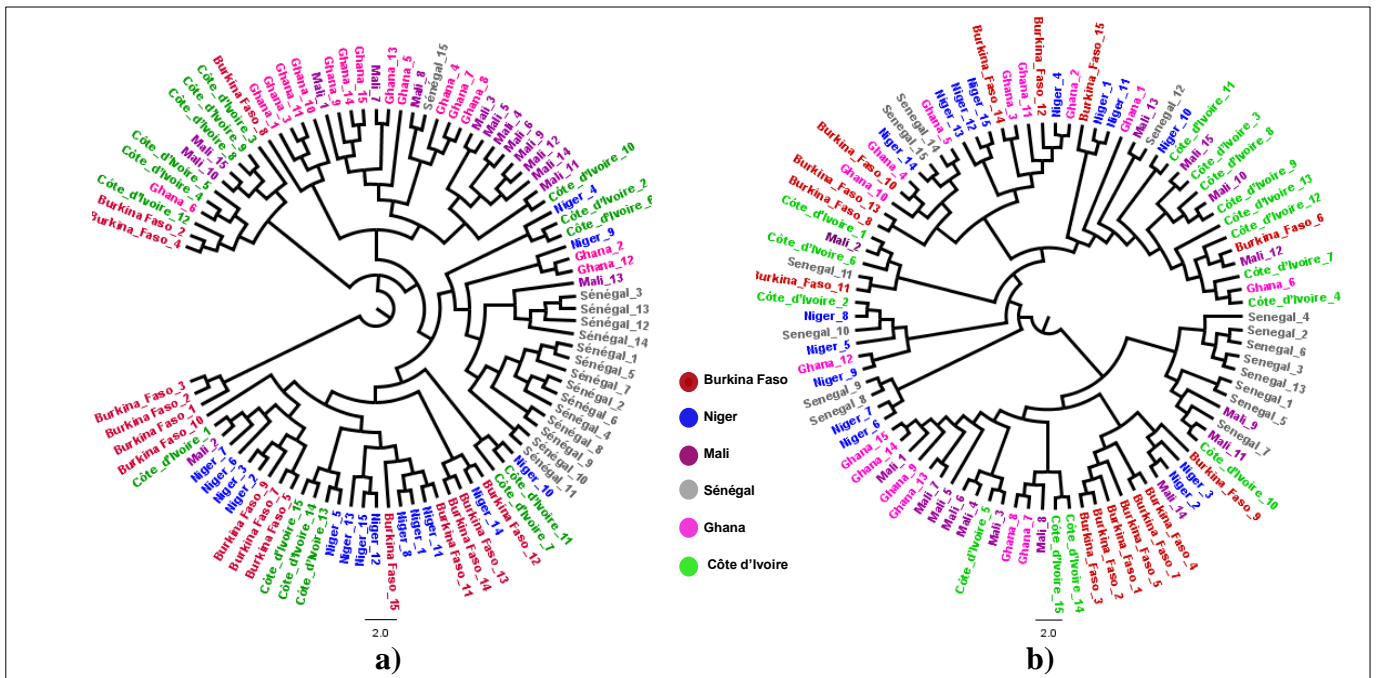
Loci	$R_{ST}$	$F_{ST}$	P-values $F_{ST}$ vs $R_{ST}$
Tous les loci	0,1040	0,1596	0,0002 ***
1A1	-0,0317	0,0440	0,0001 ***
1E1	0,1159	0,1182	0,0848 NS
1B1	0,1796	0,0941	0,0742 NS
3G7	0,8060	0,5400	0,7797 NS
3G1	0,1286	0,0509	0,3721 NS
3H6	0,0676	0,1279	0,3393 NS
3A11	-0,0003	0,0341	0,01447 *
1D10	0,0610	0,1373	0,1244 NS

NS:  $p > 0,05$  = non significatif ; \* :  $P < 0,05$  = significatif ; \*\* :  $P < 0,01$  = très significatif. \*\*\* :  $P < 0,001$  : hautement significatif.

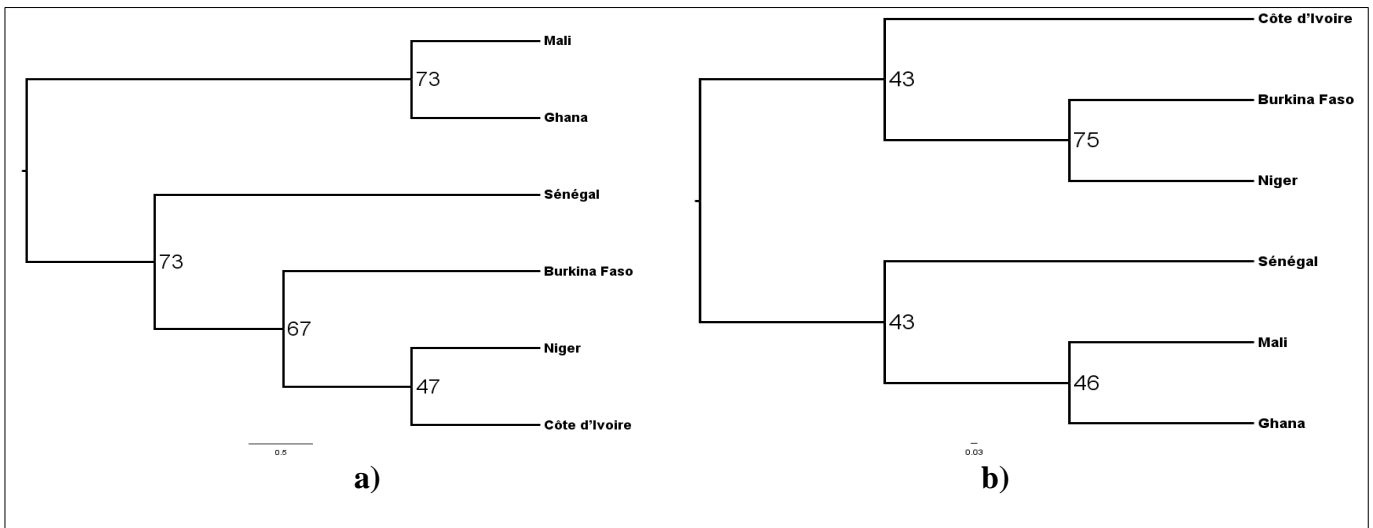
### III.2.4 RECONSTRUCTIONS PHYLOGENETIQUES

Du fait de la prépondérance du flux génique dans l'explication de la diversité génétique observée et de la signature d'un bottleneck démographique, nous avons représenté les reconstructions phylogénétiques grâce à 2 mesures de la distance génétique : la distance génétique angulaire DA (Nei *et al.*, 1983) et la distance des allèles partagés (DAS) (Chakraborty and Jin, 1993). Les arbres phylogénétiques confirment ainsi les résultats des analyses multivariées et bayésiennes mais aussi l'effet structurant du locus 3G7 qui est potentiellement sous sélection directionnelle (Figure 45 à 46). Les reconstructions phylogénétiques ont aussi groupé les populations dans un pattern similaire aux analyses de la classification ascendante hiérarchique (CAH) des matrices multi-loci du  $F_{ST}$ .





**Figure 45 :** Neighbor-Joining à partir de DAS : analyses avec (a) et sans (b) le locus 3G7



**Figure 46 :** Neighbor-Joining à partir de la distance de Nei (1983) (a) et à partir de DAS (b)

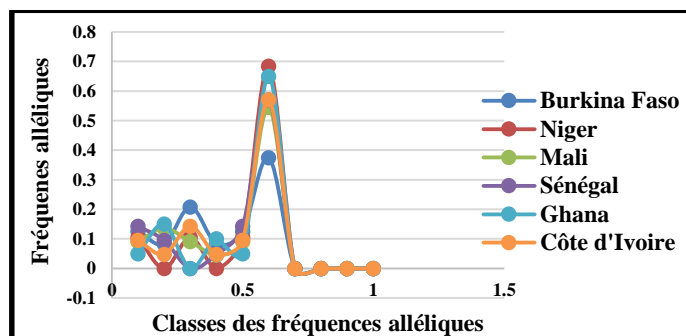
**III.2.5 EVOLUTION DEMOGRAPHIQUE : SIGNATURE DE GOULOT D'ETRANGLEMENT**

Les résultats de l'analyse du goulot d'étranglement génétique sous l'hypothèse nulle de l'équilibre dérive-mutation indiquent un excès d'hétérozygotie significatif à très significatif sous les 3 modèles IAM, SMM et TPM pour toutes les populations à l'exception de la population globale avec le modèle SMM (Tableau XXXIII). Ainsi, l'hypothèse nulle à l'équilibre mutation-dérive a été rejetée. La distribution des fréquences alléliques indique une répartition de la distribution en forme de L biaisé (Figure 47). Les allèles à fréquence modérée sont les plus nombreux (Figure 47).

**Tableau XXXIII :** Résultat des tests de signature de goulot d'étranglement récent

Paramètres	Modèles de mutation		
	SMM	IAM	TPM
<b>Populations</b>			
<b>Burkina Faso</b>	0,00391	0,00195	0,00195
<b>Niger</b>	0,00191	0,00195	0,00195
<b>Mali</b>	0,00391	0,00195	0,00195
<b>Sénégal</b>	0,00195	0,00195	0,00195
<b>Côte d'Ivoire</b>	0,00195	0,00195	0,00195
<b>Ghana</b>	0,00195	0,00195	0,00195
<b>Population globale</b>	0,27344	0,00195	0,00391

*P-values* des tests de Wilcoxon unilatéraux d'un excès d'hétérozygotie



**Figure 47 :** Spectre des fréquences alléliques des populations du Charançon du maïs

## II.3 DISCUSSION

Dans ce troisième chapitre, nous avons réalisé une étude complémentaire portant sur l'évaluation du flux génique et la structure génétique des populations de *S. zeamais* à travers l'analyse des loci microsatellites. En effet, les études récentes en biologie évolutive associent le plus souvent, deux types de marqueurs (nucléaire et mitochondrial, par exemple) dans le but d'avoir des informations plus exhaustives.

### III.3.1 VARIABILITE ALLELIQUE ET DIVERSITE GENETIQUE DES POPULATIONS

Les valeurs des indices de variabilité allélique et de polymorphisme intrapopulationnel de cette étude caractérisent une forte hétérogénéité génétique. En effet, les analyses ont indiqué un fort pourcentage de loci polymorphiques (95,83%) associé une richesse allélique globale ( $R_A=3,43$ ) et une variabilité allélique qui s'étend de 3 (3H6 et 1D10) à 7 allèles (1A1) pour une moyenne multilocus de 3,02 aboutissant à 33 allèles recensés sur 8 loci analysés, en d'un important indice de Shannon ( $I=0,85$ ). Ceci indique une variation génétique assez importante et suggère un bon potentiel adaptatif local. Ces valeurs sont dans la même gamme de valeurs comparées aux résultats de Corrêa *et al.* (2012 ; 2017) avec des populations originaires des Amériques, de l'Australie, de l'Asie et de l'Afrique (Mozambique) avec *S. zmais* et ceux de Thangaraj *et al.* (2016) avec *S. oryzae* issus de populations localisées en Chine, Australie et Kansas (USA). Par ailleurs, parmi ces loci, le locus 3G7 a la capacité informative la plus discriminante ( $PIC=0,68$ ) et le 1G3 est le locus le moins discriminant ( $PIC=0,40$ ). En moyenne, les valeurs des estimateurs ont démontré l'efficacité des marqueurs à discriminer les populations échantillonnées ( $PIC=0,54$ ) selon Botstein *et al.* (1980).

### III.3.2 DIVERSITE ET EQUILIBRE GENETIQUES DES POPULATIONS

Dans la population globale, aucun locus n'est en déséquilibre de liaison ; donc il n'y a aucune association non aléatoire entre génotypes par paires de loci. L'analyse de l'équilibre génétique de ces huit (8) loci a aussi indiqué que cinq (1B1, 3G7, 1A1, 3H6 et 3A11) présentent un écart à l'équilibre de Hardy-Weinberg ( $P < 0,05$ ) ; contre trois loci (1E1, 3G1, et 1D10) qui sont en équilibre génétique ( $P > 0,05$ ). Toutes les populations étudiées ont montré un écart significatif par rapport à l'équilibre de Hardy-Weinberg ( $P < 0,05$ ). Ce qui dénote qu'il y a un impact des forces évolutives sur les populations étudiées. Les résultats ont également montré un taux d'hétérozygotie moyenne observée ( $H_o=0,555$ ) qui est supérieure au taux d'hétérozygotie moyenne attendue ( $H_e=0,526$ ) mais non significative. Cela suggère globalement qu'il y a un excès mineur d'hétérozygotes et un déficit mineur en homozygotes au sein de la population globale par rapport aux valeurs attendues. Cette proportion de l'hétérozygotie moyenne observée est légèrement supérieure comparées aux valeurs trouvées par Corrêa *et al.* (2017) et Corrêa *et al.* (2012) au Mozambique. Il faut noter que l'hétérozygotie ou diversité allélique est un indice de performance écologique ou fitness d'un individu, c'est-à-dire sa

capacité à survivre et à se reproduire dans son biotope. En fait, l'hétérozygotie révèle une variabilité allélique d'un même gène et affaiblit la possibilité d'expression d'un allèle récessif délétère. Ceci pourrait s'avérer essentiel dans la garantie de survie des populations dans des conditions environnementales changeantes (Slate *et al.*, 2000 ; Hoglund *et al.*, 2002 ; Montaigne, 2011) car la diversité génétique d'une population est considérée comme étant le moteur de son adaptation (Montaigne, 2011). En outre, les analyses ont aussi indiqué un taux de consanguinité global modéré et hautement significatif de  $F_{IT}=8\%$ . Cette valeur du  $F_{IT}$  évoque une influence modérée de la consanguinité. Ainsi, les populations ont présenté un bon potentiel de viabilité et d'adaptation. Ce résultat est confirmé par la valeur globale du  $F_{IS}$  qui est significativement négative ( $F_{IS}=-0,063$ ). Cette valeur de  $F_{IS}$  négative indique que les individus de population globale sont moins liés que ce à quoi il fallait s'attendre avec un modèle de Hady-Weinberg. Ce qui serait lié, par exemple, à un flux migratoire important provoquant un effet d'hétérogamie. Par ailleurs, la comparaison des moyennes des indices de la diversité génétique indique que l'hétérozygotie observée et la richesse allélique de la zone sahélo-soudanienne (Nord) sont supérieures à celles trouvées à la région guinéo-congolaise (Sud). Le Nord est donc génétiquement plus riche que le Sud. Ce qui confirme les résultats des analyses avec les séquences d'ADN (Ndiaye et Sembène, 2018).

### III.3.3 REPARTITION HIERARCHIQUE ET SPATIALE DE LA DIVERSITE GENETIQUE

Globalement, les huit loci microsatellites qui ont été utilisés ont contribué significativement à la différenciation génétique globale. La comparaison des indices du  $F_{ST}$  (niveau de connectivité), du  $F_{IS}$  (consanguinité) et l'analyse du nombre de migrants par génération  $Nm$  (flux de gènes) de la population globale ont révélé la présence d'une structure génétique relativement forte, un flux génique asymétrique et globalement un niveau mineur de consanguinité. En effet, les résultats de l'analyse hiérarchique de la variance moléculaire (AMOVA) ont révélé que l'essentiel de la variation totale (97 à 98% ;  $P=0,000$ ) est due aux caractéristiques individuelles contre 8 à 9% ( $P = 0,000$ ) issue de la divergence entre populations. En outre, la valeur de la différenciation génétique globale est importante et hautement significative ( $F_{ST}=16\%$ ) révélant une forte différenciation entre les populations. Cette forte structuration de la population globale est confirmée par la valeur du déficit d'hétérozygotes global ( $F_{IT}=10,7\%$ ) qui est modérée et hautement significative. Cette différenciation et structure génétiques observée est donc due à la variance des fréquences alléliques entre les populations, appelée effet Wahlund (Wahlund, 1928). Ensuite, la comparaison de la diversité génétique intra-population entre la région sahélo-soudanienne et la région guinéo-congolaise a indiqué une diversité génétique plus importante dans le Sud que dans le Nord. En effet, les populations du Nord sont génétiquement moins différenciées et moins structurées comparées à celles du Sud. La région guinéo-congolaise montre une valeur de consanguinité modérée du  $F_{IS}=7,5\%$  et une valeur très importante de  $F_{ST}=22,1\%$  associées à un déficit en hétérozygotes global de 28%. Ce qui suggère

une connectivité très faible entre les populations du Ghana et de la Côte d'Ivoire. Ces valeurs positives du  $F_{IS}$  évoquent également une structure génétique cryptique. Alors que dans la région sahélo-soudanienne, les valeurs du  $F_{IS}=-0,123$  et du  $F_{ST}=15,6\%$  sont significatives. Ces valeurs, associées à un déficit d'hétérozygotie global faible 5,2%, mettent en exergue une absence globalement de consanguinité favorisée par un chevauchement des populations à travers des échanges d'individus. En plus, les valeurs intra-zonales ( $F_{ST}=15,6$  et  $F_{ST}=22,1\%$ ) entre les régions biogéographiques sont supérieures aux valeurs inter-zonales ( $F_{ST}=3,1\%$ ). Ainsi, la population globale constitue, en réalité, une métapopulation qui est une structure dynamique composée de plusieurs populations séparées spatialement et/ou temporellement mais connectées par le processus de la migration. De plus, cette forte diversité génétique du Nord associée au nombre de ses allèles privés, le Sud ne présentant pas d'allèles privés, pourraient déterminer le Nord comme étant la zone originelle de l'infestation. Cependant, cette hypothèse n'est pas conforme à celle des zones refuges de la savane humide dans le Sud (région Guinéo-Congolaise) (Maley, 1996 ; 2001 ; Léal, 2004) pendant le dernier maximum de glaciation. En effet, ces zones refuges sont des espaces où un petit nombre de lignées survivraient et se réfugièrent pendant les périodes défavorables du Pléistocène (aridité, glaciation). Le Nord étant considéré comme zone de recolonisation après le retour aux conditions propices à l'expansion. En outre, ces zones de forte diversité génétique reflèteraient aussi une ambiguïté car elles peuvent correspondre aussi bien à une zone refuge qu'à une zone de convergence de voies de recolonisation des lignées issues de différentes zones refuges à la fin du dernier maximum glaciaire (Dick *et al.*, 2003 ; Petit *et al.*, 2003 ; Ley *et al.*, 2014). En outre, l'analyse des matrices de différenciation et de divergence révèlent que le locus potentiellement sous sélection avait tendance à isoler ou à approfondir la divergence génétique entre le Ghana et le Mali du groupe formé par la Côte d'Ivoire, le Niger, le Burkina Faso et le Sénégal. Ces résultats ont été démontrés par ceux de l'analyse SAMOVA (Ndiaye et Sémbène, 2018). En effet, cette analyse avait montré un isolement statistiquement non significatif du Ghana puis du Mali du reste des autres populations selon le marqueur (*Cyt-B*, *COI* et séquences concaténées). Par ailleurs même si le cluster formé par le Ghana et le Mali est résolutif, leur différenciation est significativement importante ( $F_{ST}=12,6\%$ ). Ensuite, les analyses de l'AFC et de la PCoA ont confirmé la structuration de la population globale et le flux génique asymétrique entre populations. En effet, ils ont montré un regroupement les populations du Niger, du Burkina et de la Côte d'Ivoire montrant ainsi que ces populations sont très proches génétiquement ; ce qui pourrait démontrer qu'ils ont eu la même histoire évolutive. Ils ont aussi mis en exergue une différenciation et une individualisation des populations du Ghana, du Mali et du Sénégal. Ceci pourrait refléter donc une histoire évolutive différente de celle un regroupement des populations du Burkina Faso, du Niger et de la Côte d'Ivoire

La comparaison des valeurs des indices de différenciation génétique avec celles obtenues après permutations du  $R_{ST}$  (Hardy *et al.*, 2003) et du  $N_{ST}$  (Burban *et al.*, 1999) a permis de déterminer qu'à

l'échelle globale le taux de mutation n'est pas suffisamment importante pour impacter la structure génétique par la présence d'un pattern (signal) phylogéographique (Hardy *et al.*, 2003 ; Pons & Petit, 1996). Par conséquent, il n'y a pas de corrélation entre la phylogénie des différents allèles et leur géographie. Cela suggère que la différenciation observée est essentiellement dû à la migration et à la dérive génétique et donc une dispersion par un flux multidirectionnel grâce au commerce des grains entre populations (rurales surtout) ; ceci pouvant expliquer l'absence de corrélation entre la distance génétique et la distance géographique. En effet, la contribution de la diversité génétique moyenne intra-population ( $H_S$ ) à la diversité génétique totale ( $H_T$ ) est supérieure à celle de la diversité génétique inter-population ( $D_{ST}$ ). Cette caractéristique génétique est en phase avec le dynamisme commercial de la sous-région ouest-africaine qui est marquée par une multitude d'organismes d'intégration favorisant ainsi les échanges commerciaux. De plus, cette prédominance de flux génique par rapport à la mutation met aussi en évidence que une infestation récente de la sous-région ouest-africaine car le temps de divergence est court pour permettre aux mutations d'opérer (Hardy *et al.*, 2003). Ceci confirme les résultats obtenus avec l'inférence issue des analyses combinées avec les diversités haplotypique et nucléotidique (Ndiaye et Sembène, 2018). Cependant, à l'intérieur de la région guinéo-congolaise (Côte d'Ivoire vs Ghana), il y a un signal phylogéographique. Ceci peut être expliqué par le fait que même si ces 2 populations sont frontalières, ils appartiennent officiellement à des structures d'intégrations économiques différentes. Ils sont ainsi officiellement moins susceptibles d'avoir des pratiques commerciales plus importantes comparées aux autres pays de la CEDEAO.

Par ailleurs, les analyses de structure génétique basées sur un modèle génétique (modèle Hardy-Weinberg, analyses avec l'algorithme structure) ou non (analyses multivariées, principalement par la discrimination des composantes principales, DAPC) suggèrent la présence d'une structure génétique dans la population globale par l'existence de quatre *clusters* principaux. Les résultats des 2 types d'analyses sont congruents mais présentent un contenu individuel plus ou moins différent. Il faut souligner que les méthodes bayésiennes sont sujettes à moins de biais lorsqu'il n'y a pas d'isolement par la distance et la présence de barrières génétiques est correctement identifiée (Frantz *et al.*, 2009). Ce qui conforte nos résultats. Parmi les quatre *clusters* génétiques, il y a le groupe 1 qui rassemble majoritairement les individus issus du Burkina Faso, du Niger et de la Côte d'Ivoire et le groupe 2 qui englobe les spécimens du Sénégal. Tandis que le groupe 3 catégorise essentiellement des échantillons du Ghana et enfin le groupe 4 est caractérisé essentiellement par les spécimens du Mali. Cette structuration plus ou moins importante de la population globale n'est pas en phase avec celle de (Corrêa *et al.*, 2017) qui ont noté une présence très faible structure génétique mais elle est comparable avec celle de Thangaraj *et al.* (2016) réalisé sur *S. oryzae*.

Il faut aussi souligner qu'il y a une certaine congruence entre les haplogroupes identifiés par les séquences d'ADN mitochondrial et les 4 *clusters* considérés par les marqueurs nucléaires des SSR. Certaines distinctions de *clusters* ont été affinées par ces derniers comme le *cluster* essentiellement

constitués d'individus du Mali. Il y a aussi, le cluster majoritairement formés d'individus issus du Sénégal qui correspondrait à l'haplogroupe formé par les individus du Sénégal et de la République de Guinée. L'haplogroupe privé du Ghana est confirmé ; de même que l'affinité entre les populations du Burkina Faso et du Niger avec certains individus de la Côte d'Ivoire.

### III.3.4 POPULATIONS DE *SITOPHILUS ZEAMAI* SOUS SELECTION DIRECTIONNELLE

Comprendre comment les forces environnementales engendrent une variation génétique adaptative est un défi pour la lutte contre les insectes ravageurs. Ainsi, l'identification de loci en sélection divergente est une étape clé dans la compréhension du processus évolutif, car ces loci sont responsables des variations génétiques qui affectent la condition physique dans différents environnements (Feng *et al.*, 2015). Parmi les loci étudiés, 3 ont été détectés comme étant potentiellement sous sélection directionnelle ; cependant, le locus 3G7 a été systématiquement identifié comme étant sous sélection avec les trois méthodes utilisées. D'après Pérez-Figueroa *et al.* (2010), le programme BAYESCAN, en le comparant avec 2 autres programmes (DFDIST, DETSELD) pour détecter les loci sous sélection directionnelle à partir de balayages génétiques à l'aide de marqueurs dominants, semble être plus efficace que les autres méthodes pour détecter ces loci avec moins de 1% de faux positifs. En plus, les indices de consanguinité, de différenciation et de structuration génétiques ont indiqué que le locus 3G7, qui est potentiellement sous influence d'une sélection positive directionnelle, a un impact significatif dans le niveau de consanguinité au sein des populations, dans la différenciation et la structuration des populations. En effet, les valeurs de ces indices diminuent fortement en absence du locus 3G7 dans les analyses. En absence du locus 3G7, toutes les valeurs du  $F_{IS}$  sont négatives. Lorsqu'il est inclut dans l'analyse, ce locus induit des valeurs de  $F_{IS}$  positives :  $F_{IS}=0,003$  ( $P > 0,05$ ),  $F_{IS}=0,041$  ( $P > 0,05$ ) et  $F_{IS}=0,102$  ( $P > 0,05$ ) respectivement pour les populations du Burkina Faso, du Ghana et de la Côte d'Ivoire. Ce qui indique un niveau mineur de consanguinité car ces valeurs ne sont pas statistiquement significatives. D'autre part, les analyses du  $F_{ST}$  incluant ce locus montrent des valeurs oscillant entre 16,3 et 17,8%. Cependant, si ce locus est éliminé de l'analyse la valeur du  $F_{ST}$  diminue jusqu'à 8%. Par ailleurs, les résultats l'analyse hiérarchique de la variance moléculaire (AMOVA) a ont montré l'effet de la sélection directionnelle du locus 3G7 qui a entraîné une diminution de la variance génétique au sein des populations qui passe de 98% à 90% tandis que la variance génétique entre des populations a augmenté de 9% à 17% de la variation génétique réside dans la différenciation génétique entre les populations. Cela indique que le locus 3G7 a accentué la divergence entre les populations. Selon Cavalli-Sforza (1966) et Lewontin et Krakauer (1973), les loci sous une sélection directionnelle devraient indiquer de larges différences entre populations. D'après Slatkin et Barton (1989), la  $F_{ST}$  devrait être plus sensible à une adaptation locale (sélection naturelle) car la sélection entraîne un changement majeur dans les fréquences des allèles communs. En plus, les analyses au locus 3G7 ont

indiqué que les valeurs du  $R_{ST}$  et du  $N_{ST}$  sont significativement supérieures à leurs valeurs permutées respectivement ( $pR_{ST}=0,044$  et  $pN_{ST}=0,035$ ). En définitif, le locus 3G7 a été détecté comme étant un locus possédant des allèles diagnostiques particulièrement le 3G7\_185 et le 3G7\_212 car ce locus a participé fortement à la discrimination des *clusters* génétiques. Par ailleurs, même si, le charançon du maïs est caractérisé par des chevauchements de générations et un comportement grégaire, ce qui accentue la consanguinité ; il montre également une grande aptitude à la dispersion naturelle qui est accentuée par une migration passive via les activités anthropiques (Corrêa *et al.*, 2012). Cette sélection directionnelle constitue donc un avantage dans la diminution du succès invasif des populations du charançon du maïs. Selon Saccheri *et al.* (1998), la capacité adaptative due à la dispersion peut être diminuée, parfois drastiquement, par une sélection adaptative car la sélection a tendance à augmenter la consanguinité au sein des sous-populations, de même que la différenciation et la structure génétiques des populations. Ce qui entraîne, à long terme, une rupture de connectivité entre les populations et une érosion progressive de leur biodiversité ; constituant ainsi une menace à la viabilité des individus et donc des populations.

### III.3.5 EVOLUTION DEMOGRAPHIQUE

Sous l'hypothèse équilibre dérive-mutation, les résultats de l'analyse du goulot d'étranglement génétique ont indiqué globalement un excès d'hétérozygotie significatif à très significatif sous les 3 modèles (IAM, TPM et SMM;  $p < 0,01$ ) à l'exception de la population globale où la valeur du test n'est pas significative. Selon Cornuet et Luikart (1996), un goulot d'étranglement est fortement détecté par le modèle IAM, moyennement détecté par TPM et faiblement ou rarement par le modèle SSM. En outre, en cas de tests très significatifs pour les trois procédures, on peut être assez confiant (Meeûs, 2012). Ainsi, l'hypothèse nulle à l'équilibre mutation-dérive a été rejetée. Cet excès d'hétérozygotie significatif suggère la possibilité d'anciens et récents événements démographiques, i.e., un effet bottleneck (Cornuet et Luikart, 1996 ; Riccioni *et al.*, 2010). Selon Luikart et Cornuet (1998), un *bottleneck* est, en fait, plus probablement détectable quand, au moins, 8 à 10 loci microsatellite polymorphes sont utilisés ; ce qui est conforme avec nos données. En plus, la distribution des fréquences alléliques observées a démontré que les populations de *S. zeamais* ont subi un goulot d'étranglement génétique. En effet, le descripteur de la répartition de la fréquence allélique ("mode-shift") qui permet de discriminer des populations en *bottleneck* des populations stables (Luikart *et al.*, 1997) a montré que cette distribution suit une forme en L biaisée indiquant que les allèles à fréquence modérée ont été les plus nombreux. Ces *bottleneck* démographiques associés à la dérive génétique ont entraîné une réduction de la variabilité allélique. Cependant, cet épisode de réduction de la population est suivi d'une expansion démographique révélée par l'existence d'un nombre important d'allèles rares (environ 12%). De ce fait, ces résultats ont confirmé ceux trouvés avec les séquences d'ADN en ce qui concerne la démographie du passé des populations de *S. zeamais*



en Afrique de l'Ouest (Ndiaye et Sembène, 2018).

### III.3.6 FACTEURS GENETIQUES INFLUENÇANT LA DYNAMIQUE

Ainsi, des analyses effectués, les principaux facteurs génétiques influençant la dynamique des populations de *S. zeamais* en Afrique de l'Ouest sont :

- une hétérozygotie et une variabilité allélique importantes conférant un pool génétique substantielle gage d'une adaptation des populations dans leurs environnements ;
- un effet Wahlund (Wahlund, 1928) due à la variance des fréquences alléliques entre les populations ;
- une présence potentielle d'une sélection positive directionnelle du locus 3G7 possédnt au moins deux allèles diagnostiques (3G7\_185 et 3G7\_212) ;
- une dérive génétique associée une migration à travers un flux génique asymétrique et multidirectionnel mis en évidence par la présence des allèles rares (12,12%) qui sont considérés comme des indicateurs de niveau du flux génétique (Slatkin, 1985 ; Saavedra *et al.*, 1993) car leur présence massive dans la population induit un flux génique important ; et par l'analyse des réseaux de migration relative directionnelle qui a montré que un flux génique asymétrique avec la population du Burkina Faso comme point de ralliement ;
- une consanguinité modérée au sein des sous-populations attestant de leur viabilité.

Dès lors, la dispersion mise en évidence serait liée à un processus de migration passive liée aux facteurs anthropiques comme le commerce des grains. Ce dernier constitue un facteur de dispersion qui est un processus essentiel au maintien à long terme des populations. De ce fait, la dispersion joue un rôle clé dans les systèmes naturels en modélant le dynamisme spatial et évolutif des populations (Garant *et al.*, 2007 ; Chaîne *et al.*, 2013) ; et surtout elle représente un facteur avantageux car augmentant le fitness par diminution de la dépression consanguine et par accentuation du flux génétique et donc du brassage génétique. En plus, elle contrebalance les effets néfastes de la dérive génétique en permettant la persistance des petites populations et la propagation des traits adaptatifs (Manel *et al.*, 2003). Cette dispersion constitue donc un facteur majeur du succès invasif des insectes ravageurs tels que le charançon du maïs. Elle interagit avec la dérive génétique et la sélection et leur résultante module le potentiel adaptatif et la viabilité des populations (Frankham, 2005). Ce flux génique a été également identifié comme étant prédominant par rapport à la mutation dans la structuration de la variabilité génétique observée.

## CONCLUSION PARTIELLE

Cette seconde partie constitue la première étude utilisant les performances des marqueurs SSR pour l'analyse génétique des populations de *S. zeamais* en Afrique de l'Ouest. Les analyses ont montré que les marqueurs moléculaires utilisés sont efficaces pour une estimation cohérente de la diversité et de la structure génétiques dans la gestion du bioagresseur *S. zeamais*. Le génotypage de 10 loci SSR a été donc testé sur 90 individus répartis dans 6 pays d'Afrique de l'Ouest. Les loci analysés ont présenté un fort taux de polymorphisme et une capacité informative importante révélant ainsi une importante richesse allélique et une hétérozygotie substantielle associée un coefficient de consanguinité modéré. Ce qui indique un bon niveau de viabilité et un pool génétique assez important des populations de *S. zeamais* dans la sous-région ouest-africaine et donc leur capacité de survie et d'adaptation. Ces marqueurs SSR ont aussi mis en évidence l'existence d'une différenciation génétique et d'un flux génique inégalement répartis. Ce flux asymétrique est multidirectionnel mettant en exergue l'impact des échanges commerciaux et donc des structures d'intégration politico économique sur la structuration de la variabilité génétique. En effet, il existe une affinité génétique entre le Burkina Faso, le Niger et la Côte d'Ivoire et des différenciations génétiques significatives entre le Mali, le Ghana et le Sénégal. En conséquence, il existe une structure génétique sans modèle phylogéographique. En outre, la caractérisation des relations génétiques par des analyses multivariées et bayésiennes de *clustering* quantitatif ont mis en exergue la présence de quatre *clusters* génétiques fortement différenciés. Ces populations qui sont issues d'un *bottleneck* génétique récent, sont sous l'influence potentielle d'une sélection directionnelle qui, à long terme et avec une diminution voire une absence de la dispersion passive à travers l'action anthropique, pourrait constituer un facteur favorable à une forte consanguinité et une rupture progressive de la connectivité entre populations diminuant ainsi la viabilité et la capacité adaptative des populations. Les connaissances générées sur les niveaux de diversité et la structure de la population du matériel génétique de *S. zeamais* constituent une contribution importante à la gestion intégrée des populations de ce bio-agresseur et à la conservation des cultures et denrées stockées.

## CONCLUSION GENERALE

En entomologie agricole, la diversité et la structure génétiques des populations constituent des facteurs-clés car permettent de déterminer la viabilité, la connectivité et donc la capacité adaptative des insectes ravageurs. En plus, les études antérieures ont mis en exergue que ces deux données quantifiant la capacité de survie et l'adaptation des populations peuvent être affectés par des facteurs, qui agissent souvent en combinaison, tels que les changements climatiques, les paramètres environnementaux et écologiques, les barrières naturelles, les activités anthropiques à travers les échanges commerciaux et le flux génique. La finalité de ces études étant la mise en place et l'application de stratégies de gestion afin de maintenir ces organismes nuisibles en dessous d'un seuil économiquement acceptable.

Les inférences de cette étude basées sur des analyses multivariées et bayésiennes à l'aide des données de séquençage et de génotypage ont indiqué la présence d'une multitude d'haplotypes dont l'haplotype majoritaire est central. Parallèlement, les marqueurs microsatellites disposant d'une bonne capacité informative ont révélé une forte variabilité allélique et une importante diversité génétique des populations. Ce qui dénote d'un bon potentiel de viabilité et d'adaptation dans un environnement changeant. Les analyses ont indiqué que l'essentiel de la variance moléculaire est due aux caractéristiques individuelles d'une population ; qu'une proportion moins importante est liée aux différences entre populations alors que très peu dérive de la différenciation entre régions biogéographiques. Tandis que le flux génique est asymétrique et multidirectionnelle (infestation croisée) induisant une absence de corrélation entre la distance génétique et la distance géographique, d'où la présence d'une structure génétique sans pattern phylogéographique. La configuration de la variabilité génétique observée est le résultat de l'action combinée d'un déficit en hétérozygotes dû à la structuration de la population globale (l'effet Wahlund), d'une sélection directionnelle, d'un flux génique asymétrique et multidirectionnel, d'une dérive génétique et d'une dispersion passive à travers les échanges commerciaux. Les inférences bayésiennes et multivariées associées aux résultats des datations moléculaires ont suggéré l'existence de 4 à 5 *clusters* principaux qui se sont séparés après un *bottleneck* suivi d'une expansion démographique modérée depuis la période du pléistocène.

Finalement, il y a une concordance entre les résultats des analyses de marqueurs de l'ADNmt avec ceux des SSR : une différenciation génétique inégale et un flux génique multidirectionnel, plus affiné avec les SSR, d'où la présence d'une structure génétique sans signal phylogéographique. La variabilité génétique est plus importante dans le Nord d'où viendrait l'infestation. Ainsi donc, les connaissances générées sur les niveaux de la diversité et de la structure génétiques des populations de *S. zeamais* en Afrique de l'Ouest particulièrement et l'aperçu en Afrique centrale constituent une contribution importante à la gestion intégrée de cet bio-agresseur et donc à la bonne conservation des cultures et denrées stockées pour une assurance d'une sécurité alimentaire.

## RECOMMANDATIONS

De cette étude, nous pouvons tirer les recommandations suivantes :

- une meilleure connaissance de la biodiversité, les spécialistes peuvent identifier des paramètres à influencer afin de diminuer l'effet néfaste des ravageurs des produits agricoles ;
- une nécessité de définir et de renforcer les politiques d'échanges à travers des normes hygiéniques et sanitaires à mesure de limiter voire éliminer les flux de gènes entre populations d'insectes ravageurs à travers le commerce des grains ;
- une sensibilisation des populations rurales sur l'impact négatif des échanges de denrées infestées ;
- une influence prépondérante des politiques dans le respect des normes sanitaires et phytosanitaires sur les produits d'importation ;
- une gestion intégrée des populations d'insectes ravageurs en Afrique de l'Ouest et centrale devrait être réalisée en tenant compte de la différence structurelle des caractéristiques génétiques de *S. zeamais* dans les différents pays ou groupes de pays. En effet, il y a une structuration de la variabilité génétique. Cette dernière dont dépend le potentiel de survie des populations est renforcée par une augmentation du flux génétique. Dès lors, en diminuant voire éliminant les facteurs de dispersion par une migration passive favorisée par les échanges non formalisés (fraudes), il est possible de participer indirectement à l'érosion de cette variabilité génétique des populations et donc à amoindrir leur survie et leur capacité adaptative.

En outre, il opère dans ces populations de *S. zeamais* un processus érosif naturel de la diversité génétique à savoir la sélection directionnelle qui entraîne une diminution de la variabilité génétique et donc de la viabilité et de la capacité adaptative de ces insectes ravageurs. Cette force évolutive a tendance à augmenter la consanguinité, à diminuer la diversité génétique et à favoriser une rupture de connection entre les populations. Ceci constituerait un frein à leur prolifération car la variabilité génétique, la connectivité et la dispersion sont des facteurs cruciaux dans la persistance, l'accroissement et la viabilité des populations. Ainsi, la dispersion par le commerce des grains constitue un des tampons d'Achille du potentiel de conservation des produits post-récoltes.

Ainsi, cette étude a contribué :

- au plan scientifique, par une participation à la connaissance de la biodiversité de l'entomofaune de l'Afrique subsaharienne.
- au plan économique, par une contribution qui est axée sur le développement durable en mettant en exergue le rôle crucial des échanges commerciaux dans le renforcement du potentiel destructeur des insectes ravageurs en proposant de mieux formaliser les activités commerciales selon des normes hygiéniques rigoureuses ;
- au plan social, par une participation à l'élaboration des stratégies afin d'assurer la sécurité alimentaire et nutritionnelle et donc au statut sanitaire et indirectement à la stabilité sociale.

## PERSPECTIVES

Afin de mieux appréhender la biologie évolutive des insectes phytophages en général, et de *S. zeamais* en particulier, dans l'optique de déterminer un maillon faible dans leurs mécanismes de nuisance des cultures et denrées stockées, il serait intéressant d'envisager les thématiques suivantes :

- étude de la structure génétique et du flux génique à d'autres populations et zones biogéographiques d'Afrique subsaharienne ;
- caractérisation de la morphométrie de *S. zeamais* en fonction des zones agro-écologiques en Afrique de l'Ouest voire en Afrique subsaharienne ;
- étude de l'écophysiologie des populations de *S. zeamais* dans différents pays ou zones biogéographiques ;
- caractérisation et identification de nouveaux loci pour *S. zeamais* et vérifier leur transférabilité aux autres espèces de *Sitophilus* spp ;
- détermination des effets des extraits de plantes naturelles dans la protection des cultures contre *S. zeamais*...

## REFERENCES BIBLIOGRAPHIQUES

1. **Adedire C. O., (2001):** Biology, Ecology and Control of Insect Pests of Stored Cereal Grains. *in:* Ofuya TI, Lale NES, Editors. Pests of Stored Cereals and Pulses in Nigeria: Biology, Ecology and Control. Dave Collins Publications.
2. **Adedire C. O. and Lajide L., (2003):** Ability of extracts of ten tropical plant species to protect maize grains against infestation by the maize weevil, *Sitophilus zeamais* during storage. *Niger. J. Exp. Biol.*, 4 (2): 175-179.
3. **AGRHYMET, (2010) :** Le Sahel face aux changements climatiques. Enjeux pour un développement durable. Bulletin Mensuel, Numéro spécial. 43 p.
4. **Akol A. M., Chidege M. Y., Talwana H. A. L. and Mauremootoo J. R., (2011):** *Sitophilus zeamais* Motschulsky, 1855. Maize Weevil. Makerere University, Tropical Pesticides Research Institute, Makerere University and BioNET-INTERNATIONAL. Disponible sur :
5. [http://keys.lucidcentral.org/keys/v3/eafrinet/maize\\_pests/key/maize\\_pests/Media/Html/Sitophilus\\_zeamais\\_Motschulsky\\_1855\\_-\\_Maize\\_Weevil.htm](http://keys.lucidcentral.org/keys/v3/eafrinet/maize_pests/key/maize_pests/Media/Html/Sitophilus_zeamais_Motschulsky_1855_-_Maize_Weevil.htm). Consulté le 7/12/2015.
6. **Alcala N., Goudet J. and Vuilleumier S., (2014):** On the transition of genetic differentiation from isolation to panmixia: what we can learn from  $G_{ST}$  and  $D$ . *Theor. Popul. Biol.* 93, 75–84.
7. **Alleoni B. and Ferreira W., (2006):** Control of *Sitophilus zeamais* Mots., 1958 and *Sitophilus oryzae* (L., 1763) weevils (*Coleoptera*, *Curculionidae*) in stored rice grain (*Oryza sativa* L.) with insecticide pirimiphos methyl (Actellic 500 CE). 9<sup>th</sup> International Working Conference on Stored Product Protection. General Session on Stored Grain Protection. PS 10-17; pp. 4. 1234-1241.
8. **Alonso-Zarazaga M. A. and Lyal C. H. C., (1999):** A World catalogue of family and genera of *Curculionoidea* (*Insecta*, *Coleoptera*) (Excepting *Scolitidae* and *Platipodidae*). Barcelona: *Entomopraxis*. 315 p.
9. **Andolfatto P., (2001):** Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Mol Biol Evol.*, (18): 279–290.
10. **Anonyme, (2009) :** Les invertébrés continentaux des pays de la Loire-Gretia, document online, 18p.
11. **CABI, (2018) :** *Sitophilus oryzae*. Disponible sur : <http://www.cabi.org/isc/datasheet/10887> Consulté le 08/03/2018.
12. **CABI, (2017) :** *Sitophilus zeamais*. Discipline sur <https://www.cabi.org/isc/datasheet/10926> Consulté le 08/03/2018.
13. **Anonyme, (2013a):** Rice weevil, *Sitophilus oryzae* (L.) - Canadian Grain Commission. Disponible sur <http://www.grainscanada.gc.ca/storage-entrepouse/pip-irp/rw-cr-eng.htm>. Consulté le 26/01/2015.
14. **Anonyme, (2013b) :** Maize weevil *Sitophilus zeamais* (Mots.) - Canadian Grain Commission. Disponivble sur : <http://www.grainscanada.gc.ca/storage-entrepouse/pip-irp/mw-cr-eng.htm>. Consulté le 31/01/2015.
15. **Anonyme, (2016):** Diagnostic Methods for Maize weevil, *Sitophilus zeamais*. PaDIL - Plant Biosecurity Toolbox. 16 pages Disponible sur : <http://pbt.padil.gov.au/index.php?q=node/193andpbtID=217>. Consulté le 8/08/2016.
16. **Anselme C., (2006) :** Réponse immunitaire de l'hôte dans la symbiose bactérienne intracellulaire du charançon *Sitophilus zeamais*. Thèse de doctorat. INSA de Lyon. 179 p.
17. **Anselme C., Vallier A., Balmand S., Fauvarque M.-O. and Heddi A., (2006):** Host PGRP gene expression and bacterial release in endosymbiosis of the weevil *Sitophilus zeamais*. *Applied and Environmental Microbiology* (72): 6766-6772.
18. **Antao T., Lopes A., Lopes R. J, Beja-Pereira A. and Luikart G. (2008):** LOSITAN: A workbench to detect molecular adaptation based on a  $F_{st}$ -outlier method. *BMC Bioinformatics*, (9): 323.
19. **Anzala F. J., (2006) :** Devant le contrôle de la vitesse de germination chez le maïs (*Zea mays* L.) : étude de la voie de biosynthèse des acides aminés issus de l'aspartate et recherche de Quantitative Trait Loci (QTLs). Thèse de Doctorat - Université d'Angers, INH, INRA. 148 p.
20. **Arbogast R. T. and Mullen M. A., (1990):** Interaction of maize weevil (*Coleoptera: Curculionidae*) and parasitoid *Anisopteromalus calandrae* (*Hymenoptera: Pteromalidae*) in a small bulk of stored corn. *J. Econ. Entomol.*, (83): 2462–8.

21. **Arbonnier M., (2004):** Trees, Shrubs and Lianas of West African Dry Zones CIRAD, Margraf Publishers.
22. **Avise J. C., Arnold J., Ball R. M., Bermingham E., Lamb T., Neigel J. E., Reeb C. A. and Saunders N. C., (1987):** Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics.
23. **Avise, J. C., (1994):** Molecular Markers, Natural History, and Evolution. Chapman and Hall, (New York). 511 p.
24. **Avise J. C., (2000):** Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.
25. **Avise JC (2004):** Molecular markers, natural history, and evolution, 2nd edn. Sinauer, Sunderland, MA.
26. **Avise J. C., (2009):** Phylogeography: retrospect and prospect. *Journal of Biogeography*, (36): 3–15.
27. **Baliraine F. N., Bonizzoni M., Guglielmino C. R., Osir S. E., Lux A., Mulaa F. J., Gomulski L. M., Zheng L., Quilici S., Gasperi G. and Malacrida A. R., (2004):** Population genetics of the potentially invasive African fruit fly species, *Ceratitis rosa* and *Ceratitis fasciventris* (Diptera : Tephritidae). *Molecular Ecology*, (13): 683-685.
28. **Ballard J. W. O. and Whitlock M. C., (2004):** The incomplete natural history of mitochondria. *Molecular Ecology*, (13) : 729–744.
29. **Ballouche A, Akoègninou A, Neumann K, Salzmann U. and Sowunmi A. 2000 :** Le Projet “Dahomey Gap”: une contribution à l’histoire de la végétation du Sud-Bénin et du Sud-Ouest du Nigeria. *Berichte des Sonderforschungsbereichs 268, Band 14, Frankfurt A.M;* 237-251.
30. **Bandelt H. J., Forster P. and Röhl A., (1999):** Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, (16): 37–48.
31. **Barthe S., (2012) :** Signatures de transitions démographiques des arbres de la Forêt Tropicale Humide du plateau des Guyanes. Thèse de Doctorat. Université des Antilles et de la Guyane, 218p.
32. **Barton NH. and Slatkin M., (1986):** A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56: 409-415.
33. **Beaumont M. A., Zhang W., and Balding D. J., (2002):** Approximate Bayesian computation in population genetics. *Genetics*, (162): 2025–2035.
34. **Beaumont M. and Rannala B., (2004):** The Bayesian revolution in genetics. *Nature Reviews Genetics*, 5(4): 251–261.
35. **Beaumont M. and Panchal M., (2008):** On the validity of nested clade phylogeographical analysis. *Mol Ecol.*, 17:2563–2565
36. **Beaumont M. A., Cornuet J.-M., Marin J.-M., and Robert C. P., (2009):** Adaptive approximate Bayesian computation. *Biometrika*, 96 (4): 983–990.
37. **Beaumont M. A., 2010:** Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41: 379–406.
38. **Beckett S. J., Fields P. G. and Subramanyam B. H., (2007):** Disinfestation of stored products and associated structures using heat. In: *Heat Treatments for Postharvest Pest Control* (Tang J., Mitcham E., Wang S. and Lurie S., eds.), CAB International, UK, pp. 182-237.
39. **Bekon K. A., (1984) :** Biologie du développement et comportement alimentaire de *Tribolium castaneum* (Herbst) (Coleoptera : Tenebrionidae) sur les sémences de céréales. Thèse de Doctorat Ingénieur. 167 p.
40. **Bekon K. and Fleurat-Lessard F., (1989):** Evolution des pertes en matière sèche des grains dues à un ravageur secondaire -*Tribolium castaneum* (Herbest), (Coleoptera ; Tenebrionidae), lors de la conservation des céréales, céréales en région chaudes, AUPELF-UREF, Ed.John Libbey Eurotext, Paris, pp. 97-104
41. **Belkhir K., Borsa P., Chikhi L., Raufaste N. and Bonhomme F., (2004) :** GENETIX v. 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
42. **Bell A., Mück O. and Schneider H., (1998) :** La protection intégrée des denrées stockées est une affaire rentable. GTZ, Eschborn, Germany. 42 p.

43. **Bernays E.A. and Chapman R.F., (1994):** Host-plant selection by phytophagous insects. Chapman and Hall, London.
44. **Bhargava A. and Fuentes F. F., (2010):** Mutational dynamics of microsatellites. *Molecular Biotechnology*, (44): 250-266.
45. **Birky C. W. Jr, Fuerst P. and Maruyama T., (1989):** Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, (121): 613–627.
46. **Birky, C. W., (2001):** The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*, (35) : 125-148.
47. **Blein R., Soulé B. G., Dupaigne B. F. and Yérima B. (2008) :** Les potentialités agricoles de l’Afrique de l’Ouest (CEDEAO). Fondation pour l’agriculture et la ruralité dans le monde. 116 p.
48. **Blum M. G. B., (2012) :** Statistique bayésienne et applications en génétique des populations. Statistics. Université de Grenoble.
49. **Boivin G., (1996) :** Évolution et diversité des insectes parasitoïdes. *Antennae*. Numéro spécial: 6-12
50. **Boore J. L., (1999):** Animal mitochondrial genomes. *Nucleic Acids Res*, 27 (8): 1767-1780.
51. **Bossart J. L. and Prowell D. P., (1998):** Genetic estimates of population structure and gene flow: limitation, lessons, and new directions. *Trends Ecol Evol*, (13): 202-206.
52. **Bostein D. R. L, White M, Skolnick R and Davis W., (1980):** Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics* 32: 314-331
53. **Bouchard P., Bousquet Y., Davies A. E., Alonso-Zarazaga M. A., Lawrence J. F., Lyal C. H. C., Newton A. F., Reid C. A. M., Schmitt M., Slipinski S. A. and Smith A. B. T., (2011):** Family-group names in *Coleoptera (Insecta)*. *ZooKeys* 88: 1–972.
54. **Boudreaux H. B., (1969):** The identity of *Sitophilus oryzae*. *Ann. Entomol. Soc. Am.*, 62: 169-172.
55. **Boura A.F., (2006) :** Sensibilité des différentes souches de *Sitophilus zeamais* au maïs blanc CMS 8504 et efficacité des pratiques traditionnelles de stockage. Mémoire de DEA, ENSAI, Université de Ngaoundéré. 61 p.
56. **Bourbonnais G., (2013) :** Coléoptères. Cégep de Sainte-Foy, TBE.
57. **Breton S. and Beaupré H. D., (2007) :** Un système de transmission de l’ADN mitochondriale sexuellement équitable. *Med Sci (Paris)*; 23 (11): 1038–1040.
58. **Breton S., Burger G., Stewart D. T. and Blier P., (2006):** Comparative analysis of gender associated, complete mitochondrial genomes in marine mussels (*Mytilus* spp.). *Genetics*, 172 (2): 1107–1119.
59. **Brochier C., Bapteste E., Lopez Ph. and Phillippe Hervé :** Phylogénies d’espèces à partir d’un grands nombre de gènes. 5p.
60. **Brower A. V. Z., (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 of the United States of America*, 91, (14): 6491-6495.
61. **Bryja J., Granjon L., Dobigny G., Patzenhauerová H., Konečný A., Duplantier J.M., Gauthier P., Colyn M., Durnez L., Lalis A. and Nicolas V. 2010:** Plio-Pleistocene history of West African Sudanian savanna and the phylogeography of the *Praomys daltoni* complex (*Rodentia*): the environment/geography/genetic interplay. *Mol. Ecol.* 19: 4783–4799.
62. **Buchner P., (1965):** Endosymbiosis of Animals with Plant Microorganisms. New York, Interscience.
63. **Burban C., Petit R. J., Carcreff E. and Jactel H., (1999):** Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (*Homoptera: Matsucoccidae*) in relation to the genetic structure of its host. *Molecular ecology* (8): 1593-1602.
64. **Burnham K. P. and Anderson D. R., (2002):** Model selection and multimodel inference: a practical information theoretic approach, 2nd ed., Springer, New York.
65. **Burnham K. P. and Anderson D. R., (2004):** « Multimodel inference: understanding AIC and BIC in Model Selection », *Sociological Methods and Research*, p. 261–304
66. **Cao H., (2007):** Molecular Markers for weevil (*Sitophilus zeamais*) population genetics. Disponible sur « <http://www.biotechniques.org/students/2007/cao/> ». Consulté le 09/07/2016.
67. **Carlton J.T., (1996):** Biological invasions and cryptogenic species, *Ecology*, 77, p.1653-1655.



68. **Carvalho M. O., Pires I., Barbosa A., Barros G., Riudavets J., Garcia A.C., Brites C. and Navarro S., (2012):** The use of modified atmospheres to control *Sitophilus zeamais* and *Sitophilus oryzae* on stored rice in Portugal. *J. Stored Prod. Res.*, (50): 49-56.
69. **Cavalli-Sforza LL (1966):** Population structure and human evolution. *Proc R Soc Lond B Biol Sci* 164, 362-379.
70. **Cavalli-Sforza L. L. and Edwards A. W. F., (1967):** Phylogenetic analysis: models and estimation procedures. *Evolution*, (32): 550-570.
71. **CEDEAO, (2008) :** La politique agricole régionale de l’Afrique de l’Ouest : l’ECOWAP. Commission de la CEDEAO. 12 p.
72. **Chaine S., Legendre Stephane and Jean Clobert (2013):** The co-evolution of multiply-informed dispersal: information transfer across landscapes from neighbors and immigrants. *Peer J.* 1: e44.
73. **Chaisaeng P., Chongrattanameteekul W., Visarathanonth P. and Vajarasathira B., (2010):** Laboratory studies on control of the maize weevil *Sitophilus zeamais* by the parasitoid *Anisopteromalus calandrae*. *Science Asia*, (36): 6–11.
74. **Chakraborty R. and Jin L., (1993):** A unified approach to study hypervariable polymorphism: statistical consideration of determining relatedness and population distance. In *DNA fingerprinting: State of the Science* ed. By Pena S. D. J., Chakraborty R., Epplen and Jeffrey A. J. Birkhauser Verlag Basel/Switzerland, 153-175.
75. **Chao L., (1990):** Fitness of RNA virus decreased by Muller's ratchet. *Nature*, (348): 454-455.
76. **Chapuis M-P. and Estoup A., (2007):** Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, (24): 621-631.
77. **Charles H., (1997) :** Aspect moléculaire de la bactérie symbiotique principale du charançon des céréales *Sitophilus oryzae* (Coléoptère, Curculionidae) et étude de ses interactions avec son hôte. Thèse de Doctorat - INSA de Lyon. 134 p.
78. **Chistiakov D. A., Hellemans B. and Volckaert F. A. M., (2006):** Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture*, (255): 1 –29.
79. **Corander J. and Tang J., (2007):** Bayesian analysis of population structure based on linked molecular information. *Mathematical Biosciences*, (205): 19-31.
80. **Cornelius E. W., Bani R. J., Boateng B. A., Egyir I. S., Josaih M. N., Obeng-Ofori D., Ofori-Anim J. and Timpo S. E., (2008) :** Postharvest Science and Technology, Smartline Ltd. University of Ghana –Legon.
81. **Cornuet J. M. and Luikart G., (1996):** Description and evaluation of two tests for detecting recent bottlenecks. *Genetics*, 144: 2001-2014.
82. **Cornuet J.M. and Luikart G., (1997):** Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144: 2001-2014.
83. **Corrêa A. S., Vinson C. C., Guedes R.N.C. and Oliveira L.O., (2012):** Characterization and transferability of microsatellite markers in three weevil pest species of the genus *Sitophilus* (Coleoptera: Curculionidae). Departamento de Entomologia, Universidade Federal de Viçosa, Viçosa, MG 36570-000, Brazil. *Molecular Ecology Resources*. Permanent genetic resources notes.
84. **Corrêa, A.S., Vinson, C.C., Braga, L.S., Guedes, R.N.C. and de Oliveira, L.O. (2016):** ‘Ancient origin and recent range expansion of the maize weevil *Sitophilus zeamais*, and its genealogical relationship to the rice weevil *S. oryzae*’, *Bulletin of Entomological Research*, 107(1), pp. 9–20.
85. **Couteron P., (1996) :** Les secteurs intermédiaires entre domaines soudanien et sahélien en Afrique occidentale : simples transitions ou réalité à part entière? in Guillaumet J.-L, (1996) *Phytogéographie tropicale : réalités et perspectives*. Paris: ORSTOM, p. 39-50.
86. **Cox, A.J. and Hebert P. D. N., (2001):** Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Molecular Ecology* (10): 371–386.
87. **Coyne J. A. and Orr H. A., (2004):** Speciation. Sinauer Associates; Sunderland, MA.
88. **Crandall K. A., (1996):** Multiple interspeciestransmissions of human and simian T-cell leukaemia/lymphoma virus type I sequences. *Mol. Biol. Evol.*, (13) : 115–131
89. **Cruz J. F., Troude F., Griffon D. and Hubert J. P., (1988):** Conservation des grains en régions chaudes. - 2<sup>e</sup> édition «Techniques rurales en Afrique» - CEEMAT/CIRAD, Paris, 545 p.

90. **Csillery K., Blum M. G.B., Gaggiotti O. E. and François O., (2010):** Approximate Bayesian Computation (ABC) in practice. Article in press. Cell press. 9p.
91. **Cunningham C.W., Zhu H. and Hillis D.M., (1998):** Best-fit maximum-likelihood models for phylogenetic inference: Empirical tests with known phylogenies. *Evolution*, (52): 978-987.
92. **Cutter A. D., (2006):** Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics Society of America*. DOI: 10.1534/genetics.105.048207.
93. **Cyr F., (2012) :** Les impacts de la dispersion historique sur la variabilité génétique à différentes échelles spatiales : connaître l'histoire pour mieux comprendre le présent. Thèse de Doctorat d'Etat es Sciences Université de Montréal (Canada), 141p.
94. **Da Silva A. A., Braga L. S., Guedes R. N. C. and Tavares M. G., (2015):** Cytogenetic analyses using C-banding and DAPI/CMA3 staining of four populations of the maize weevil *Sitophilus zeamais* Motschulsky, 1855 (*Coleoptera, Curculionidae*). *Comparative Cytogenetics* 9: 89–102.
95. **Dal Bello G., Padin S., Lopez L. C. and Fabrizio M., (2001):** Laboratory evaluation of Chemical-biological control of the rice weevil (*Sitophilus oryzae* L.) in stored grains. *J. Stored Prod. Res.*, (37): 77-84.
96. **Daly H.V., Doyen J.T. and Purcell III A.H., (1998):** Introduction to Insect Biology and Diversity, 2nd ed. Oxford University Press, Oxford, New York.
97. **Danho M., Gaspar C. and Haubruge E., (2002):** The impact of grain quantity on the biology of *Sitophilus zeamais* Motschulsky (*Coleoptera: Curculionidae*): oviposition, distribution of eggs, adult emergence, body weight and sex ratio. *Journal of Stored Products Research*, (38): 259–266.
98. **Darriba D., Taboada G.L., Doallo R. and Posada D., (2012):** JModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9 (8), 772.
99. **De Lima C. P. F., (1979):** The conduct of field infestation surveys and the economic use of their results. *Proceedings of 1st International Working Conference on Stored Products Entomology* 47-65.
100. **Decaëns T., Porco D. and Rougerie R., (2013) :** Le barcoding ADN : un outil pour étudier la biodiversité des invertébrés terrestre. Disponible sur <http://www.sfecologie.org/regards/2013/10/15/r50-barcoding-adn-decaens-et-al/>. Consulté le 22/7/2015.
101. **Delobel B. and Grenier A. M., (1993):** Effect of non-cereal food on cereal weevils and tamarind pod weevil (*Coleoptera: Curculionidae*). *J. Stored Prod. Res.*, (29): 7–14.
102. Delobel A. and Tran M : 1993 : Les coléoptères des denrées alimentaires entreposées dans les régions chaudes, Paris, éditions ORSTOM. 425 p.
103. **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 (I): 101-102.
104. **Delsuc F. and Douzery E. J P., (2004) :** Les méthodes probabilistes en phylogénie moléculaire : (2) L'approche bayésienne. *Biosystema*, 22, Avenir et pertinence des méthodes d'analyse en phylogénie moléculaire pp.75-86.
105. **Delvare G. and Aberlenc H. P., (1989) :** Les insectes d'Afrique et d'Amérique tropicale. Clés pour la reconnaissance des familles. Montpellier : CIRAD-GERDAT, 302 p.
106. **Denton M., (2010) :** Evolution: Une théorie en crise : p. 180; éd. Flammarion, 389 p.
107. **Dick C.W., Abdul-Salim K. and Bermingham E., (2003):** Molecular Systematic Analysis Reveals Cryptic Tertiary Diversification of a Widespread Tropical Rain Forest Tree. *The American naturalist* 162.
108. **Dickerson, (2003):** Specialty Corns. Disponible sur [http://aces.nmsu.edu/pubs/\\_h/H232/welcome.html](http://aces.nmsu.edu/pubs/_h/H232/welcome.html). Consulté le 14/12/2015.
109. **Diome T., Thiaw C., Ndong A., Sarr M., Kane M. and Sembène M., (2012):** Haplotype diversity of *Tribolium castaneum* H. (*Coleoptera, Tenebrionidae*) pest of stored millet in Senegal. Vol. 6 (13), pp. 192-199.
110. **Diome T, Thiaw C, Sarr M, Ndong A, Kane M, Cissé N. and Sembène M, (2013):** Genetic diversity of *Tribolium castaneum* (Herbst) population in storage infrastructures and agro ecological zones in Senegal. *International Journal of Biosciences* (9): 248-258.

111. **Diome, (2015)** : Biodémographie et diversité génétique des populations de *Tribolium castaneum* H. (Coleoptera, Tenebrionidae) ravageur des grains de mil (*Pennisetum glaucum* Brown) stockés au Sénégal. Thèse de Doctorat. Université Cheikh Anta Diop de Dakar (Sénégal), 197p.
112. **Dobzhansky T., (1937)**: Genetics and the origin of species. Columbia University Press; New York, NY.
113. **Donnelly P. and Tavaré S., (1986)**: The ages of alleles and a coalescent. *Adv Appl Prob*, (18): 1-19.
114. **Douady C. J., Delsuc F., Boucher Y., Doolittle W. F. and Douzery E. J., (2003)**: Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.*, (20): 248–254.
115. **Douglas A. E., (1989)**: Mycetocyte symbiosis in insects. *Biol Rev Camb Philos Soc.*, 64 (4): 409–434.
116. **Doutt R. L. (1959)**: The biology of parasitic hymenoptera. *Ann. Rev. Entomol.*, (4): 161-182.
117. **Douzery, E.J.P., Snell, E.A., Baptiste, E., Delsuc, F. and Philippe, H., (2004)**: The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. USA* 101:15386-15391.
118. **Drummond A. J., Ho S. Y. W., Phillips M. J. and Rambaut A., (2006)**: Relaxed phylogenetics and dating with confidence. *PLoS Biol.*, 4:e88.
119. **Drummond A. J., Rambaut A. and Suchard M. A., (2015)**: BEAST: Bayesian Evolutionary Analysis by Sampling Trees. V. 1.8.2. Available on: <http://beast.bio.ed.ac.uk/>.
120. **Drummond A. J., Rambaut A. and Xie W., (2011)**: Relaxed Phylogenetics and Dating with Confidence. 18p.
121. **Dufumier M., (2010)** : Agro-écologie et développement durable. ISDA, Montpellier : 28-30 June. 20 p.
122. **Duminil J., (2006)** : Etudes Comparatives de la Structure Génétique des Plantes. Thèse de doctorat. Université Henri Poincaré, Nancy I (France), 215 p.
123. **Dupanloup I., Schneider S. and Excoffier L., (2002)**: A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11 (12): 2571-81.
124. **Durand J. D., Bianco P. G., Laroche J. and Gilles A., (2003)**: Insight into the origin of endemic Mediterranean ichthyofauna : phylogeography of *Chondrostoma* genus (Teleostei, Cyprinidae). *Journal of Heredity*, 94 (4): 315-328.
125. **Earl, D. A. and vonHoldt, Br. M., (2012)**: STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* vol. 4 (2) pp. 359-361 doi: 10.1007/s12686-011-9548-7.
126. **Egal S., Hounsa A., Gong Y. Y., Turner P. C., Wild C. P., Hall A. J., Hell K. and Cardwell, K. F., (2005)** : Dietary exposure to aflatoxin from maize and groundnut in young children from Benin and Togo, West Africa. *International Journal of Food Microbiology* (104): 215-224.
127. **Ehrlich P.R. and Raven P.H., (1964)**: Butterflies and plants: a study in coevolution. *Evolution* 18: 586-608.
128. **El Mousadik A. and Petit R. J., (1996)**: High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* L. Skeels) endemic to Morocco. *Theoret. Appl. Genet.*, (92): 832-836.
129. **Ellegren H., (2000a)**: Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* 16, 551–558.
130. **Ellegren H., (2000b)**: Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genetics*, 24 (4): 400-402.
131. **Ellegren H., (2004)**: Microsatellites: simple sequences with complex evolution. *Nature Reviews. Genetics* 5, 435-445.
132. **CACID, (2012)**: L'état du commerce en Afrique de l'Ouest. Rapport annuel, Dakar (Sénégal), 117 p.
133. **Epule T.E., Ford J. D., Lwasa Sh. and Lepage L., (2017)**: Climate change adaptation in the Sahel *Environmental Science and Policy* 75: 121–137.

134. **Erwin, T. L., (1997):** Biodiversity at its utmost: Tropical forest beetles. In: Biodiversity II: Understanding and Protecting Our Biological Resources, Reaka-Kudla M. L., W.D.E., and Wilson E. O., Editor. 1997, Joseph Henry Press: Washington. p. 27-40.
135. **ESASU, (2008) :** Sénégal: enquête de sécurité alimentaire en situation d'urgence en milieu rural. 15 pages.
136. **Esposti D. M., De Vries S., Crimi M., Ghelli A., Patarnello T. and Meyer A., (1993) :** Mitochondrial cytochrome b: evolution and structure of the protein. *Biochim Biophys Acta* 1143:243–271
137. **Estoup A, Jarne P. and Cornuet J-M (2002):** Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* (11): 1591-1604.
138. **Estoup A., Tailliez C., Cornuet J. M. and Solignac M. (1995):** Size homoplasmy and mutational processes of interrupted microsatellites in two bee species. *Apis mellifera* and *Bombus terrestris* (*Apidae*). *Mol. Biol. Evol.* 13: 1074- 1084.
139. **Evanno G., Regnaut S. and Goudet J., (2005):** Detecting the number of *clusters* of individuals using the software Structure: a simulation study. *Mol. Ecol.*, (14): 2611-2620.
140. **Excoffier L. and Heckel G., (2006):** Computer programs for population genetics data analysis: a survival guide. *Nat Rev Genet* (7): 745–758.
141. **Excoffier L. and Lischer H. E., (2010): ARLEQUIN suite ver. 3.5:** A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.
142. **Excoffier L., Smouse, P.E. and 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.
143. **FAO, (1993) :** Le maïs dans la nutrition humaine, N° T0395, version 25.
144. **FAO, (1996):** Rome Declaration on World Food Security and World Food Summit Plan of Action. World Food Summit 13-17 November, Rome (Italy).
145. **FAO, (2002):** Global Forest Resources Assessment 2000. Disponible sur :
146. <http://www.fao.org/docrep/004/Y1997E/y1997e00.htm#Contents> . Consulté le 21/01/2017.
147. **FAO, (2003) :** Gestion de la fertilité des sols pour la sécurité alimentaire en Afrique subsaharienne, Organisation des Nations Unies pour l'Alimentation et l'Agriculture, Rome (Italie), 54p.
148. **FAO, (2004) :** Sahel : Situation météorologique et état des cultures. Système Mondial d'Information et d'Alerte Rapide Rapport No 2 - 16 juillet 2004.
149. **FAO, (2011a):** Missing Food: the Case of Postharvest Grain Losses in Sub-saharan Africa.
150. **FAO, (2011b) :** Produire plus avec moins. Guide à l'intention des décideurs sur l'intensification durable de l'agriculture paysanne. Rome, 112p.
151. **FAO, FIDA and PAM, (2013) :** L'état de l'insécurité alimentaire dans le monde. Les multiples dimensions de la sécurité alimentaire. Rome. FAO.
152. **FAO, (2014) :** Résidus agricoles et sous-produits agro-industriels en Afrique de l'ouest Etat des lieux et perspectives pour l'élevage. 73 p.
153. **FAO, FIDA and PAM, (2015) :** L'état de l'insécurité alimentaire dans le monde 2015. Objectifs internationaux 2015 de réduction de la faim: des progrès inégaux. Rome.
154. **Farrell B. D., (1998):** "Inordinate fondness" explained: Why are there so many beetles? *Science* (281): 555-559.
155. **Farrell B. D., Sequeira A. S., O'Meara B. C., Normark B. B., Chung J. H. and Jordal B. H., (2001):** The evolution of agriculture in beetles (*Curculionidae: Scolytinae* and *Platypodinae*). *Evolution* 55 (10): 2011–2027.
156. **Félix, (2008) :** Ecologie chimique et approche phylogénétique chez trois espèces de Lépidoptères africains du genre *Busseola* (*Noctuidae*). Thèse de Doctorat. Université Paris XI, Faculté de médecine paris-sud. 202p.
157. **Felsenstein J., (1981):** Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, (17): 368-376.
158. **Felsenstein J., (1985):** Confidence of limits on phylogenies: an approach using bootstrap. *Evolution* 39: 783-791.

159. **Felsenstein, J. (2005)**: PHYLIP (phylogeny inference package), version 3.6. Distributed by Author. Department of Genome Sciences, University of Washington, Seattle.
160. **Feng X.-J., Jiang G.-F. and Fan Zh., (2015)**: Identification of outliers in a genomic scan for selection along environmental gradients in the bamboo locust, *Ceracris kiangsu*. *Sci. Rep.* 5, 13758; doi: 10.1038/srep13758
161. **FIDA, (2011)** : Rapport 2011 sur la pauvreté rurale. Nouvelles réalités, nouveaux défis : de nouvelles chances pour la prochaine génération. Rome.
162. **FIDA, (2012)** : La petite agriculture durable : nourrir le monde, protéger la planète. Conseil des gouverneurs.
163. **Fisher R. A., (1930)**: The genetical theory of natural selection Clarenton Press.
164. **Foll M. and Gaggiotti OE (2008)**: A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977-993
165. **Foll M., (2012)**: BayeScan v2. 1 User Manual. *Ecology*, 20, 1450–1462
166. **Fontaine M. C., (2008)** : Structure et évolution des populations de marsouins communs des eaux palearctiques occidentales. Thèse de doctorat. Faculté des Sciences DSGE. Université de Liège. 126 pages.
167. **Foulley J.-L. and Ollivier L., (2006a)**: Estimating allelic richness and its diversity. *Livest. Sci.*, (101): 150-158.
168. **Foulley J. L. and Ollivier L., (2006b)**: Diversité génétique et richesse allélique: concepts et application à des races bovines. Rencontre autour des Recherches sur les Ruminants 13, 227-230.
169. **Fox J. and Weisberg S. (2011)**: An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
170. **Frankham R., (1996)**: Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, 10: 1500–1508.
171. **Frankham R., (2005)**: Genetics and extinction. *Biological Conservation*, 126, 131–140.
172. **Frantz A. C., Cellina S., Krier A., Schley L., and Burke T., (2009)**: Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: *clusters* or isolation by distance? *Journal of Applied Ecology* 46:493–505.
173. **Fu Y.-X., (1997)**: Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147: 915-925.
174. **Futuyma D. J., (2005)**: Evolution Sinauer Associates; Sunderland, MA.
175. **Gallo D., Nakano O., Silveira Neto S., Carvalho R. P. L., Batista G. C., Berti filho E., Parra J. R. P., Zucchi R. A., Alves S. B. and Vendramin J. D., (1988)**: Manual de Entomologia Agrícola. São Paulo: Agrônômica Ceres,
176. **Gallo D., Nakano, O., Silveira Neto S., Carvalho R.P.L., Baptista G. C., Berti Filho E., Parra J.R.P., Zucchi R. A., Alves S.B. and Vendramin J. D., (2002)**: Entomologia Agrícola (Agricultural Entomology). Piracicaba, SP, Fealq, 920 p.
177. **Galtier N. and Romiguier J., (2015)**: Parental investment predicts genetic diversity in animal species. *MéDecine Sciences: M/S.* 31: 17-9. PMID 25658723 DOI: 10.1051/medsci/2015310100
178. **Garant, D., S. E. Ford and A. P. Hendry (2007)**: The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functionnal Ecology* 21: 434-443.
179. **Gelman A., Carlin J. B., Stern H. S. and Rubin D. B., (2004)** : Bayesian Data Analysis, 2nd Edition. Boca Raton, FL, USA: Chapman and Hall/CRC.
180. **Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A. and Rubin D. B., (2014)** : Bayesian Data Analysis. Third Edition. Boca Raton, FL, USA: Chapman and Hall/CRC.
181. **Gernhard T., (2008)**: The conditioned reconstructed process. *Journal of theoretical biology*, 253 (4), 769-778.
182. **Girod C., (2010)** : Conséquences génétiques des variations climatiques du Quaternaire et distribution des espèces forestières Néotropicales : L'exemple du palmier *Astrocaryum sciophilum*. Thèse de doctorat, Université Pierre et Marie Curie, 130p.
183. **Girod C., Vitalis R., Leblois R. and Freville H., (2011)**: Inferring Population Decline and Expansion from Microsatellite Data: a Simulation-Based Evaluation of the MSVAR Method. *Genetics* 188: 165–179.

184. **Gong Y. Y., Cardwell K. F., Hounsa A., Egal S., Turner P. C., Hall A. I. and Wild C. P. (2002):** Dietary aflatoxin exposure and impaired growth in young children from Benin and Togo: Cross sectional study. *British Medical Journal* (325): 20-21.
185. **Gong, Y.Y., Hounsa, A, Egal, S., Turner, P.c., Sutcliffe, A. E., Hall, A. J., Cardwell, K.F. and Wild, C. P., (2004):** Postweaning exposure to aflatoxin results in impaired child growth : A longitudinal study in Benin, West Africa. *Environmental Health Perspectives* (112): 1334-1338.
186. **Grant W. S. and Bowen B. W., (1998):** Shallow Population Histories in Deep Evolutionary Lineages of Marine Fishes: Insights from Sardines and Anchovies and Lessons for Conservation. *The American Genetic Association*, (89): 415-426.
187. **Gray M. W., Burger G. and Lang B. F., (2001):** The origin and early evolution of mitochondria. *Genome Biol.* 2 (6): reviews 1018.
188. **Greenbaum G., Templeton A. R., Zarmi Y. and Bar-David S., (2014):** Allelic richness following population founding events - a stochastic modeling framework incorporating gene flow and genetic drift. *PLoS ONE* 9 (12): e115203.
189. **Grenier A. M., Nardon C. and Nardon P., (1994):** The role of symbiotes in flight activity of *Sitophilus* weevils. *Entomol. Exp. Appl.*, (70) : 201-208.
190. **Grenier A. M., Nardon P. and Bonnot G., (1986) :** Importance de la symbiose dans la croissance des populations de *Sitophilus oryzae* L. (Coléoptère Curculionidae). Etude théorique et expérimentale. *Acta Oecologica Oecol. Applic.*, (7): 93-110.
191. **Grimaldi D and Engel M., (2005):** Evolution of the Insects. Cambridge University Press, Cambridge, UK.
192. **Grivet D., (2002) :** Phylogéographie et évolution moléculaire comparée d'arbres forestiers à l'aide des marqueurs chloroplastiques. Thèse. Université de Nancy I. France.
193. **Gu X., Fu Y. X. and Li W. H., (1995):** Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.*, 12: 546-557
194. **Guèye M. T., Seck D., Wathelet J.-P. and Lognay G., (2011) :** Typologie des systèmes de stockage et de conservation du maïs dans l'est et le sud du Sénégal. *Biotechnol. Agron. Soc. Environ.*, (16) No 1: 49-58.
195. **Guindon S. and Gascuel O., (2003):** A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood, *Syst. Biol.*, (52): 696-704.
196. **Gullan P. J. and Cranston P. S., (2005):** The insects: an Outline of Entomology. Third edition Blackwell, Oxford, UK.
197. **Gum B., Gross R. and Kuehn R., (2005):** Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. *Mol. Ecol.*, 14, 1707-1725.
198. **Haldane, J. B. S., (1932):** The Causes of Evolution. Longmans, Green and Co., London, 60 pp.
199. **Hall T., (2001):** BioEdit version 7.2.5. Department of Microbiology, North Carolina State.
200. **Halstead D. G. H., (1963a):** External sex differences in stored products Coleoptera. *Bulletin of Entomological Research*, 54: 119-134.
201. **Halstead D. G. H., (1963b):** The separation of *Sitophilus oryzae* (L.) and *S. zeamais* Motschulsky (Col., Curculionidae), with a summary of their distribution. *Entomol. Mon. Mag.*, (99): 72-74.
202. **Hardy G. H., (1908):** Mendelian proportions in a mixt population. *Science*, 28, 49-50.
203. **Hardy O. J. and Vekemans X., (2002):** SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2, 618-620.
204. **Hardy O. J., Charbonnel N., Fréville H. and Heuertz M., (2003):** Microsatellites allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* 163, 1467-1482.
205. **Harpending H. C., Sherry S. T., Rogers A. R. and Stoneking M., (1993):** Genetic structure of ancient human populations. *Curr. Anthropol.*, (34) : 483-496.
206. **Harpending, R. C., (1994):** Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66: 591-600
207. **Hartl D. L. and Clark A. G., (1989):** Principles of Population Genetics," 2nd ed., Sinauer, Sunderland, MA.

208. **Hartl D. L. and Clark A. G., (2007):** Principles of Population Genetics, 4th ed. Sinauer Associates Inc., Sunderland, 542 pp.
209. **Hartl D. L., (1994) :** Génétique des populations. Médecine-Sciences, éditions Flammarion. 305p.
210. **Haryadi Y. and Fleurat-Lessard F., (1994):** Factors affecting survival and development of *Sitophilus oryzae* (L.) in rice grain pericarp layers. *in*: Highley, E., Wright, E.J., Banks, H.J., Champ, B.R. (Eds.), Proceedings of the Sixth International Working Conference on Stored Products Protection. CAB International, Wallingford, United Kingdom, pp. 525–527.
211. **Hebert P. D. N, Stoeckle M. Y., Zemlak T. S. and Francis C. M., (2004).** - Identification of birds through DNA barcodes. *Plos Biology* 2, 1657-1663.
212. **Heddi A., Charles H., Khatchadourian C., Bonnot G. and Nardon P., (1998):** Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA. *J. Mol. Evol.*, (47): 52-61.
213. **Heddi, A., Grenier, A.M., Khatchadourian, C., Charles, H. and Nardon, P., (1999):** Four intracellular genomes direct weevil biology: Nuclear, mitochondrial, principal endosymbionts, and *Wolbachia*. *Proc. Natl. Acad. Sci., USA*, (96): 6814-6819.
214. **Heddi A., Charles H. and Khatchadourian C., (2001):** Intracellular bacterial symbiosis in the genus *Sitophilus*: the 'biological individual' concept revisited. *Res. Microbiol.*, (152): 431-437.
215. **Heddi A., Vallier A., Anselme C., Xin H., Rahbe Y. and Wäckers F., (2005):** Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis. *Cell Microbiol.*, 7 (2): 293-305.
216. **Heddi A., (2009) :** L'immunité dit "OUI" aux bactéries symbiotiques, mais les garde sous contrôle ! *Biofutur*, Vol 28, (299): 36-39.
217. **Heinrigs P., (2010):** Security Implications of Climate Change in the Sahel Region: Policy considerations. SWAC/OECD 32 p.
218. **Helbig J., (1998):** Ability of naturally occurring parasitoids to suppress the introduced pest *Prostephanus truncatus* (Horn) (Coleoptera, Bostrichidae) in traditional maize stores in Togo. *J Stored Prod Res* (34): 287–95.
219. **Hell K., Cardwell K. F., Setamou M. and Poehling H. M., (2000):** The influence of storage practices on aflatoxin contamination in maize in four agro-ecological zones of Benin, West Africa. *Journal of Stored Products Research* (36): 365-382.
220. **Hell K., Fandohan P., Ranajit B., Kiewnick S., Sikora R. and Cotty P. J., (2008) :** Pre- and postharvest management of aflatoxin in maize: an African perspective. *in* Leslie J. F., Bandyopadhyay R., Visconti A., *Mycotoxins : detection methods, management, Public health and agricultural trade.* 219-229.
221. **Hell, K., Mutegi, C. and Pandohan, P., (2010):** Aflatoxin control and prevention strategies in maize for Sub-Saharan Africa. *Julius-Kühn-Archiv* (425): 534–541.
222. **Hendry A. P., Wenburg J. K., Bentzen P., Volk E. C. and Quinn T. P., (2000):** Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290, 516–519.
223. **Hewitt G. H., (2000):** The genetic legacy of Quaternary ice ages. *Nature* 405: 907–913.
224. **Hewitt G.H., (2004):** Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans., Ser. B* 359: 183–195.
225. **Hammond D. S., (2005):** Tropical forests of the guiana shield Ancient forest in a modern world. *in*: Hammond DS. Biophysical features of the Guiana Shield. Ed Hammond DS Pp15-194
226. **Hidayat P., Phillips T.W. and Ffrench-Constant R. H., (1996):** Molecular and morphological characters discriminate *Sitophilus oryzae* and *S. zeamais* (Coleoptera: Curculionidae) and confirm reproductive isolation. *Ann. Entomol. Soc. Am.* (89): 645–652.
227. **Hill D. S. and Waller L. M., (1990):** Pest and Disease of Tropical Crops. Volume 2: Field handbook. *Longman Scientific and Technical.*
228. **Hillis D. M. and Bull J. J., (1993):** An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, 42 (2) 182-192.
229. **Ho S. Y. W., and Lo N., (2013):** The insect molecular clock. *Australian Journal of Entomology*, (52): 101–105.

230. **Hoglund, J., Piertney, S. B., Alatalo, R. V., Lindell, J., Lundberg, A. and Rintamaki, P. T., (2002).** Inbreeding depression and male fitness in black grouse. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 269, 711-715.
231. **Holsinger K. E. and Weir, B. S., (2009).** Genetics in geographically structured populations: defining, estimating and interpreting  $F$  *Nature Rev. Genet.*10, 639–650
232. **Honek A., (1993):** Intraspecific variation in body size and fecundity of insects: a general relationship. *O.ikos* 67, 483–492.
233. **Hong K-J., Lee W., Park Y-J. and Yang J-O., (2018):** First confirmation of the distribution of rice weevil, *Sitophilus oryzae*, in South Korea. *Journal of Asia-Pacific Biodiversity*, Vol. 11, Issue 1: 69-75
234. **Hudson R. R., Kreitman M. and Aguade M., (1987):** A test of neutral molecular evolution based on nucleotide data. *Genetics* 116, 153-159.
235. **Hudson R. R. and Slatkin M., (1991):** Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129: 555-562.
236. **Hudson R. R., Slatkin M. and Maddison W. P., (1992b):** Estimation of Levels of Gene Flow From DNA Sequence Data. *Genetics* (132): 583-589.
237. **Huelsenbeck J. P. and Crandall K. A., (1997):** Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics* Vol. (28): 437-466
238. **Huelsenbeck J. P. and Rannala B., (1997):** Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science*. Apr 11; 276 (5310): 227-32.
239. **Huelsenbeck J. P., (2000):** Likelihood-Based Inference of Phylogeny. 56p.
240. **Huelsenbeck J. P., Ronquist F., Nielsen R. and Bollback J. P., (2001):** Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, 294: 2310-2314.
241. **Huelsenbeck, J. P. and Ronquist F., (2001):** Mrbayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, (17): 754-755.
242. **Hurvich C. M. and Tsai C-L., (1989):** Regression and time series model selection in small samples. *Biometrika*. (76): 297–307.
243. **Huvet A., (2000) :** Ressources génétiques et phylogéographie des huîtres creuses *Crassostrea gigas* et *C. angulata* : variabilité, différenciation et adaptation des populations naturelles et introduites. Thèse de doctorat. Université François Rabelais TOURS. U.F.R. Sciences (France). 388 p.
244. **Huxley J., (1942):** The modern synthesis. George Allen and Unwin Ltd; London, UK. *Evolution*.
245. **Iloba and Ekkrakene, (2006):** Comparative assessment of insecticidal effect of *Azadiractha indica*, *Hyptis suaveolens* and *Ocimum gratissimum* on *Sitophilus zeamais* and *Callosobruchus maculatus*. *J. Biol. Sci.*, 6 (3): 626 – 63.
246. **Isman M. B., (2006):** Botanical insecticides, deterrents, and repellents in modern agriculture and an increasingly regulated world. *Annual Review of Entomology*, (51): 45–66.
247. **Isman M. B., (2008):** Perspective botanical insecticides: For richer, for poorer. *Pest Management Science*, (64): 8–11.
248. **Jarne P. and Lagoda P. J. L., (1996):** Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, (11): 424-429.
249. **Jolliffe I. T. & Cadima J., (2016):** Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci.*; 374 (2065) :20150202.
250. **Jombart T., (2008)** adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.
251. **Jombart T. and Ahmed I. (2011):** adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. doi: 10.1093/bioinformatics/btr521
252. **Jombart T., Devillard S. and Balloux F., (2010):** Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.*11 (1): 94.
253. **Jost L., (2008):**  $G_{ST}$  and its relatives do not measure differentiation. *Mol. Ecol.*, 17, 4015–4026.
254. **Joyce P. and Tavare S., (1995):** The distribution of rare alleles. *J. Math. Biol.*, 33: 602-618.
255. **Kalinowski S.T., Taper M.L. and Marshall T.C., (2007):** Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16: 1099-1106



256. **Kartavtsev YP (2011a)**: Sequence divergence at mitochondrial genes in animals: applicability of DNA data in genetics of speciation and molecular phylogenetics. *Mar Genomics* 4: 71-81.
257. **Kartavtsev YP (2011b)**: Divergence at *Cyt-b* and *Co-I* mtDNA genes on different taxonomic levels and genetics of speciation in animals. *Mitochondrial DNA* 22: 55-65.
258. **Keenan K, McGinnity P, Cross TF, Crozier W.W. and Prodöhl PA., (2013)**: diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* 4, 782–788.
259. **Kébé K., (2013)** : Diversité génétique, histoire évolutive et bioécologie de *Callosobruchus maculatus* F. (*Coleoptera, Bruchinae*) ravageur de graines de niébé (*Vigna unguiculata* Walp). Thèse de Doctorat. Université Cheikh Anta Diop de Dakar (Sénégal), 198p.
260. **Kébé K., Alvarez N., Espíndola A., Justy F., Olivieri I. and M. Sembène (2016)**: Insights into the genetic structure of the cowpea pest *Callosobruchus maculatus* in Africa. *J Pest Sci* 89: 449.
261. **Kébé, K., Alvarez, N., Tuda, M., Arnqvist, G., Fox, C. W., Sembène, M., and Espíndola, A. (2017)**: Global phylogeography of the insect pest *Callosobruchus maculatus* (*Coleoptera: Bruchinae*) relates to the history of its main host, *Vigna unguiculata*. *Journal of Biogeography*, 44 (11), 2515-2526.
262. **Kehe M., (1975)** : Expérimentation pour la mise au point d'un protocole d'essai de substances insecticides pour la lutte contre les charançons des grains : *Sitophilus granarius* (L) et *Sitophilus oryzae* (L). Mémoire D. A. A. ENSAM, Montpellier, 55p.
263. **Kenway H. and Carrott J., (2006)**: Insect species associations characterize past occupation sites. *Journal of Archeological Sciences* (33) : 1452–1473.
264. **Kimura M., (1968)**: Evolutionary rate at the molecular level. *Nature*, 217: 624-626.
265. **Kimura M., (1969)**. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux. *Genetics* 61: 893-903.
266. **Kimura M., (1980)**: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* (16) 111-120.
267. **Kimura M., (1983a)**: The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, 384 pp.
268. **Kimura M., (1983b)**: Rare variant alleles in the light of the neutral theory. *Mol. Biol. Evol.*, 1, 84-93
269. **King J. L. and Jukes T. H., (1969)**: Non-Darwinian evolution. *Science*, 164: 788-798.
270. **King H., Conlong D. E., 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.
271. **Kingman, J. F. C., (1982)**: On the genealogy of large populations. *Journal of Applied Probability* (19): 27-43.
272. **Kingman J. F. C., (2000)**: Origins of the Coalescent: 1974-1982. *Genetics* (156), 1461-1463.
273. **Kosakovsky-Pond S.L., Frost S.D. and Muse S.V., (2005)**: HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21, 676–679.
274. **Kosakovsky Pond S.L. and Muse S.V., (2005)** - Site-to-site variation of synonymous substitution rates. *Mol Biol Evol* 22: 2375–2385.
275. **Koskinen, M.T., Nilsson, J., Veselov, A.J., Potutkin, A.G., Ranta, E. and Primmer, C. R., (2002)**: Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus*, *Salmonidae*. *Heredity* 88, 391–401.
276. **Kottek M., Grieser J., Beck C., Rudolf B. and Rubel F., (2006)**: World map of the Köppen–Geiger climate classification updated *Meteorol. Z.*, 15 pp. 259-263.
277. **Knowles LL. and Maddison WP., (2002)**: Statistical phylogeography. *Mol. Ecol*; 11: 2623–2635.
278. **Kudom A. A., Mensah B. A. and Botchey M. A., (2011)**: Aqueous neem extract versus neem powder on *Culex quinquefasciatus*: Implications for control in anthropogenic habitats. *Journal of Insect Science*, 11 (142).
279. **Kumar S., Stecher G., and Tamura K (2016)**: MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets *Molecular Biology and Evolution* (33): 1870-1874.

280. **Kuschel G., (1961):** On problems of synonymy in the *Sitophilus oryzae* complex (30th contribution, *Col. Curculionidae*). *Annu. Mag. Nat. Hist. Sen* 13, (4): 241-244.
281. **Kuschel G., (1978):** Notes on the identity of *Sitophilus zeamais* Motschulsky based on type material examination. *J. Nat. Hist.* (12): 231.
282. **Lacy R. C., (1987):** Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation biology* (1): 143-158.
283. **Ladhari A., Laarif A., Omezzine F. and Haouala R., (2013):** Effect of the extracts of the spiderflower, *Cleome arabica*, on feeding and survival of larvae of the cotton leafworm, *Spodoptera littoralis*. *Journal of Insect Science*, 13 (61).
284. Lange K., (2002): Mathematical and statistical methods for genetic analysis. *Springer Science & Business Media*.
285. **Langella O., (2002):** Population genetic software, POPULATIONS 1.2.30.
286. **Larsen J., (2012) :** La récolte record de céréales en 2011 ne permet pas de reconstituer les stocks mondiaux. Publication du *Earth Policy Institute*: Indicateur Eco-Economie : Céréales, USA.
287. **Le Conte J., (1950) :** Le maïs hybride aux Etats-Unis d'Amérique. Archives de l'Institut de Recherche Agronomique de l'Indochine, N° 5. 187p.
288. **Léal M. E., (2004):** The African rain forest during the Last Glacial Maximum, an archipelago of forests in a sea of grass, PhD thesis, Wageningen University, Wageningen.
289. **Lee M. S. Y., (2003):** Species concepts and species reality: salvaging a Linnaean rank. *Journal of Evolutionary Biology*, (16): 179-188.
290. **Lefèvre C., (2004) :** Caractérisation et phylogénie des bactéries symbiotiques intracellulaires des charançons de la famille des *Dryophthoridae*. Thèse de doctorat. INSA de Lyon. 124 p.
291. **Lefèvre Th., Raymond M. and Thomas F., (2016):** Biologie évolutive. 2ème édition. 1000 pages.
292. **Levinson H. Z. and Levinson A. R., (1994):** Origin of grain storage and insect species consuming desiccated food. *Anz. Schädlingsk. Pfl. UmwSch.* (67): 47–60.
293. **Lewontin R.C. and Krakauer J., (1973):** Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74, 175-195.
294. **Ley A. C., Dauby G., Köhler J., Wypior C., Röser M. and Hardy O. J., (2014):** Comparative phylogeography of eight herbs and lianas (*Marantaceae*) in central African rainforests. *Front. Genet.* 5: 403.
295. **Li S., (1996):** Phylogenetic tree construction using Markov chain Monte Carlo. Ph. D. Dissertation, Ohio State University.
296. **Li W. H., Wu C. I. and Luo C. C., (1985):** A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular Biology and Evolution* 2, 150-174.
297. **Li Y-C., Korol A. B., Fahima T., Beiles A. and Nevo E., (2002):** Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* (11): 2453-2465.
298. **Likhayo P. W. and Hodges R. J., (2000):** Field monitoring *Sitophilus zeamais* and *Sitophilus oryzae* (*Coleoptera: Curculionidae*) using refuge and flight traps baited with synthetic pheromone and cracked wheat. *J Stored Prod Res* (36): 341-353
299. **Linder H. P., Klerk H. M. de, Born J., Burgess N. D., Fjeldsa J. and Rahbek C. (2012):** The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography* 39: 1189–1205
300. **Longstaff, B. C., (1981):** Biology of the grain pest species of the genus *Sitophilus* (*Coleoptera: Curculionidae*): A critical review. *Protection Ecology* (2): 83-130.
301. **Lopez M. D., Jordán M. J. and Pascual-Villalobos M. J., (2008):** Toxic compounds in essential oils of coriander, caraway and basil active against stored rice pests. *Journal of Stored Products Research.*, (44): 273–278.
302. **Lorenzen, E. D., Heller, R., and Siegismund, H. R. (2012):** Comparative phylogeography of African savannah ungulates. *Molecular Ecology*, 21, 3656-3670.
303. **Lu W. and Hanrahan M., (2014):** Phylogenetic Analysis Using Bayesian Model. Proceedings of 2014 Zone 1 Conference of the American Society for Engineering Education (ASEE Zone 1).

304. **Lucas E. and Riudavets J., (2002):** Biological and mechanical control of *Sitophilus oryzae* (Coleoptera: Curculionidae) in rice. *J. Stored Prod. Res.* (38): 293–304.
305. **Luck G. W., Daily G. C. and Ehrlich P.R., (2003):** Population diversity and ecosystem services. *Trends in Ecology and Evolution*, 18, 331-336.
306. **Luikart, G. and J. M. Cornuet (1998):** Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*. 12: 228-237.
307. **Luikart G. L., Allendorf F. W., Cornuet J-F. and Sherwin W. B., (1998):** Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, 89 (3): 238-247.
308. **Lunt D. H., Zhang D. X., Szimura J. M. and Hewitt G. M., (1996):** The insect cytochrome oxidase unit I gene : evolutionnary pattern and conserved primers for phylogenetics studies. *Insect Molecular Biology*. 5 (0), 153-165.
309. **Lv C., Zhong B., Zhong G., Weng Q., Chen S., Hu M., Sun X. and Qin W., (2012):** Four botanical extracts are toxic to the hispine beetle, *Brontispa longissima*, in laboratory and semi-field trials. *Journal of Insect Science*, 12 (58).
310. **Lynch M. and Crease T. J., (1990):** The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 7: 377-394.
311. **Lynch M., (2006):** Streamlining and simplification of microbial genome architecture. *Annu Rev Microbiol*, (60): 327-349.
312. **Lyons-Weiler J., Hoelzer G. A. and Tausch R. J., (1 998):** Optimal outgroup analysis. *Biological Journal of the Linnean Society* (64): 493-511.
313. **MacArthur R. H. and Wilson E. O., (1967):** The theory of island biogeography. Princeton University Press, Princeton.
314. **Macauley H. and Ramadjita T., (2015) :** Les cultures céréalières : riz, maïs, millet, sorgho et blé.
315. **Magnussen A. and Parsi M. A., (2013):** Aflatoxins, hepatocellular carcinoma and public health. *World Journal of Gastroenterology*:19 (10):1508-1512.
316. **Mahé F., (2009) :** Phylogénie, éléments transposables et evolution de la taille des génomes chez les lupins. *Biomolécules*. Université Rennes, 253 p.
317. **Maley J., (1991):** The African rain forest vegetation and palaeoenvironments during late Quaternary. *Climate Change* 19: 79-98.
318. **Maley J., (1996):** The African rain forest - main characteristics of changes in vegetation and climate from the upper cretaceous to quaternary. *Proceedings of the Royal Society of Edinburgh: Essays on the ecology of the Guineo-Congo rain forest*, Huddersfield, UK, 104B: 31-73.
319. **Maley J., (2001) :** La destruction catastrophique des forêts d'Afrique centrale survenue il y a environ 2500 ans exerce encore une influence majeure sur la répartition actuelle des formations végétales. *Systematics and Geography of Plants* 71: 777-796.
320. **Maltagliati F., Di Giuseppe G., Barbieri M., Castelli A. and Dini F., (2010):** Phylogeography and genetic structure of the edible sea urchin *Paracentrotus lividus* (Echinodermata: Echinoidea) inferred from the mitochondrial *cytochrome b* gene. *Biological Journal of the Linnean Society*, 100: 910–923.
321. **Manel, S., M. K. Schwartz, G. Luikart and P. Taberlet (2003):** Landscape genetics: combining *Landscape Ecology* and population genetics. *Trends in Ecology and Evolution* 18 (4): 189-197.
322. **Mantel N., (1967):** Detection of disease clustering and a generalized regression approach. *Cancer Res.*, (27): 209–220.
323. **Margulis L., (1981):** Symbiosis in cell evolution. Freeman, W. H., San Francisco. p. 206-227.
324. **Maruyama T. and Fuerst P. A., (1984):** Population bottlenecks and non-equilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics*, (108): 745-763.
325. **Maruyama T. and Fuerst P. A., (1985):** Population bottlenecks and non-equilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, (111):675-689.
326. **Marvaldi A. E., Sequeira A. S., O'Brien C. W. and Farrell B. D., (2002):** Molecular and morphological phylogenetics of weevils (Coleoptera, Curculionoidea): Do niche shifts accompany diversification? *Systematic Biology* 51: 761-785.

327. **Marvaldi A.E. and Morrone J. J., (2000):** Phylogenetic systematics of weevils (*Coleoptera: Curculionoidea*): A reappraisal based on larval and adult morphology. *Insect Systematics and Evolution* 31 (1): 43-58.
328. **Mason L. J., (2003):** Grain Insect Fact Sheet, E-237-W: Rice, Granary, and Maize Weevils *Sitophilus oryzae* (L.), *S. granarius* (L.), and *S. zeamais* (Motsch). Purdue University, Department of Entomology.
329. **Mau B., (1996):** Bayesian phylogenetic inference via Markov chain Monte Carlo methods. Ph. D. Dissertation, University of Wisconsin.
330. **Mayr E., (1942):** Systematics and the origin of species. Columbia Univ. Press, New York.
331. **Mbaye, (2014) :** Gestion des risques liés à l'utilisation des pesticides de synthèse dans la conservation des produits halieutiques transformés. Thèse de Doctorat d'Etat es Sciences Université Cheikh Anta Diop de Dakar (Sénégal), 228p.
332. **McDonald J. H. and Kreitman M., (1991):** Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* (351): 652-654.
333. **Mestre J. and Chiffaud J., (2005):** Milieu: Végétation. Disponible sur : <http://acrida.info/div/Veg.htm>. Consulté le 23/01/2017.
334. **Metcalf, C.L., and Flint, W.P., (1962):** Destructive and useful insects: their habit and control. New York: London: Mc Graw Hill, 1087 p.
335. **Michalakis Y. and Excoffier L., (1996):** A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, 142: 1061-1064.
336. **Mittermeier R. A., Myers N., Gill P. C. and Mittermeier C. G., (2000):** Hotspots : Earth's Richest and Most Endangered Terrestrial Ecoregions. Mexico City: CEMEX. Myers N. 1988. Threatened biotas: "Hot spots" in tropical forests. *The Environmentalist* (8): 1-20
337. **Montaigne W., (2011) :** Diversité génétique et adaptation au milieu chez les arbres forestiers tropicaux : Etude chez le genre *Virola* (Myristicaceae). Thèse de doctorat. Université des Antilles et de la Guyane. 177 p.
338. **Montesinos J. L. V., (1998) :** Evolución de la fitofagia en los insectos, Boletín de la Real Sociedad Española de Historia Natural (Actas), (95) : 23-30.
339. **Montllor C., Maxmen A. and Purcell A., (2002):** Facultative bacterial endosymbionts benefit pea aphid *Acyrtosiphon pisum* under heat stress. *Ecol Entomol.*, (27): 180-195.
340. **Moraes M. M., Milléo J., Artoni R. F. and Almeida M. C., (2003):** Análise citogenética de duas espécies do gênero *Sitophilus* (Curculionidae): Cariótipo e meiose. Proceedings of the 49° Congresso Brasileiro de Genética, Águas de Lindóia, september 16-19, 2003. Águas de Lindóia, 174.
341. **Moran N. A., (1996):** Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci USA*, (93) : 2873-2878.
342. **Murrell B., Moola S, Mabona A., Weighill T., Sheward D., Kosakovsky Pond S.L. and Scheffler K., (2013):** FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol.* 30 (5): 1196-205
343. **Myers N., (1988):** Threatened biotas: "Hot spots" in tropical forests. *The Environmentalist* (8): 1-20
344. **Myers N., (1990):** The biodiversity challenge: Expanded hot spots analysis. *The Environmentalist* (10): 243-256.
345. **Myers N., Mittermeier R. A., Mittermeier C. G., Da Fonseca G. A. and Kent J., (2000):** Biodiversity hotspots for conservation priorities. *Nature* 403: 853-858.190.
346. **Nagylaki T., (1998):** Fixation indices in subdivided populations. *Genetics*, (148): 1325-1332.
347. **Nardon, P., (1963):** Les possibilités d'emploi des radiations dans la lutte contre les insectes. *Phytoma*, (144) : 7-12.
348. **Nardon, P., (1973) :** Obtention d'une souche asymbiotique chez le charançon *Sitophilus sasakii* Tak : différentes méthodes d'obtention et comparaison avec la souche symbiotique d'origine. *C. R. Acad. Sci. Paris*, (277D): 981-984.
349. **Nardon P. and Charles H., (2001):** Morphological aspects of symbiosis. In: Cellular origin and life in extreme habitats. "Symbiosis, mecanismes and model systems", Seckbach, J., Editor. Kluwer Academic Publishers: Dordrecht, Netherlands. p. 13-44.

350. **Nardon C. and Nardon P., (2002):** New characters to distinguish larvae and adults of the two sibling species: *Sitophilus oryzae* (L.) and *S. zeamais* Mots. (*Coleoptera, Dryophthoridae*). *Annales de la Societe*.
351. **Narum S. R. and Hess J. E. (2011):** Comparison of  $F_{ST}$  outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11: 184–194. doi:10.1111/j.1755-0998.2011.02987.x
352. **Ndiaye A., Fofana A., Ndiaye M., Mbaye D. F., Sène M., Mbaye I. and Chantereau J., (2005) :** Les céréales. *in: Bilan de la recherche agricole et agroalimentaire au Sénégal*. Dakar: ISRA, ITA, CIRAD, 241-256.
353. **Ndiaye, (2012) :** Impact des zones agro-écologiques et agro-climatiques sénégalaises et Guinéennes sur la Diversité et la structuration génétiques de *Sitophilus* spp (coléoptère, Curculionidae) ravageurs du maïs stocké. Mémoire de fin d'étude – Université Cheikh Anta Diop de Dakar. 42 pages.
354. **Ndiaye A., (2014) :** Caractérisation génétique, évaluation des flux de gènes pour différents écotypes de la bruche *Caryedon serratus*, ravageur de stocks d'arachide. Thèse de Doctorat. Université Cheikh Anta Diop de Dakar (Sénégal), 151p.
355. **Ndiaye M. R., Ndong A., Thaiw C., Diome T. and Sembène M., (2014) :** Haplotypes of Beetles *Sitophilus zeamais* and *Sitophilus oryzae*, Storage Pests of Maize in Senegal and Republic of Guinea. University Cheikh Anta DIOP of Dakar. *International Journal of Science and Advanced Technology* (ISSN 2221-8386) Vol. 4 (6): 17-30.
356. **Ndiaye M. R. and Sembène M., (2018):** Genetic Structure and Phylogeographic Evolution of West African Populations of *Sitophilus zeamais* (*Coleoptera, Curculionidae*). *Journal Stored Products Research*, Volume 77, Pages 135-143. Elsevier.
357. **Ndong A., THIAW C., Diome T., Diallo B., Kane M., Sarr M. and Sembène M., (2015). –** Barcoding: Comparison of Variation Degree of COI and Cytochrome b Mitochondrial Markers in Two Species. *Primary Maize Pests (Sitophilus zeamais and Sitophilus oryzae)*.
358. **Ndong A., (2015) :** Caractérisation génétique de *Sitophilus* spp., ravageur des stocks de maïs et essais d'éradication du virus de la mosaïque du manioc *Manihot esculenta* par la culture de tissus *in vitro*. Thèse de Doctorat. Université Cheikh Anta Diop de Dakar (Sénégal), 314p
359. **Needleman S. B. and Wunsch C. D., (1970):** A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48 (3): 443-453.
360. **Nei M (1972):** Genetic distance between populations. *The American Naturalist*, 106 (949), 283-292.
361. **Nei M., (1973):** Analysis of gene diversity in subdivided populations. *PNAS* 70, 3321-3323.
362. **Nei M., Maruyama T. and Chakraborty R., (1975).** The bottleneck effect and genetic variability in populations. *Evolution*, 29: 1-10.
363. **Nei M., (1977):** F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* Lond. 41: 225-233.
364. **Nei M., (1978).** Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89: 583-590.
365. **Nei M. and 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.
366. **Nei M and Chesser R. K., (1983):** Estimation of fixation indices and gene diversities. *Annals of Human Genetics*, 47, 253–259.
367. **Nei M, Tajima F. and Tatenno Y., (1983):** Accuracy of estimated phylogenetic trees from DNA analysis. *Journal of Molecular evolution*, 19, 153-170.
368. **Nei M. and Gojobori T., (1986):** Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* (3): 418-426.
369. **Nei M., (1987).** *Molecular evolutionary genetics*, Columbia University Press, New York. NY, USA.
370. **Newman S. M., Zelenaya-Troitskaya O, Perlman P. S. and Butow R. A., (1996):** Analysis of mitochondrial DNA nucleoids in wild-type and a mutant strain of *Saccharomyces cerevisiae* that lacks the mitochondrial HMG box protein Abf2p. *Nucleic Acids Res* 24 (2): 386-393.
371. **Ngamo L. S. T., Hance T., (2007):** Diversité des ravageurs des denrées et méthodes alternatives de lutte en milieu tropical. *Tropicicultura*. 25 (4): 215-220

372. **Nicolas V., Granjon L., Duplantier J.-M., Cruaud C. and Dobigny G. (2009):** Phylogeography of spiny mice (genus *Acomys*, Rodentia: Muridae) from the south-western margin of the Sahara with taxonomic.
373. **Noosidum A. and Sangprajan S., (2014):** Insecticidal efficacy of diatomaceous earth against *Sitophilus zeamais* Motschulsky (Coleoptera: Curculionidae) on stored maize in Thailand. *11<sup>th</sup> International Working Conference on Stored Product Protection*, pp. 820-827.
374. **Nordborg M., (2004):** Coalescent Theory. In: *Handbook of Statistical Genetics*. John Wiley and Sons, Ltd.
375. **Nukenine E. N., Monglo B., Awason I., Ngamo L.S.T., Thuenguem F. F. N., and Ngassoum M.B., (2002):** - Farmer's perception on some aspects of maize production and infestation levels of stored maize by *Sitophilus zeamais* in the Ngaoundéré region of Cameroon. *J. Biol. Biochem Sci.* 12 (1): 18-30.
376. **Obata H., Manabe A., Nakamura N., Onishi T. and Senba Y., (2011):** A New Light on the Evolution and Propagation of Prehistoric Grain Pests: The World's Oldest Maize Weevils Found in Jomon Potteries, Japan. *PLoS ONE* 6 (3): 14785.
377. **OCDE, (1998):** Pour Préparer l'avenir de l'Afrique de l'Ouest : Une vision à l'horizon 2020. Etude des perspectives à long terme en Afrique de l'Ouest. 160 p.
378. **OCDE, (2006) :** La zone écologique fragile des pays du sahel. 12 p.
379. **OCDE, (2007) :** L'Afrique de l'Ouest : Une région en mouvement Une région en mutation. Une région en voie d'intégration. 66 p.
380. **OCDE and CEDEAO, (2008):** Élevage et marché régional au Sahel et en Afrique de l'Ouest. Potentialités et défis. Etude réalisée dans le cadre du partenariat entre la Commission de la CEDEAO et le Secrétariat du CSAO/OCDE sur l'avenir de l'élevage au Sahel et en Afrique de l'Ouest.
381. **OCDE, (2013) :** Note de synthèse de l'étude « Peuplement, marché et sécurité alimentaire ». Les Cahiers de l'Afrique de l'Ouest de l'OCDE/CSAO. 7 p.
382. **Ojo J. A. and Omoloye A. A., (2012):** Rearing the maize weevil, *Sitophilus zeamais*, on an artificial maize-cassava diet. *J. Insect Sci.* (12): 1-9.
383. **Oliver K. M., Russell J. A., Moran N. A. and Hunter M. S., (2003):** Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc Natl Acad Sci USA*, (100): 1803-1807.
384. **Olson D.M., Dinerstein E., Wikramanayake E.D., Burgess N.D., Powell G.V.N., Underwood E.C., D'Amico J.A., Itoua I., Strand H.E., Morrison J.C., Louks C.J., Allnutt T.F., Ricketts T.H., Kura Y., Lamoreux J.F., Wettengel W.W., Hedao P. and Kassem K.R., (2001):** Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience*, 51 pp. 933-938.
385. **OMS (2001) :** Changements climatiques et maladies à transmission vectorielles, une analyse régionale. Bulletin de l'Organisation mondiale de la santé, Recueil d'articles n°4, 2001 [http://whqlibdoc.who.int/recueil\\_articles/2001/RA\\_2001\\_4\\_62-72\\_fre.pdf](http://whqlibdoc.who.int/recueil_articles/2001/RA_2001_4_62-72_fre.pdf)
386. **Orémus M., Garrigue C. and Cleguer C., (2011) :** Isolement et diversité génétique des dugongs de Nouvelle-Calédonie. Rapport non publié pour la Direction de l'Environnement de la Province Sud, Nouvelle-Calédonie: 40p
387. **Orti G., Pearse D. E. and Avise J., (1997):** Phylogenetic assessment of length variation at a microsatellite locus. *Proceedings of the National Academy of Sciences of the United States of America*, 94 (20): 10745-10749.
388. **Paes J. L., Faroni L. R. D.'A., Dhingra O. D., Cecon P. R. and Silva T. A., (2012):** Insecticidal fumigant action of mustard essential oil against *Sitophilus zeamais* in maize grains. *Crop Protection*, (34): 56–58.
389. **Paetkau D., Calvert W., Stirling I., and Strobeck C., (1995):** Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol* (4): 347–354.
390. **Paetkau, D., R. Slade, M. Burden and A. Estoup (2004):** Genetic assignment methods for the direct, real time estimation of migration rate: a simulation based exploration of accuracy and power. *Molecular Ecology* 13 (1): 55-65.
391. **Page R., D. M. and Holmes E. C., (1998):** Molecular Evolution: A Phylogenetic Approach. Abingdon, UK: Blackwell Science.

392. **Paliwal R. L., Granados G., Lafitte H. R., Violic A. D. and Marathee J. P., (2000):** Tropical Maize: Improvement and Production. In: FAO Plant Production and Protection Series (FAO), vol. 28, p. 374.
393. **Panchal M and Beaumont MA., (2007):** The automation and evaluation of nested clade phylogeographic analysis. *Evol*; 61: 1466–1480.
394. **Papachristos D. P. and Stamopoulos D. C., (2002):** Repellent, toxic and reproduction inhibitory effects of essential oils vapours on *Acanthoscelides obtectus* (Say) (Coleoptera: Bruchidae). *Journal of Stored Products Research*, (38): 117-128.
395. **Peakall R. and Smouse P. E., (2012):** GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, (28): 2537-2539.
396. **Peck D.R, Congdon D.C., (2004):** Reconciling historical processes and population structure in the sooty tern *Sterna fuscata*. *J. Avian Biol.* (35): 327–335.
397. **Pedersen B. V., (2002) -** European bumblebees (Hymenoptera: Bombini): Phylogenetic relationships inferred from DNA sequences. *Insect Syst. Evol.*, vol. 33 (pg. 361-386)
398. **Pérez-Mendoza J., Baker J. E., Arthur F. H. and Flinn P. W., (1999):** Effects of Protect-It on Efficacy of *Anisopteromalus calandrae* (Hymenoptera: Pteromalidae) Parasitizing Rice Weevils (Coleoptera: Curculionidae) in Wheat. *Biol. Control. Environ. Entomol.*, 28 (3): 529-534.
399. **Pérez-Brocal V., Gil R., Ramos S., Lamelas A., Postigo M., Michelena J. M., Silva FJ, Moya A. and Latorre A, (2006) :** A small microbial genome: the end of a long symbiotic relationships? *Science*, (314): 312-313.
400. **Pérez-Figueroa A., García-Pereira M. J., Saura M., Rolán-Alvarez E. and Caballero A., (2010):** Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, (23): 2267–2276. doi:10.1111/j.1420-9101.2010.02093.x.
401. **Perna N.T. and Kocher T.D., (1995) -** Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* 41: 353-35.
402. **Peters, A. E., Bavishi, A., Cho, H., and Choudhary, M. (2012).** Evolutionary constraints and expression analysis of gene duplications in *Rhodobacter sphaeroides* 2.4.1. *BMC Research Notes*, 5, 192.
403. **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. and Vendramin G. G., (2003) :** Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300: 1563-1565.
404. **Petit R. J., (2007):** The coup de grâce for the nested clade phylogeographic analysis? *Molecular Ecology*.  
**Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2017) :** `_nlme: Linear and Nonlinear Mixed Effects Models_`. R package version 3.1-131, <URL: <https://CRAN.R-project.org/package=nlme>>.
405. **Piry S., Alapetite A., Cornuet J. M., Paetkau D., Baudouin L. and Estoup A., (2004):** GENECLASS 2, a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95, 536–9.
406. **Plarre R., (2010):** An attempt to reconstruct the natural and cultural history of the granary weevil, *Sitophilus granarius* (Coleoptera: Curculionidae). *European Journal of Entomology*, (107): 1-11.
407. **Pons O, Petit R. J., (1996):** Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144, 1237-1245.
408. **Posada D., Crandall K. A. and Templeton A. R., (2000):** Geodis: A program for the cladistic nested analysis of the geographic distribution of genetic haplotypes. *Molecular Ecology*, (9): 487-488.
409. **Posada D. and Crandall K.A., (2001):** Selecting the best-fit model of nucleotide substitution. *Systematic Biology*, 50 (4): 580-601.
410. **Posada D., Baxevanis A.D., Davison D.B., Page R.D.M., Petsko G.A., Stein L.D. and Stormo G.D., (2003) -** Using Modeltest and PAUP\* to select a model of nucleotide substitution, *Current Protocols in Bioinformatics*, New York: John Wiley and Sons(pg. 6.5.1-6.5.14)

411. **Posada D. and Buckley Th. R., (2004):** Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology*, 53(5), 793-808.
412. **Prates H. T., Santos J. P., Waqui J. M., Fabris J.D., Oliveira A., Forster J. E. and Embrapa, (1998):** Insecticidal activity of monoterpenes against *Rhyzoperhta domonica* (F.) and *Tribolium castaneum* (Herbst). *Journal of Stored Products Research*, (34), IV, 243-249.
413. **Pritchard J. K., Seielstad M. T., Perez-Lezaun A. and Feldman M. W., (1999):** Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution* (16): 1791–1798.
414. **Pritchard J.K., Stephens M. and Donnelly P., (2000):** Inference of population structure using multilocus genotype data. *Genetics*, (155): 945-959.
415. **Queller D. C., Strassmann J. E. and Hughes C. R., (1993):** Microsatellites and kinship. *Tree* (8): 285-288.
416. **Quérouil S., Verheyen E., Dillen M. and 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.
417. **R Core Team, (2015):** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
418. **R Core Team, (2017):** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
419. **Rambaut A., (2000):** Estimating the rate of molecular evolution: Incorporating noncontemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* (16): 395-399.
420. **Rambaut A., (2014):** Figtree, a graphical viewer of phylogenetic trees. V. 1.4.2.
421. **Ramirez-Soriano, A., Ramos-Onsins, S.E., Rozas, J., Calafell, F. and Navarro, A. (2008):** Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics*, 179, 555–567.
422. **Ramos-Onsins S. E. and Rozas J., (2002):** Statistical properties of new neutrality tests against population growth, *Mol. Biol. Evol.* 19, 2092–2100.
423. **Rand, D. M., (2001):** The units of selection on mitochondrial DNA. *Annual Review of Ecology and Systematics*, vol. 32, p. 415-448.
424. **Rannala B. and Yang Z., (1996):** Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J. Mol. Evol.*, (43) : 304-311.
425. **Rannala B. et Mountain J. L., (1997):** Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA* 94: 9197-9201.
426. **Ratnadass and Sauphanor B. (1989) :** Les pertes dues aux insectes sur les stocks paysans de céréales en Côte d’ivoire in *Céréales en régions chaudes*. AUPELF-UREF, Eds John Libhey Eurotext (Paris). 47-56 p.
427. **Rees D. P., (1996):** *Coleoptera in: Integrated Management of Insects in Stored Products* (eds Subramanyam 162 BH, Hagstrum DW), pp. 1- 40. Marcel Dekker, New York.
428. **Rees D. P., (2004):** *Insects of Stored Products*. Manson Publishing, Ltd., UK (ed.).
429. **Reydon T. A. C., (2004):** Why does the species problem still persist? *Bio Essays* 26 (3): 300-305.
430. **Riccioni G, Landi M, Ferrara G, Milano I, Cariani A, Zane L, Sella M, Barbujani G and Tinti F (2010):** Spatio-temporal population structuring and genetic diversity retention in depleted atlantic bluefin tuna of the Mediterranean Sea. *Proc Natl Acad Sci USA* 107:2102-2107.
431. **Rice T. K., Schork N. J. and Rao D. C., (2008):** Methods for handling multiple testing. *Adv Genet*, 60:293–308.
432. **Ricklefs R. E., (1987):** Community diversity: Relative Roles of Local and Regional Processes. *Science* 235, 167-171.
433. **Rio R. V. M., Lefevre C., Heddi A. and Aksoy S., (2003):** Comparative genomics of insect-symbiotic bacteria: influence of host environment on microbial genome composition. *Applied and Environmental Microbiology*, (69): 6825-6832.
434. **Rodriguez F., Oliver J. L., Marin A. and Medina J. R., (1990):** The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*. (142): 485–501.



435. **Rogers A. R. and Harpending H., (1992):** Population growth makes waves in the distribution of pairwise genetic difference. *Molecular Biology and Evolution*, 9: 552-569.
436. **Ronce O. (2007):** How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics* 38: 231–25
437. **Rosegrant M. W., Msangi S., Ringler C., Sulser T. B., Zhu T., and Cline S. A., (2008):** International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): 2014 ASABE – CSBE/SCGAB Annual International Meeting Paper Page 14 Model Description. International Food Policy Research Institute, Washington, D.C. Disponible sur : <http://library.ifpri.info/>. Consulté le 01/08/2016.
438. **Rosenberg N. A. and Nordborg M., (2002):** Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.*, 3: 380–390.
439. **Rossetto M, Kooyman R, Sherwin W and Jones R (2008):** Dipersal limitation, rather than bottlenecks or habitat specificity, can restrict the distribution of rare and endangered rainforest trees. *American Journal of Botany*, **95**, 321-329.
440. **Rousset F., (1996):** Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* (142): 1357-1362.
441. **Rousset F., (1997):** Genetic differentiation and estimation of gene flow from F-Statistics under Isolation by Distance. *Genetics*, 145 (4): 1219-1228.
442. **Rousset F., (2008):** Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, (8): 103-106.
443. **Librado P. and Rozas J., (2009):** DNAsp v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25:1451-1452.
444. **Rozas J., Librado P. and Sanches-Delbarrio J. C., (2010):** DNA sequence polymorphism. Version 5.10.01.
445. **Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W. and Hanski I., (1998):** Inbreeding and extinction in a butterfly metapopulation. *Nature*. 392:491-494
446. **Saitou N. and Nei M., (1987):** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology Evolution*, (4): 406-425.
447. **Salemi M. and Vandamme A.M., (2003).** The Phylogenetic Handbook: a Practical Approach to DNA and Protein Phylogeny. 1st ed. Cambridge University Press, Cambridge, 430 pp.
448. **Santiago E. and Caballero A., (2005):** Variation after a selective sweep in a subdivided population. *Genetics* 169, 475–483.
449. **Saraste, M., (1999):** Oxidative phosphorylation at the fin de siècle. *Science*, Vol. 283, (5407): 1488-1493.
450. **Sarr A. G. R. J., Dia C. A. K. M., Ndiaye M. R. and Sembène M., (2016) :** Genetic Structure of Two *Sitophilus (Coleoptera, Curculionidae)* Species According to Storage Infrastructures and Agro-ecological Areas. University Cheikh Anta DIOP of Dakar. International Journal of Science and Advanced Technology (ISSN 2221-8386) Vol. 6 (4): 1-11.
451. **SAS Institute Inc. (1998):** STATVIEW ver. 5.0. Disponible sur : <http://statview.software.informer.com/5.0/>
452. **Sauvion N., Calatayud P.-A., Thiéry D. and Marion-Poll F., (2013) :** Interactions insectes-plantes ISBN 978-2-7592-2018-2 Coédition Quæ IRD Éditions, 784 pages.
453. **Saxena K.-N., Pala-Okeyo A., Seshu-Reddy K.-V., Omolo E.-O. and Ngode L., (1990):** Insect pest management and socio-economic circumstances of small-scale farmers for food crop production in western Kenya: A case study. *Insect Sci. Applic.* (10): 443-462.
454. **Schlenker W. and Lobell D.B., (2010):** Robust negative impacts of climate change on African agriculture. *Environ. Res. Lett.* 5 014010 doi: 10.1088/1748--9326/5/1/014010.
455. **Schmidt HA, Strimmer K, Vingron M. and von Haeseler A. (2002) -** TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* (18): 502-504.
456. **Schneider, P.M., Bender, K., Mayr, W.R., Parson, W., Hoste, B. and Decorte R., (2004):** STR analysis of artificially degraded DNA results of a collaborative European exercise. *Forensic Sci. Int.* (139): 123–134.

457. **Schuelke (2000)**: An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* (18): 233-234.
458. **Seck, D., (2009)** : Stockage, conservation et transformation des récoltes en zones tropicales. Notes de cours. 31 p.
459. **Selkoe K. A. and Toonen R. J., (2006)**: Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* (9) : 615-629.
460. **Sembène M., (2000)** : Variabilité de l'Espacement Interne Transcrit (ITS) de l'ADN ribosomique et polymorphisme des locus microsatellites chez la bruche *Caryedon serratus* (Olivier, 1790) : différenciation en races hôtes et infestation de l'arachide au Sénégal. Thèse de Doctorat d'Etat es Sciences Université Cheikh Anta Diop de Dakar (Sénégal), 180p.
461. **Sembène M., Kébé K. and Delobel A. (2012a)** : Effet structurant de la plante hôte chez la bruche de l'arachide, *Caryedon serratus* (Olivier, 1790) (Coleoptera : Bruchidae). *Biotechnologie Agronomie Societe Environnement*, 16 (1) : 3-11.
462. **Sembène M., (2012b)** : De la naissance des concepts à la génétique évolutive des populations. Cours de master II de spécialité « Génétique des populations ». Université Cheikh Anta DIOP de Dakar.
463. **Seri-Kouassi, B. Ph., Bekon K. A. and Diomande T., (1987)** : Effet de l'alimentation, de l'éclairage et de la température sur des paramètres biologiques de développement d'*Ephestia kuehniella* (Zell.) (*Lepidoptera*: Pyralidae). *Ann. Univ. Abidjan. Série Sc. Nat Biol.* 22 (B): 261-272.
464. **Sézonlin M., Dupas S., Le Rü B. Le Gall P., Moyal P., Calatayud P.-A., Giffard I., Faure N. and 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.
465. **Sézonlin, M., (2006)** : Phylogéographie et génétique des populations du foreur de tiges de céréales *Busseola fusca* (fuller) (*Lepidoptera*, *noctuidae*) en Afrique subsaharienne, implications pour la lutte biologique contre cet insecte. Thèse de doctorat de l'Université de paris XI – Orsay. 161p.
466. **Sherwin WB, Jobot F, Rush R and Rossetto M (2006)**: Measurement of biological information with applications from genes to landscapes. *Molecular Ecology*, **15**, 2857-2869.
467. **Shi Q-H., Zhao F., Hao J-S. and Yang Q., (2013)**: Complete mitochondrial genome of the Common Evening Brown, *Melanitis leda* Linnaeus (*Lepidoptera*: *Nymphalidae*: *Satyrinae*). Mitochondrial DNA, Informa UK, Ltd.
468. **Shimodaira H., (2002)**: An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.*, (51): 492-508.
469. **Simmons R. B. and Weller S. J., (2001)**: Utility and evolution of cytochrome b in insects. *Molecular Phylogenetics and Evolution*, 20 (2): 196-210.
470. **Simon C., Frati F., Beckenbach A., Crespi B., Liu H. and Flook P., (1994)**: Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annal of the Entomological Society of America*, (87): 651-702.
471. **Singh N., Bedi R., Garg R. and Singh J., (2009)**: Physicochemical, thermal and pasting properties of fractions obtained during three successive reduction milling of different corn types. *Food Chem.* 113, 7177.
472. **Slate, J., Kruuk, L. E. B., Marshall, T. C., Pemberton, J. M. and Clutton-Brock, T. H., (2000)**. - Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus*). *Proceedings of the Royal Society of London Series B Biological Sciences*, 267, 1657-1662.
473. **Slatkin M., (1985)**: Rare alleles as indicators of gene flow. *Evolution* (39): 53–65.
474. **Slatkin M. and Barton N. H., (1989)**: A Comparison of 3 Indirect Methods for Estimating Average Levels of Gene Flow. *Evolution* 43, 1349-1368.
475. **Slatkin M., (1987)**: Gene flow and the geographic structure of natural populations. *Sciences*, (236): 787-792.
476. **Slatkin M., (1991)**: Inbreeding coefficients and coalescence times. *Genetical Research* (58): 167-175.
477. **Slatkin M., (1994)**: Linkage disequilibrium in growing and stable populations. *Genetics*, (137): 331-336.

478. **Slatkin M (1995):** A measure of population subdivision based on microsatellite allele frequencies. *Genetics* (139): 1463-1463.
479. **Smith S. G. and Virkki N., (1978):** Animal cytogenetics. Insecta 3. Coleoptera 5. Gebrüder Borntraeger, Berlin, 366 pp.
480. **Smith L., (1993):** Host-size preference of the parasitoid *Anisopteromalus calandrae* [Hym: Pteromalidae] on *Sitophilus zeamais* (Col.: Curculionidae) larvae with a uniform age distribution. *Entomophaga* 1993, Volume 38, Issue 2, pp 225-233.
481. **Smith J., Weber G., Manyong V. M. and Fakorede M. A. B., (1997):** Fostering sustainable increases in maize productivity in Nigeria. Chapter 8 in *Africa's emerging maize revolution*, edited by D. Byerlee, and C.K. Eicher, Lynne Rienner Publishers, London, UK.
482. **Smith D. and Kenward H., (2011):** Roman grain pests in Britain: Implications for grain supply and agricultural production. *Britannia* (42): 243–262.
483. **Soltis, P. S. and Soltis D. E., (2003):** Applying the bootstrap in phylogeny reconstruction. *Stat. Sci.* (18): 256–267.
484. **Spicer G. S., (1995):** Phylogenetic utility of the mitochondrial cytochrome oxidase gene: molecular evolution of the *Drosophila buzzatii* species complex. *J. Mol. Evol.*, 41: 479-759.
485. **Steel, M.A., Lockhart, P.J. and Penny, D., (1993):** Confidence in evolutionary trees from biological sequence data. *Nature* 364, 440–442.
486. **Stejskal V. and Kucerova Z., (1996):** The effect of grain size on the biology of *Sitophilus granarius* L. (Coleoptera: Curculionidae). I. Oviposition, distribution of eggs and adult emergence. *Journal of Applied Entomology* 120, 143–146.
487. **Sterken R., Kiekens R., Coppens E., Vercauteren I., Zabeau M., Inzé D., Flowers J. and Vuylsteke M., (2009):** "A population genomics study of the Arabidopsis core cell cycle genes shows the signature of natural selection". *Plant Cell*, 21 (10): 2987–98.
488. **Stouthamer R., Breeuwers J. A. J. and Hurst G. D. D., (1999):** *Wolbachia pipiensis*: Microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.*, (53): 71-102.
489. **Strimmer K. and von Haeseler A. (1997):** Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences of the United States of America* 94: 6815-6819.
490. **Strong, D.R., Lawton, J.H. and Southwood T.R.E., (1984):** Insects on plants: community patterns and mechanisms. Harvard University Press, Cambridge, Massachusetts.
491. **Strosnider H., Azziz-Baumgartner E, Banziger M, Bhat RV, Breiman R, Brune MN, DeCock K, Dilley A, Groopman J, Hell K, Henry SH, Jeffers D, Jolly C, Jolly P, Kibata GN, Lewis L, Liu X, Luber G, McCoy L, Mensah P, Miraglia M, Misore A, Njapau H, Ong CN, Onsongo MT, Page SW, Park D, Patel M, Phillips T, Pineiro M, Pronczuk J, Rogers HS, Rubin C, Sabino M, Schaafsma A, Shephard G, Stroka J, Wild C, Williams JT, and Wilson D., (2006) :** Workgroup report: public health strategies for reducing aflatoxin exposure in developing countries. *Environmental Health Perspectives* (114): 1898–1903.
492. **Sundqvist L, Keenan K, Zackrisson M, Prodöhl P. and Kleinhans D., (2016):** Directional genetic differentiation and relative migration. *Ecol. Evol.* 6, 3461–3475. (doi:10.1002/ece3.2096)
493. **Suzuki R. and Shimodaira H., (2014):** pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R package version 1.2-2.
494. **Suzuki Y., Glazko G. V., and Nei M., (2002):** Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proc. Natl. Acad. Sci., USA* 99: 16138–16143
495. **Swofford D.L., Olsen, G.J., Waddell P.J., and Hillis D.M., (1996) -** Phylogenetic inference in: Hillis D.M., C. Moritz and B.K. Mable (Editeurs), *Molecular Systematics*, second edition. Sinauer Associates, Sunderland, Massachusetts, pp. 407-514.
496. **Symonds M. R. E. and Moussalli A., (2011):** A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behav Ecol Sociobiol* (65): 13–21.
497. **Tajima F., (1989a):** Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.

498. **Tajima F., (1989b):** The effect of change in population size on DNA polymorphism. *Genetics*: (123): 597-601.
499. **Takehana Y., Nagai N., Matsuda M. and Sakaizumi M., (2003):** Geographic Variation and Diversity of the Cytochrome *b* Gene in Japanese Wild Populations of Medaka, *Oryzias latipes*. *Zoological science*: 1279–1291.
500. **Tamura, K., and M. Nei, (1993):** Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.*, (10): 512-526.
501. **Tapondjou L.A., Adler C., Bouda H. and Fontem D.A., (2002):** Efficacy of powder and essential oil from *Chenopodium ambrosioides* leaves as post-harvest grain protectants against six-stored product beetles. *Journal of Stored Products Research*, (38): 395-402.
502. **Templeton A. R., Boerwinkle E. and Sing C. F., (1987):** A Cladistic Analysis of Phenotypic Associations with Haplotypes Inferred from Restriction Endonuclease Mapping. I. Basic Theory and an Analysis of Alcohol Dehydrogenase Activity in *Drosophila*. *Genetics*. 117 (2): 343-351. 67.
503. **Templeton A. R., (1993):** The “Eve” hypothesis: A genetic critique and reanalysis. *American Anthropologist*. (95): 51-72.
504. **Templeton A. R. and Sing C. F., (1993):** A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* (134): 659-669.
505. **Templeton A. R., Routman E. and Phillips C. A., (1995):** Separating population structure from population history: cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782.
506. **Templeton A. R., (1998):** Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* 7: 381-397.
507. **Templeton A.R., (2004):** Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, 13, 789–809.
508. **Templeton A.R., (2011):** Inference key for the nested haplotype tree analysis of geographical distances. Disponible sur [http://darwin.uvigo.es/download/geodisKey\\_06Jan11.pdf](http://darwin.uvigo.es/download/geodisKey_06Jan11.pdf).
509. **Thangaraj S. R., Subramaniam Ch., McCulloch G. A., Subbarayalu M. and Walter G. H., (2016):** Development of microsatellite markers and a preliminary assessment of population structuring in the rice weevil, *Sitophilus oryzae* (L.). *Journal of Stored Products Research*, 66: 12-17.
510. **Thompson, J., Higgins, D. and Gibson, T., (1997):** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22: 4673-4690.
511. **Thiruvankadan A.K, Jayakumar V., Kathiravan P. and Saravanan R., (2014):** Genetic architecture and bottleneck analyses of Salem Black goat breed based on microsatellite markers. *Veterinary World*, 7: 733-737
512. **Throne J. E., (1994):** Life history of immature maize weevils (*Coleoptera: Curculionidae*) on corn stored at constant temperatures and relative humidities in the laboratory. *Environmental Entomology*, (23): 1459–1471.
513. **Thuillet, A.-C., D. Bru, J. David, P. Roumet, S. Santoni, P. Sourdille and T. Bataillon (2002).** Direct Estimation of Mutation Rate for 10 Microsatellite Loci in Durum Wheat, *Triticum turgidum* (L.) Thell. ssp durum desf. *Molecular Biology and Evolution* 19: 122-125.
514. **Tipping P. W., Mikalajczak E. L., Rudringuez J. G., Poneleit C. G. and Legg D. E., (1987) :** Effects of whole corn kernels and extracts on the behaviour of maize weevil (*Coleoptera: Curculionidae*). *Entomol. Soc. Am.* (85): 1010-1013.
515. **Trematerra P., (2009):** Preferences of *Sitophilus zeamais* to different types of Italian commercial rice and cereal pasta. *Bulletin of Insectology* (62): 103-106.
516. **Trematerra P., (2013):** Aspects related to decision support tools and Integrated Pest Management in food chains. *Food Control* (34): 733-742.
517. **Tsuchida T, Koga R. and Fukatsu T., (2004):** Host plant specialization governed by facultative symbiont. *Science*; (303): 1989.
518. **Turesson G. 1922a.** The species and the variety as ecological units. *Hereditas* 3: 100–113.

519. **Turesson G. 1922b.** The genotypic response of the plant species to habitat. *Hereditas* **3**: 211–350.
520. **Usher M. B., (2007):** Conserver la diversité biologique européenne dans le contexte du changement climatique. Sauvegarde de la nature, no 149. 33 pages.
521. **Uvarov B. P., (1957):** The aridity factor in the ecology of desert locusts and grasshoppers of the Old World. Arid Zone Research (VIII). Human and Animal Ecology: Reviews of Research, UNESCO, Paris, pp. 164-98 in Popov G. B., (1996) : Quelques effets de la sécheresse sahélienne sur la dynamique des populations acridiennes. *Sécheresse* (7): 91–97.
522. **Vallier A., Vincent-Monégat C., Laurençon A. and Heddi A., (2009):** RNAi in the cereal weevil *Sitophilus* spp: systemic gene knockdown in the bacteriome tissue. *BMC Biotechnology*, 9 (1), 44.
523. **Van Oosterhout C, Hutchinson WF, Wills DPM. and Shipley P., (2004):** Microchecker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes*. (4): 535–538.
524. **Vásquez-castro J., De Baptista G. C., Trevizan L. R. P. and Gadanha Jr C. D., (2009) :** Flight Activity of *Sitophilus oryzae* (L) and *Sitophilus zeamais* Motsch (*Coleoptera: Curculionidae*) and its Relationship with Susceptibility to Insecticides. *Neotropical Entomology* 38(3):405-409.
525. **Venkatesan T., Jalali S. K., Srinivasmurthy K., Rabindra R. J. and Dasan C.B., (2007):** Economics of production of *Goniozus nephantidis* (Muesebeck), an important parasitoid of coconut black-headed caterpillar, *Opisina arenosella* (Walker) for bio-factories. *Biological Control.*, (21) : 53–58.
526. **Villers P., (2014):** Aflatoxins and safe storage. *Frontiers in Microbiology*. 5: 158.
527. **Volobouev VT, Aniskin VM, Lecompte E. and Ducroz J. F., (2002).** Patterns of karyotype evolution in complexes of sibling species within three genera of African murid rodents inferred from the comparison of cytogenetic and molecular data. *Cytogenetic and Genome Research* (96): 261–275.
528. **Von Haeseler A., Sajantila A. and Pääbo S., (1996):** The genetical archaeology of the human genome. *Nat. Genet.* 14, 135-140.
529. **Wagner T., (2002):** Late Cretaceous to early Quaternary organic sedimentation in the eastern Equatorial Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology* (179): 113-147.
530. **Wahlund S (1928):** Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunktd der Vererbungslehre aus betrachtet. *Hereditas*, II, 65-106.
531. **Wang C. Y., Feng Y. and Chen X., (2012):** Complete sequence and gene organization of the mitochondrial genome of *Batocera lineolata* Chevrolat (*Coleoptera: Cerambycidae*). Vol. 57, N° 27: 3578 3585.
532. **Wang D-P, Wan H-L, Zhang S. and Yu J., (2009):** gamma-MYN: a new algorithm for estimating Ka and Ks with consideration of variable substitution rates. *Biology Direct* 4: 20.
533. **Waples R. S., Antao T. and Luikart G., (2014):** Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, 114: e164822.
534. **Wares, J.P. and Cunningham, C.W. (2001a):** Comparative phylogeography and historical ecology of the North Atlantic intertidal. *Evolution*, 55, 2455-2469.
535. **Wares, J.P. and Cunningham, C.W. (2001b):** Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution*, 12, 2455–2469.
536. **Watterson GA and Guess HA, (1977):** Is the most frequent allele the oldest? *Theor Popul Biol.*, (11): 141-160.
537. **Webster M. S., Marra P. P., Haig S. M., Beusch S. and Hormes S.T., (2002):** Links between worlds: unravelling migratory connectivity. *Trends in Ecology and Evolution*, 17: 76-83.
538. **Weir B.S. and Cockerham C. C., (1984):** Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.
539. **Weinberg W., (1908):** Über den Nachweis der Verebung beim Menschen. *Jahresh. Verein f. Vaterl. Naturk in Wüttemberg*, 64: 368-382.
540. **Wen B. and Brower J. H., (1995):** Competition between *Anisopteromalus calandrae* and *Choetospila elegans* (*Hymenoptera: Pteromalidae*) at different parasitoid densities on immature rice weevils (*Coleoptera: Curculionidae*) in wheat. *Biol Contr* (5) 151–7.

541. **Wernegreen J. J., (2002):** Genome evolution in bacterial endosymbionts of insects. *Nature Reviews Genetics*, (3): 850-861.
542. **Wickham H., (2016):** ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
543. **Williams, J. H., (2011):** Aflatoxin as a public health factor in developing countries and its influence on HIV and other diseases. *Human Aflatoxicosis in Developing Countries: A review of toxicology, exposure, potential health consequences and interventions*. Peanut Collaborative Research Support Program, University of Georgia.
544. **Wright S., (1931):** Evolution in mendelian populations. *Genetics* 16, 97-159.
545. **Wright S., (1943):** Isolation by distance. *Genetics* 28, 114-138.
546. **Wright S., (1951):** The genetical structure of populations. *Ann. Eugen. (Lond.)* 15, 323-354.
547. **Wright S., (1969):** Evolution and genetics of populations, Volume 2. The theory of gene frequencies. University of Chicago Press, Chicago, IL.
548. **Wright S., (1978):** Evolution and the Genetics of Populations. Vol. 4. Variability within and among Natural Populations, University of Chicago Press.
549. **Wright S., Keeling J. and Gillman L., (2006):** The road from Santa Rosalia: A faster tempo of evolution in tropical climates.
550. **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): 1356
551. **White F., (1986) :** La végétation de l'Afrique : mémoire accompagnant la carte de végétation de l'Afrique. Paris, ORSTOM, 384 p.
552. **Xia X., Xie Z., Salemi M., Chen L. and Wang Y., (2003):** An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* 26:1-7.
553. **Xia X. and Lemey P., (2009):** Assessing substitution saturation with DAMBE. Pp. 615-630 in Philippe Lemey, Marco Salemi and Anne-Mieke Vandamme, eds. *The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny*. 2nd edition Cambridge University Press.
554. **Xia X., (2013):** DAMBE 5: A comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution* (30): 1720-1728.
555. **Yang Z.-Y., P. Huang and Wu G.-X., (1989):** Observations on the karyotypes of *Sitophilus oryzae* and *Sitophilus zeamais* and their hybrid offsprings. *Acta Entomol. Sin.* (32): 406-409 (in Chinese).
556. **Yang Z., (1996):** "Maximun likelihood models for combined analysed of multiple sequence data". *J. Mol. Evol.*, (42): 587-596.
557. **Yang Z. and Nielsen R., (2000):** Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Molecular Biology and Evolution* 17, 32-43.
558. **Yasutomi K, Umeya K. Tokyo: Zenkoku Nouson Kyouiku Kyoukai, (2000):** An illustrated encyclopedia of house pests.
559. **Yeates DK, Cameron SL and Trautwein M., (2012):** A view from the edge of the forest: recent progress in understanding the relationships of the insect orders. *Australian Journal of Entomology* (51): 79-87.
560. **Yingjie, Z., Quangong H., Lailm Z., Zhanmmg Z. and Yongjia J., (1999) :** Identification of *Sitophilus* three Curculionidae) by protein specificity. *Proceedings of the 7th International Working Conference on Stored-product Protection, 14-19 October 1998, Beijing, China. House of Science and Technology Chengdu, China, - Volume 1.*
561. **Yu J, G Pressoir, WH Briggs, IV BI, M Yamasaki, et al. (2006):** A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genetics* 38: 233-208.
562. **Zaghloul O. A., Magda B. El- Kady, Hossam F. El-Wakil, Salwa M. S. Ahmed and Marwa I. Mackled, (2012):** Biological and genetical studies on the rice weevil, *Sitophilus oryzae* (L.) (Curculionidae: Coleoptera), in Egypt. *Research Journal of Agriculture and Biological Sciences*, 8 (2): 92-97.
563. **Zhang J., Rosenberg H. F. and Nei M., (1998):** Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proceedings of the National Academy of Sciences (USA)* (95): 3708-3713.
564. **Zhang D-X. and Hewitt G. M., (2003). -** Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* (12) 563-584.

565. **Zhao X., Li N., Guo W., Hu X., Liu Z., Gong G., Wang A., Feng J. and Wu C., (2004):** Further evidence for paternal inheritance of mitochondrial DNA in the sheep (*Ovis aries*). *Heredity* 93 (4): 399-403.
566. **Zimmerman E. C., (1993):** Australian Weevils. Volume III, 1993, Melbourne: CSIRO Publication. 854 p. 122.
567. **Zuckerkindl E. and Pauling L. B., (1962):** "Molecular disease, evolution, and genetic heterogeneity". In Kasha, M. and Pullman, B (editors). *Horizons in Biochemistry*. Academic Press, New York. pp. 189–225.

## ANNEXES

## Annexe 1 : Tests de neutralité sélective au sein des populations

Populations	Prob	Stat	Populations	Prob	Stat
Burkina Faso	0,166	0,975	Burkina Faso	1	-0,782
Niger	1	-0,579	Niger	0,210	0,809
Caméroun	1	0	Caméroun	1	-1,071
Côte d'Ivoire	1	-1,004	Côte d'Ivoire	1	0,000
Ghana	1	-2,645	Ghana	1	-1,071
Mali	1	0,000	Mali	1	0,000
Centrafrique	1	0,000	Centrafrique	1	-1,123
Sénégal	1	-1,801	Sénégal	1	0,000
République de Guinée	1	-1,798	République de Guinée	1	-1,071

*CytB**COI*

**Prob** : La probabilité de rejeter l'hypothèse nulle de stricte neutralité ( $d_n = d_S$ ) en faveur de l'hypothèse alternative ( $D_n > d_S$ ), au niveau de 5%. **Stat.** : statistique de test ( $dN - dS$ ) est indiquée dans la colonne  $dS$  et  $dN$  sont les nombres de substitutions synonymes et non synonymes par site, respectivement.

Annexe 2 : Les distances génétiques intra-populations et la matrice des distances génétiques inter-populations basée sur Kimura-2-parameter avec le *Cyt-B* (en gras) and le *COI* (en italique)

Distances intra-populations			Distances inter-populations								
			1	2	3	4	5	6	7	8	9
1. BF	<b>0,014</b>	<i>0,0018</i>		<i>0,0002</i>	<i>0,0000</i>	<i>0,0001</i>	<i>0,0002</i>	<i>0,0010</i>	<i>0,0000</i>	<i>0,0002</i>	<i>0,0000</i>
2. CAM	<b>0,000</b>	<i>0,0035</i>	<b>0,003</b>		<i>0,0002</i>	<i>0,0000</i>	<i>0,0000</i>	<i>0,0024</i>	<i>0,0003</i>	<i>0,0000</i>	<i>0,0001</i>
3. CI	<b>0,001</b>	<i>0,0011</i>	<b>0,003</b>	<b>0,000</b>		<i>0,0000</i>	<i>0,0002</i>	<i>0,0012</i>	<i>0,0000</i>	<i>0,0002</i>	<i>0,0000</i>
4. GH	<b>0,012</b>	<i>0,0004</i>	<b>0,013</b>	<b>0,010</b>	<b>0,010</b>		<i>0,000</i>	<i>0,0020</i>	<i>0,0001</i>	<i>0,0000</i>	<i>0,0000</i>
5. RG	<b>0,004</b>	<i>0,0000</i>	<b>0,004</b>	<b>0,001</b>	<b>0,000</b>	<b>0,011</b>		<i>0,0024</i>	<i>0,0003</i>	<i>0,0000</i>	<i>0,0001</i>
6. ML	<b>0,000</b>	<i>0,0000</i>	<b>0,003</b>	<b>0,000</b>	<b>0,000</b>	<b>0,010</b>	<b>0,001</b>		<i>0,0009</i>	<i>0,0024</i>	<i>0,0013</i>
7. NG	<b>0,018</b>	<i>0,0025</i>	<b>0,001</b>	<b>0,002</b>	<b>0,002</b>	<b>0,013</b>	<b>0,003</b>	<b>0,002</b>		<i>0,0003</i>	<i>0,0000</i>
8. RCA	<b>0,000</b>	<i>0,0000</i>	<b>0,003</b>	<b>0,000</b>	<b>0,000</b>	<b>0,010</b>	<b>0,001</b>	<b>0,000</b>	<b>0,002</b>		<i>0,0001</i>
9. SN	<b>0,012</b>	<i>0,0010</i>	<b>0,005</b>	<b>0,002</b>	<b>0,002</b>	<b>0,012</b>	<b>0,002</b>	<b>0,002</b>	<b>0,004</b>	<b>0,002</b>	
GSO	<b>0,006</b>	<i>0,0009</i>	GSA vs GSO								
GSA	<b>0,012</b>	<i>0,0014</i>	<b>0,001/0,0002</b>								
			Population globale		<b>0,010/0,0012</b>						

**BF** (Burkina Faso) ; 2 : **CAM** (Caméroun) ; 3 : **CI** (Côte d'Ivoire) ; 4 : **RCA** (Centrafrique) ; 5 : **GH** (Ghana) ; 6 : **ML** (Mali) ; 7 : **NG** (Niger) ; 8 : **SN** (Sénégal) ; 9 : **RG** (République de Guinée) ; **GSA vs GGC** : Groupe des Pays du Sahel versus Groupe des pays du domaine guinéo-congolais.

## Annexe 3 : Les distances génétiques intra-populations et la matrice des distances génétiques inter-populations basée sur Kimura-2-parameter avec les Séquences concaténées.

Distances intra-populations		Distances inter-populations								
1. BF	0,0111									
2. CAM	0,0019	0,0018								
3. CI	0,0009	0,0016	0,0001							
4. GH	0,0084	0,0061	0,0049	0,0049						
5. RG	0,0026	0,0021	0,0005	0,0003	0,0053					
6. ML	0,0000	0,0022	0,0012	0,0006	0,0016	0,0058				
7. NG	0,0182	0,0003	0,0013	0,0011	0,0017	0,0059	0,0016			
8. RCA	0,0000	0,0018	0,0000	0,0001	0,0050	0,0049	0,0016	0,0013		
9. SN	0,0062	0,0029	0,0016	0,0016	0,0013	0,0065	0,0023	0,0025	0,0016	
GSO	0,0045	GSA vs GSO								
GSA	0,0097	<b>0,0010</b>								
		Population globale		<b>0,0076</b>						

**BF** (Burkina Faso) ; 2 : **CAM** (Caméroun) ; 3 : **CI** (Côte d'Ivoire) ; 4 : **RCA** (Centrafrique) ; 5 : **GH** (Ghana) ; 6 : **ML** (Mali) ; 7 : **NG** (Niger) ; 8 : **SN** (Sénégal) ; 9 : **RG** (République de Guinée) ; **GSA vs GGC** : Groupe des Pays du Sahel versus Groupe des pays du domaine guinéo-congolais.



**Annexe 4** : Matrice des  $F_{ST}$  par paires de populations (en-dessous de la diagonale) et la matrice des  $Nm$  (au-dessus de la diagonale) par paires de populations avec le *Cyt-B*.

	1	2	3	4	5	6	7	8	9
<b>1. BF</b>		1,3544	0,7935	0,5947	0,5481	1,0231	10,4839	1,2834	1,0920
<b>2. CAM</b>	0,2696*		inf	inf	0,3835	inf	3,0665	3,4317	1,5986
<b>3. CI</b>	0,3866*	0,0000		4,5087	0,2554	14,4444	1,4860	1,9538	1,7632
<b>4. RCA</b>	0,4568*	0,0000	0,0100		0,1934	inf	1,0964	1,5722	0,6798
<b>5. GH</b>	0,4771*	0,5660*	0,6619*	0,7211*		0,3105	0,6302	0,5160	0,3800
<b>6. ML</b>	0,3283*	0,0000	0,0335	0,0000	0,6169*		2,0714	2,5486	1,1875
<b>7. NG</b>	0,0455	0,1402*	0,2518*	0,3132*	0,4424*	0,1944*		1,8080	1,8738
<b>7. SN</b>	0,2804*	0,1272	0,2038*	0,2413*	0,4948*	0,1640*	0,2166*		3,1325
<b>9. RG</b>	0,3141*	0,2383*	0,2209*	0,4238*	0,5682*	0,2963*	0,2106*	0,1377*	

**BF** (Burkina Faso) ; **2 : CAM** (Caméroun) ; **3 : CI** (Côte d'Ivoire) ; **4 : RCA** (Centrafrique) ; **5 : GH** (Ghana) ; **6 : ML** (Mali) ; **7 : NG** (Niger) ; **8 : SN** (Sénégal) ; **9 : RG** (République de Guinée) ; <sup>ns</sup> : non significatif ( $P > 0,05$ ) ; les valeurs accompagnées du symbole \* sont statistiquement significatives.

**Annexe 5** : Matrice des  $F_{ST}$  par paires de populations (en-dessous de la diagonale) et la matrice des  $Nm$  (au-dessus de la diagonale) par paires de populations avec le *COI*.

	1	2	3	4	5	6	7	8	9
<b>1. BF</b>		5,3749	inf	1,0073	4,5109	0,40373	inf	Inf	2,4697
<b>2. CAM</b>	<b>0,0851</b>		3,6458	2,4852	6,5236	0,2767	5,4861	3,1343	24,0000
<b>3. CI</b>	<b>0,0000</b>	<b>0,1206</b>		1,2613	7,1454	0,2662	inf	Inf	2,6667
<b>4. RCA</b>	<b>0,3317*</b>	<b>0,1675</b>	<b>0,2839</b>		13,9286	0,0000	0,9454	1,8750	Inf
<b>5. GH</b>	<b>0,0998</b>	<b>0,0712</b>	<b>0,0654</b>	<b>0,0346</b>		0,0521	3,3355	13,9286	Inf
<b>6. ML</b>	<b>0,5533*</b>	<b>0,6438*</b>	<b>0,6526*</b>	<b>1,0000*</b>	<b>0,9057*</b>		0,6109	0,2469	0,0000
<b>7. NG</b>	<b>0,0000</b>	<b>0,0835</b>	<b>0,0000</b>	<b>0,3459*</b>	<b>0,1304</b>	<b>0,4501*</b>		Inf	2,4167
<b>7. SN</b>	<b>0,0000</b>	<b>0,1376</b>	<b>0,0000</b>	<b>0,2105*</b>	<b>0,0347</b>	<b>0,6694*</b>	<b>0,0000</b>		3,8710
<b>9. RG</b>	<b>0,1684</b>	<b>0,0204</b>	<b>0,1579</b>	<b>0,0000</b>	<b>0,0000</b>	<b>1,0000*</b>	<b>0,1714</b>	<b>0,1144</b>	

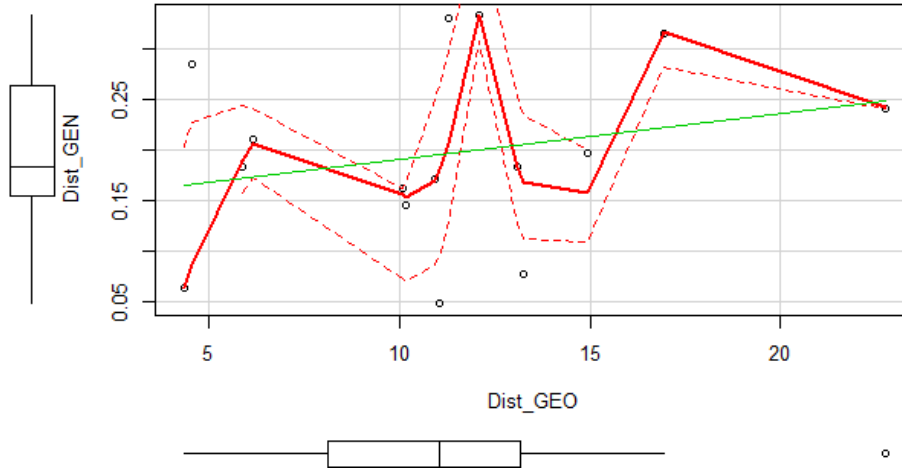
**BF** (Burkina Faso) ; **2 : CAM** (Caméroun) ; **3 : CI** (Côte d'Ivoire) ; **4 : RCA** (Centrafrique) ; **5 : GH** (Ghana) ; **6 : ML** (Mali) ; **7 : NG** (Niger) ; **8 : SN** (Sénégal) ; **9 : RG** (République de Guinée) ; <sup>ns</sup> : non significatif ( $P > 0,05$ ) ; les valeurs accompagnées du symbole \* sont statistiquement significatives.

**Annexe 6** : Matrice des  $F_{ST}$  par paires de populations (en-dessous de la diagonale) et la matrice des  $Nm$  (au-dessus de la diagonale) par paires de populations avec les gènes contiguës.

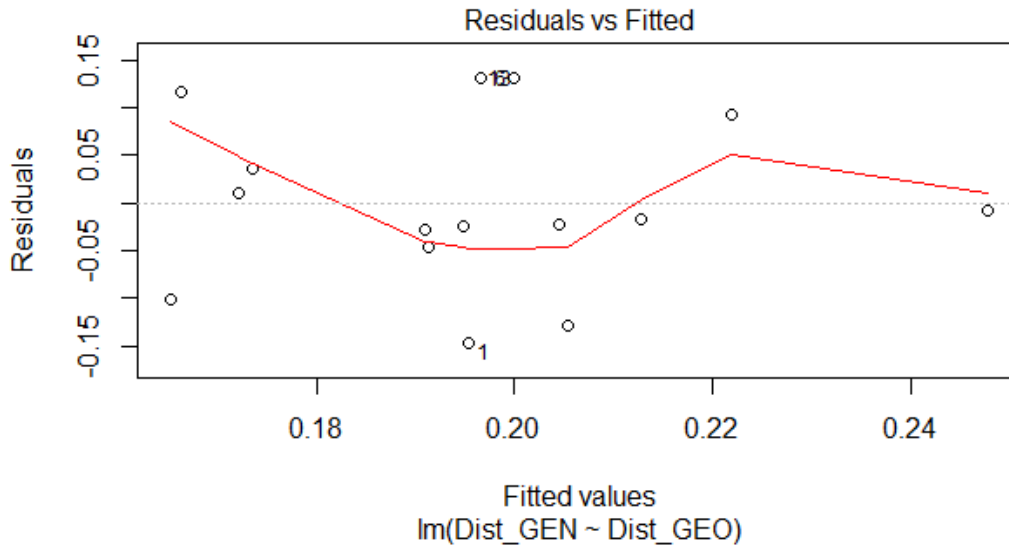
	1	2	3	4	5	6	7	8	9
<b>1. BF</b>		1,6189	1,0099	0,6504	0,5957	0,8688	13,5076	1,0613	1,2974
<b>2. CAM</b>	0,2360*		4,4742	2,8636	0,4663	0,2767	2,8750	1,3856	2,2144
<b>3. CI</b>	0,3312*	0,1005		1,8428	0,3076	0,4198	1,5845	0,9785	1,7520
<b>4. RCA</b>	0,4347*	0,1487	0,2134*		0,2169	0,0000	0,9724	0,7048	0,6175
<b>5. GH</b>	0,4563*	0,5174*	0,6191*	0,6975*		0,2703	0,7146	0,4343	0,4306
<b>6. ML</b>	0,3653*	0,6438*	0,5436*	1,0000*	0,6491*		1,4278	0,7028	0,3028
<b>7. NG</b>	0,0357	0,1482*	0,2399*	0,3396*	0,4117*	0,2594*		1,3613	2,0315
<b>7. SN</b>	0,3202*	0,2652*	0,3382*	0,4150*	0,5352*	0,4157*	0,2686*		1,8155
<b>9. RG</b>	0,2782*	0,1842*	0,2220*	0,4474*	0,5373*	0,6228*	0,1975*	0,2159*	

**BF** (Burkina Faso) ; **2 : CAM** (Caméroun) ; **3 : CI** (Côte d'Ivoire) ; **4 : RCA** (Centrafrique) ; **5 : GH** (Ghana) ; **6 : ML** (Mali) ; **7 : NG** (Niger) ; **8 : SN** (Sénégal) ; **9 : RG** (République de Guinée) ; <sup>ns</sup> : non significatif ( $P > 0,05$ ) ; les valeurs accompagnées du symbole \* sont statistiquement significatives.

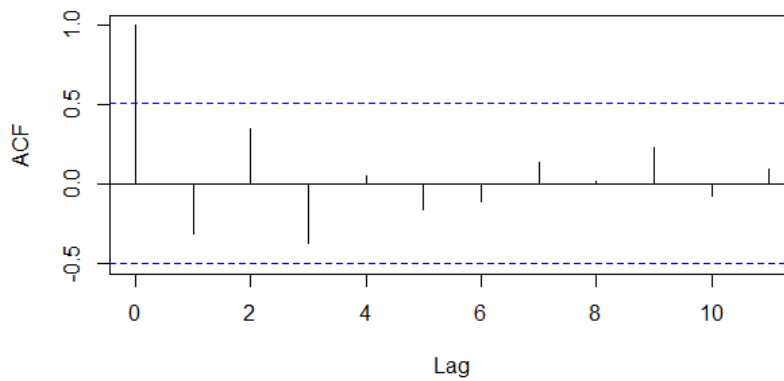
**Annexe 7 :** Evaluation à posteriori de l'hypothèse de linéarité sur les résidus



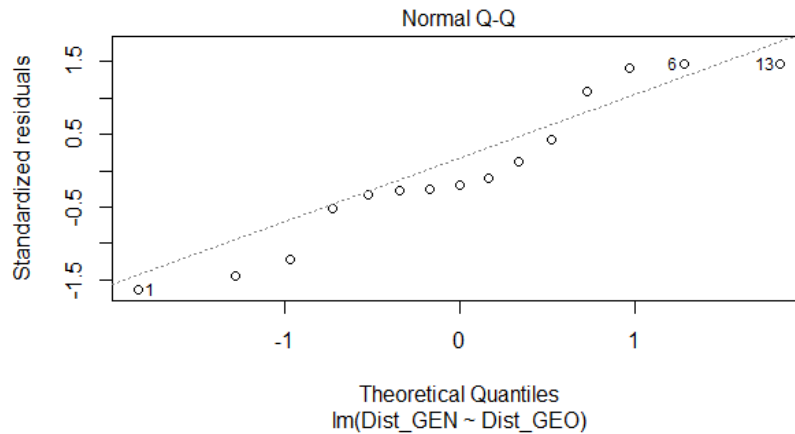
**Annexe 8 :** Evaluation à posteriori de l'hypothèse de linéarité sur les résidus



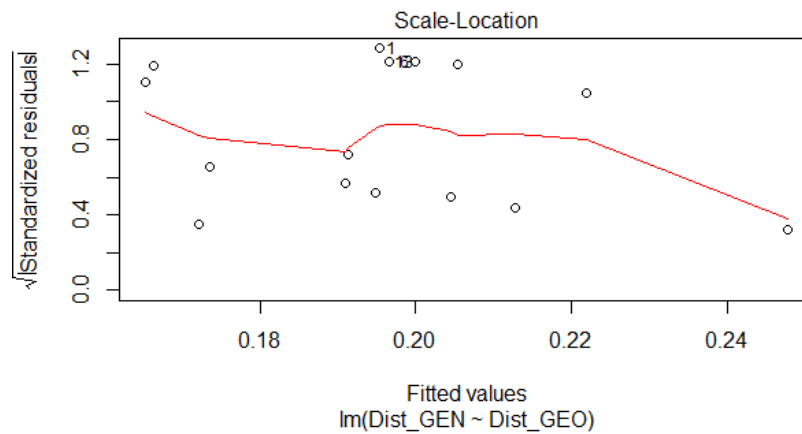
**Annexe 9 :** Evaluation de l'hypothèse d'indépendance des résidus par un « lag plot »



La p-value du test de durbin Watson : P=0,612.

**Annexe 10** : Evaluation de l'hypothèse de normalité des résidus à l'aide d'un QQplot

La p-value du test de normalité des résidus Shapiro-Wilk :  $P=5,93.10^{-12}$ .

**Annexe 11** : Evaluation de l'hypothèse d'homogénéité des résidus

La p-value du test d'homogénéité des résidus de Breush-Pagan:  $P=0,722$ .

**Annexe 12** : PCR en Duplex de *Sitophilus zeamais*

Protocole 'tailing' avec M13-amorce F + M13 marqué à l'IRD-700 ou 800

Marqueur : Microsatellite en Simplex

	[Initiale]	Unité		Qté finale	Unité		Vol. d'ADN par réaction
ADN	5	ng/ $\mu$ l	ADN	10,0	ng	ADN	2,00 $\mu$ l
	[Initiale]	Unité	[Finale]	Unité		Vol. de réactifs pour le mix	
Tp	10	X	Tp	1,0	X	Tp	99 $\mu$ l
MgCl2	50	mM	MgCl2	1,0	mM	MgCl2	20 $\mu$ l
dNTP	2000	$\mu$ M	dNTP	200,0	$\mu$ M	dNTP	99 $\mu$ l
Am 1-M13	10	$\mu$ M	Am F-M13	0,1	$\mu$ M x 2	Am F-M13	8 $\mu$ l
Am 2	10	$\mu$ M	Am R	0,1	$\mu$ M x 2	Am R	8 $\mu$ l
M13 marqué	10	$\mu$ M	M13 marqué	0,1	$\mu$ M	M13 marqué	8 $\mu$ l
Taq	2	U/ $\mu$ l	Taq	0,1	U/ $\mu$ l	Taq	50 $\mu$ l
						H2O	501 $\mu$ l

Nombre de réactions : 90

Volume de réaction : 10  $\mu$ l

Marge pour le mix : 10 %

Volume du Mix : 792,0  $\mu$ l

Volume à distribuer : 8,00  $\mu$ l

## PUBLICATIONS ET COMMUNICATIONS

## 1° PUBLICATIONS

1°\_ **Ndiaye M. R.**, Ndong A., Thiaw C., Diome T. and Sembène M., (2014) : Haplotypes of Beetles *Sitophilus zeamais* and *Sitophilus oryzae*, Storage Pests of Maize in Senegal and Republic of Guinea. *International Journal of Science and Advanced Technology*, 4 (6): 17-30.

*International Journal of Science and Advanced Technology (ISSN 2221-8386)*  
http://www.ijst.com

Volume 4 No 6 June 2014

### Haplotypes of Beetles *Sitophilus zeamais* and *Sitophilus oryzae*, Storage Pests of Maize in Senegal and Republic of Guinea.

**Mama Racky NDIAYE**

BIOPASS Laboratory ; Department of Animal Biology, Faculty of Science and Technology, PO Box 5005 Dakar, Senegal.

Email: [kiiraa12@gmail.com](mailto:kiiraa12@gmail.com)

Tel: +221 77 730 10 10

**Assane NDONG**

BIOPASS Laboratory ; Department of Animal Biology, Faculty of Science and Technology, PO Box 5005 Dakar, Senegal.

**Toffene DIOME**

BIOPASS Laboratory ; Department of Animal Biology, Faculty of Science and Technology, PO Box 5005 Dakar, Senegal.

**Cheikh THIAW**

Regional Centre for Studies on the Improvement of Plant Adaptation to Drought (CERAAS), Khombole Road, PO Box 3320, Thies, Senegal.

**Mbacke SEMBENE**

BIOPASS Laboratory, IRD Bel Aire, P.B. 1386, Dakar. Laboratory of Entomology and Acarology, Department of Animal Biology, Faculty of Science and Technology, University Cheikh Anta Diop of Dakar, Senegal.

Email: [mbacke.sembene@ird.fr](mailto:mbacke.sembene@ird.fr)

#### Abstract

Maize is the first worldwide production and a staple food for many African countries. These countries face problems related to adequacy and adaptation storage systems made vulnerable by disparaging product insects. And losses due to seed-weevils of corn can exceed 25% of the harvest or even 40% in six months of storage. The objective of this study is to determine the influence of agro-ecological and agro-climatic areas on the genetic and structure diversity of *Sitophilus* spp by the PCR-sequencing technique of mitochondrial gene, cytochrome B.

The study highlights the presence of two species in the samples : *Sitophilus zeamais* and

*Sitophilus oryzae* with 18 and 16 haplotypes respectively in seven localities in Senegal and Republic of Guinea. In Senegal, the two species have high haplotype, nucleotide diversity characterizing population stability with a large effective. However in Guinea, *S. zeamais* is characterized by a rapidly growing population from an ancestral population with low numbers and where time is not enough to find a great diversity between haplotypes. Networks and phylogenetic trees of maximum parsimony and maximum likelihood indicate that the two species would be structured according to agro-ecological and agro-climatic areas.

**Keywords :** *Sitophilus zeamais* ; *Sitophilus oryzae* ; Haplotypes ; Agro-ecological areas ; agro-climatic areas ; Senegal ; Republic of Guinea ; PCR-Sequencing ; Cytochrome B.

#### I. Introduction

Native at Central America, specifically in Mexico, maize (*Zea mays*, L.) is introduced in Africa in the sixteenth century by Portuguese

explorers [29 ; 49]. It is widely grown as a cereal for its grain rich in starch (72-73% of its weight) and represents the first cereal production in front of rice and wheat [5 ; 9 ; 28].

2°\_ Bambou A. E., Thiaw C, **Ndiaye M. R.**, Diallo Y., Gueye Momar T. and Sembène M., (2015) – Effectiveness of crushed fresh leaves of *Boscia senegalensis*, Lam Ex. Poir. harvested in Dakar and Richard-Toll on *Sitophilus zeamais* (Motschulsky, 1855). *South Asian J Exp Biol*, 5 (1): 1-5.

South Asian J Exp Biol; 5 (1): 1-5; 2015




**South Asian Journal**  
of Experimental Biology

ISSN: 2230-9799      Vol. 5, Issue 1, Page 1-5      <http://www.sajeb.org>

---

**REGULAR ARTICLE**

**Effectiveness of crushed fresh leaves of *Boscia senegalensis* Lam Ex. Poir. harvested in Dakar and Richard Toll on *Sitophilus zeamais* (Motschulsky)**

**Asta Eliane Bambou<sup>1\*</sup>, Cheikh Thiaw<sup>2</sup>, Mama Racky Ndiaye<sup>3</sup>, Younoussa Diallo<sup>3</sup>, Momar Talla Gueye<sup>3</sup>, Mbacké Sembène<sup>1,4</sup>**

<sup>1</sup>Department of Animal Biology, Cheikh Anta Diop University, Dakar 5005, Senegal  
<sup>2</sup>Senegalese Institute of Agricultural Research/National Center of Agronomics Research (ISRA- CNRA); Box: 33 Senegal  
<sup>3</sup>Dakar Institute of Food Technology Hann-Dakar, BP 2703, Senegal  
<sup>4</sup>CBGP-BIOPASS; IRD Bel Air, B.P.: 1380 Dakar Senegal

ARTICLE INFO	ABSTRACT
<p><i>Article History:</i>            Received: 19 Nov 2014            Revised: 8 Dec 2014            Accepted: 11 Dec 2014</p> <p><i>*Corresponding Author:</i>            Email: elianeasta@hotmail.com            Phone: +221 77 402 78 22</p> <p><i>Keywords:</i> <i>Sitophilus zeamais</i>; <i>Boscia senegalensis</i>; maize; insecticidal activity; locality</p>	<p>The maize is one of the most produced worldwide cereals. Its conservation in stock is attacked by many pests that cause huge loss to the qualitative and quantitative level. Several alternative methods of control resulting in the insecticidal plants replace chemical pesticides. Among these methods, fumigation has proved very effective on insects' food stored. In the present work, fumigating insects of <i>Sitophilus zeamais</i> to different doses of fresh crushed leaves of <i>Boscia senegalensis</i> showed strong insecticidal activity of this plant depending on the locality. Leaves of two localities, Dakar and Richard Toll have been tested. Statistical analysis of the mortality rate expressed in cumulative frequencies showed a significant difference (<math>p &lt; 0.0001</math>) between the groups tested and control groups. Leaves from Richard Toll gave 100% mortality within 24 h to 1 g / l while those collected in Dakar gave 100% in 24 h to a maximum dose of 9 g / l. We can explain this by significant differences in the rate of glucocapparin content in the leaves of this plant according to the locality and period of harvest. It should be noted that crops leaves were not done during the same periods. In this work, the best season to obtain a high efficiency is during dry season.</p>

**1. Introduction**

A third of the food production for consumption in the world is lost or wasted, reaching about 1.3 billion tons per year (Ouédraogo and Théodore Kaboré, 2014). Post harvest losses comprise both weight losses and quality losses. According to FAO, post grain harvest losses are estimated at € 2.7 billion per year (FAO, 2011). Those caused by storage pests represent about 10% on average for an annual monetary value of nearly US\$ 58 billion (Goergen, 2003). Faced with these losses, chemical pesticides remain effective method of control but cause serious problems such as addiction insects and the selection of resistant strains (Benhalima et al., 2004), poisoning, pollution of the environment with ecological disturbances (Regnault-Roger, 2002) and the adverse effects on human health. Alternative control methods against pests such rewarding plants insecticidal have been developed.

1

3°\_ Ndong A., Thiaw C., Diome T., **Ndiaye M. R.**, Kane M. and Sembène M., (2015) Diversity and Genetic Structure of *Sitophilus* Spp. Haplotypes, Primary Pest of Stored Maize in Senegal and Guinea : Genetic Impact of Storage Infrastructure and Agro-Climatic Zones. International Journal of Research Studies in Biosciences (IJRSB), 3 (1): 122-140 pp.

*International Journal of Research Studies in Biosciences (IJRSB)*  
 Volume 3, Issue 1, January 2015, PP 122-140  
 ISSN 2349-0357 (Print) & ISSN 2349-0365 (Online)  
 www.arcjournals.org

### Diversity and Genetic Structure of *Sitophilus* Spp. Haplotypes, Primary Pest of Stored Maize in Senegal and Guinea: Genetic Impact of Storage Infrastructure and Agro-Climatic Zones

#### <sup>1</sup>Assane Ndong

Faculty of sciences and Technology  
 Department of Animal Biology  
 BIOPASS Laboratory UMR 022 IRD-CBGP  
 University Cheikh Anta DIOP  
 Dakar, Senegal  
 assanendong85@yahoo.fr

#### <sup>2</sup>Cheikh Thiaw

Senegalese Institute of Agricultural Research  
 Regional Center for the Study on  
 Improvement of Adaptation to Drought  
 (ISRA/CERAAS), Thies Escale, Senegal.  
 thiacheikh@hotmail.com

#### <sup>3</sup>Toffene Diome

Faculty of sciences and Technology  
 Department of Animal Biology  
 BIOPASS Laboratory UMR 022 IRD-CBGP  
 University Cheikh Anta DIOP  
 Dakar, Senegal  
 eltof@yahoo.fr

#### <sup>4</sup>Mama Raky Ndiaye

Faculty of sciences and Technology  
 Department of Animal Biology  
 BIOPASS Laboratory UMR 022 IRD-CBGP  
 University Cheikh Anta DIOP  
 Dakar, Senegal  
 kiiraa10@hotmail.fr

#### <sup>5</sup>Mamadou Kane

BIOPASS Laboratory  
 UMR 022 IRD-CBGP  
 Bel-Air, Dakar, Senegal  
 Mamadou.kane@ird.fr

#### <sup>6</sup>Mbacke Sembene

Faculty of Sciences and Technology  
 Department of Animal Biology  
 BIOPASS Laboratory UMR 022 IRD-CBGP  
 University Cheikh Anta DIOP,  
 Dakar, Senegal, mbacke.sembene@ird.fr

**Abstract:** In Senegal, as in most neighbouring Sahelian states, corn is one of the major food crops and its promotion is one of the basic objectives of food security defined in most countries. While food security means producing enough by promoting new seed varieties, it also involves storage of this production. The post-harvest component was then wrongly little considered in Senegal. Thus *Sitophilus* spp. causes losses ranging from 25 to 40% in farmers' stocks. However, their biology and physiology are poorly understood because of their endogenous development and the lack of specific gene molecular manipulation. The aim of this study is to identify both the different species in the genus *Sitophilus* infesting stored maize in Senegal and Republic of Guinea and the number, the polymorphic diversity and genetic structure of haplotypes that exist there, through the analysis of cytochrome B gene. On the other hand, it is to determine the genetic impact of storage infrastructure and agro-climatic zones in the affinities between allopatric populations and population dynamics. In Senegal the genus *Sitophilus* is composed of *Sitophilus zeamais* and *Sitophilus oryzae* found specifically in areas that are their own and without overlapping sympatric. They are characterized respectively by an average genetic differentiation of 5.26% based on 18 haplotypes and 6.68% with 16 haplotypes. In addition to the genetic structure of this genus based on agro-climatic zones, genetic distances between populations of *S. zeamais* are negatively correlated with geographic distance. There is a decreasing gradient of genetic diversity based on storage infrastructure starting from stores of traders (market) to the kitchen passing respectively through the producer stores, attics, fields and huts. The high genetic diversity observed in stores and markets is the result of the effect of food by cross infection between cereals.

**Keywords:** *Sitophilus zeamais*, *Sitophilus oryzae*, *Sitophilus* spp, storage infrastructure, agro-climatic zone, Cytochrome B, Genetic diversity, Maize.

4°\_ Sarr A. G. R. J., Dia A. K. M., Ndiaye M. R. and Sembène M., (2016): Genetic Structure of Two *Sitophilus* (*Coleoptera*, *Curculionidae*) Species According to Storage Infrastructures and Agro-ecological Areas. International Journal of Science and Advanced Technology, 6 (4): 1-11.

## Genetic Structure of Two *Sitophilus* (*Coleoptera*, *Curculionidae*) Species According to Storage Infrastructures and Agro-ecological Areas

Adiouma Georges Robert Jacques SARR\*

Searcher Student

Department of Animal Biology,

Faculty of Sciences and Technics,

University Cheikh Anta DIOP of Dakar, Senegal.

PO Box 5005 Dakar, Senegal

[adiouma1987@hotmail.fr](mailto:adiouma1987@hotmail.fr)

Mama Racky NDIAYE

Searcher Student

Department of Animal Biology,

Faculty of Science and Technics,

University Cheikh Anta DIOP of Dakar, Senegal.

PO Box 5005 Dakar, Senegal.

[mamaracky.ndiaye@ucad.edu.sn](mailto:mamaracky.ndiaye@ucad.edu.sn)

Cheikh Abdou Khadre Mbacké DIA

Searcher Student

Department of Animal Biology,

Faculty of Sciences and Technics,

University Cheikh Anta DIOP of Dakar, Senegal.

PO Box 5005 Dakar, Senegal.

[cheikhabdoukhardembacke.dia@ucad.edu.sn](mailto:cheikhabdoukhardembacke.dia@ucad.edu.sn)

Mbacké SEMBENE

Titular professor

Department of Animal Biology,

Faculty of Science and Technics,

University Cheikh Anta DIOP of Dakar, Senegal.

PO Box 5005 Dakar, Senegal.

[mbacke.sembene@ird.fr](mailto:mbacke.sembene@ird.fr)

**Abstract** — Senegal and the Republic of Guinea are made up in majority of agricultural population, which most often stores grain after harvest. During this storage, a major party is attacked by insect pests as *Sitophilus zeamais* and *Sitophilus oryzae* commonly known as weevils. They can cause losses up to 20-30%. Thus the objective of this study is to determine the genetic structuring according to the agro-ecological zones and infrastructure of storage of two species of *Sitophilus* (*Sitophilus zeamais* and *Sitophilus oryzae*). This is done by PCR-Sequencing of the gene for Cytochrome b and Cytochrome oxidase I. The results indicate that the two species are genetically different. At the level of the Cyt. b, *S. zeamais* present a genetic structuring at the level of some infrastructure preparations and for the COI, genetic structuring was observed based on infrastructure preparations and agro-ecological zones.

**Keyword:** *Sitophilus zeamais*; *Sitophilus oryzae*; PCR-Sequencing; Cytochrome b; Cytochrome oxidase; Storage infrastructure; agro-ecological zone

1. **INTRODUCTION** Senegal and the Republic of Guinea consist mostly of agricultural population with respectively 80% [13] and 82% [1]. Both have virtually the same crops and are in Sahel where the dry season lasts most of the year, crop storage is a question of survival. However during this storage a large part is attacked by insect which are responsible of losses that may exceed 30% [8]; [9]. And they promote thereby the colonization of stocks by fungi like *Aspergillus flavus* producing the aflatoxin [9]. These losses are mostly the consequences of activities of genus *Sitophilus* commonly called weevil such as *Sitophilus zeamais* and *Sitophilus oryzae* which are similar morphologically. Moreover, in rural areas where techniques of storage are little developed, *S. zeamais* can cause losses post-harvest up to 90% for 5 months of storage [10]. Face these many damage, several control techniques have been developed from chemical struggles with consequences on the environment, the physical struggles

5°\_Ndiaye M. R., Ndiaye N.P., (2017) : Différenciation et Structure Génétique des Populations de *Sitophilus* spp (*Coleoptera, Curculionidae*), Ravageurs de Céréales Stockées en Afrique de l'Ouest. Actes du Collège doctoral inter-régional en Biotechnologies végétales et Agroalimentaire de l'AUF, pp 55-60.

Actes du collège doctoral :  
**Biotechnologies végétales  
et Agro-alimentaire**

**AGENCE  
UNIVERSITAIRE  
DE LA FRANCOPHONIE**

BIOTECHNOLOGIES VÉGÉTALES

### Structuration Génétique & Relations Phylogéographiques entre Populations de *Sitophilus zeamais* (*Coleoptera, Curculionidae*), Ravageur du Maïs Stocké en Afrique de l'Ouest

Mama Racky Ndiaye<sup>1</sup>, Ndéye Penda Ndiaye<sup>1</sup>, Adjourma G. R. J. Sarr<sup>1</sup>,  
Cheikh A. K. Dia<sup>1</sup>, Cheikh Thiaw<sup>1</sup>, Fatimata Mbaye<sup>1</sup> & Mbacké Sembere<sup>2</sup>

1. Laboratoire de BIOPASS, Campus IRD Bel-Air BP 1386 Dakar ; Département de Biologie animale, Faculté des Sciences & Techniques (UCAD) ; Boite postale : 5005 Dakar (Sénégal).  
2. Centre National de Recherches Agronomiques-Institut Sénégalais de Recherches Agricoles (ISRA-CNRA) de Bambey (Sénégal).

Correspondance : mamaracky.ndiaye@ucad.edu.sn ; mbacke.sembere@ird.fr

**Résumé**

Dans les pays en développement, 40-50% des pertes post-récoltes sur le maïs sont dues aux insectes, en majorité des charançons. Une bonne gestion des cultures et des céréales stockées est donc indispensable et passe par une meilleure compréhension des mécanismes qui régissent l'évolution et la structuration des populations. L'objectif de cette étude est de déterminer la structuration génétique et les relations phylogéographiques des populations de *Sitophilus zeamais* en Afrique de l'Ouest. Le Cytocrome B a été partiellement séquencé. Après alignement et correction des séquences, les paramètres de diversité et de structuration génétiques ont été calculés et les relations phylogéographiques ont été établies. L'étude révèle la présence de 21 haplotypes. L'haplotype majoritaire représente 63,63% de la population globale et est originaire de l'Ouest à l'Afrique centrale. Les analyses d'AMOVA ( $F_{ST}=0,14$ ,  $P<0,05$  &  $F_{ST}^{ind}=0,10$ ,  $P<0,001$ ) et de PERMUT ( $N_{ST}=0,301$  &  $G_{ST}=0,199$  ;  $P<0,05$  ;  $N_{ST}^{ind}=0,647$  &  $G_{ST}^{ind}=0,369$  ;  $P<0,01$ ) indiquent une structuration génétique et un signal phylogéographique entre les zones sénégalaise et soudanaise. Cependant elle est plus marquée entre populations avec, en moyenne, 70% de la variance totale alors que seulement 14% est attribuée à une différenciation entre régions biogéographiques. L'analyse SAMMNA ( $\beta_{ST}=0,80$  ;  $P<0,033$ ), le réseau d'haplotypes et l'analyse bayésienne suggèrent l'existence de 2 lignées phylogénétiques présentant une barrière au flux génétique. L'évolution démographique par les indices de diversité, les tests démographiques et les indices de mismatch distribution révèlent que la population globale a subi dans le passé un goulot d'étranglement suivi d'une expansion démographique modérée au Pléistocène. Enfin, une étude est en perspective pour faire une analyse combinée avec le gène du COI suivi de la caractérisation de la structuration génétique et de l'estimation du flux génétique avec le génotypage des marqueurs microsatellites.

**Mots clés :** *Sitophilus zeamais* ; Structuration génétique ; Relations phylogéographiques ; Méthode bayésienne ; Cytocrome B ; Pays ouest-africains.

**Introduction**

Originaire d'Amérique centrale plus précisément au Mexique et introduit en Afrique vers le XVIII<sup>ème</sup> siècle par les explorateurs portugais, le maïs (*Zea mays*, L., 1753) est une denrée de base pour beaucoup de pays africains. Cependant comme toutes les céréales, le maïs n'est pas épargné des dégâts causés par les insectes ravageurs de stocks et plus particulièrement, les charançons qui peuvent dégrader 25% voire 40% de la récolte en 6 mois de stockage<sup>(1)</sup>. Étant des insectes primaires, ils ouvrent la voie aux insectes secondaires comme *Tribolium confusum* (Du Val, 1864) mais aussi aux champignons qui, en plus, sécrètent des substances nuisibles et cancérogènes telles que l'aflatoxine, compromettant ainsi la

55





6°\_Ndiaye M. R. and Sembène M., (2018): Genetic Structure and Phylogeographic Evolution of West African Populations of *Sitophilus zeamais* (Coleoptera, Curculionidae). In Press, *Journal of Stored Products Research*. 77: 135-143.

Journal of Stored Products Research 77 (2018) 135–143

Contents lists available at ScienceDirect

**Journal of Stored Products Research**

journal homepage: [www.elsevier.com/locate/jSpr](http://www.elsevier.com/locate/jSpr)

## Genetic structure and phylogeographic evolution of the West African populations of *Sitophilus zeamais* (Coleoptera, Curculionidae)

Mama Racky Ndiaye<sup>a</sup>, Mbacké Sembène

Département de biologie animale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop de Dakar, Equipe GENESPOF, PB: 5005 Dakar, Senegal

**ARTICLE INFO**

Article history:  
Received 13 April 2018  
Accepted 16 April 2018

Keywords:  
*Sitophilus zeamais*  
Sub-saharan Africa  
Biogeographical regions  
Phylogeography  
Genetic structure  
Mitochondrial genes

**ABSTRACT**

Delimiting populations of a species in space and assessing the degree of connectivity between them are essential steps to the development of management scenarios. Populations more or less spatially isolated and reproductively independent should be considered because local genetic luggage allows adaptation to specific conditions. This is especially true in beetles such as *Sitophilus zeamais* (Motschulsky, 1855), which have an extraordinary plasticity and potential destruction of stored cereals. The objective of this study was to characterize genetic structure and phylogeographic evolution of *S. zeamais* populations associated with maize in two biogeographical regions of Africa, using genetic markers (sequencing) associated with Bayesian approaches. Portions of mitochondrial gene were sequenced from individuals sampled in West and Central Africa. A total of 112 (Cytochrome *B*), 109 (Cytochrome oxidase *I*) and 99 (concatenated matrix) sequences were used for analysis. An analysis of sequences revealed the presence of 30 haplotypes. The analyses of sequences exhibited a rather high level of polymorphism ( $Hd = 0.789 \pm 0.039$ ;  $\pi = 0.005 \pm 0.001$ ). A high level of haplotype diversity and low nucleotide diversity were observed in West Africa ( $Hd = 0.522 \pm 0.023$ ;  $\pi = 0.007,0012$ ). These both parameters were twice lower in Central Africa populations ( $Hd = 0.512 \pm 0.085$ ;  $\pi = 0.003 \pm 0.001$ ). The results indicated also a genetic structure without phylogeographic pattern. Most of molecular variances were due to individual characteristics in a population; whereas a little molecular divergence was related to a differentiation between the biogeographical regions in West and Central Africa. These populations were characterized by private and common haplotypes. The combined analyses of Bayesian phylogenetic, molecular dating and neutrality tests have suggested the existence of 4 main clusters that separated from themselves from a bottleneck followed by a moderate demographic expansion since Pleistocene period.

© 2018 Elsevier Ltd. All rights reserved.

**1. Introduction**

Maize (*Zea mays* L., 1753) is originated in Central America, most likely in Mexico, and imported to Africa in the sixteenth century by Portuguese explorers. It is a staple food for many African countries. Like all cereals, maize is susceptible to damage caused by stored product weevils which can destroy 25%–40% of stored harvest in 6 months (Nukerine et al., 2002). Among these weevils considered as primary pest, *Sitophilus zeamais* (Motschulsky, 1855), the maize weevil, facilitates infestations of secondary pests such as *Tribolium confusum* (Du Val, 1808) and also promotes fungal contamination, including aflatoxin contamination that compromises food safety

and health of vulnerable groups in Africa. It is estimated that more of 200 million people in Sub-Saharan Africa (ASS) depend on corn as a source of safety food and economic welfare (Macaulay and Ramadja, 2015). In addition, among the first 20 countries that generate 96% of the total production of maize in ASS, seven are in West Africa (FAOSTAT, 2015). In addition, previous studies have shown the extent of damage caused by *Sitophilus* weevils (Nukerine et al., 2002).

The economic importance of cereals as a maize is related to commercial exchange between rural populations. This conveyance of grains from one country to another one could be accompanied by introduction of infested products (Delobel, 1995; Ndiaye, 2014), which contributes to the spread of this insect pest in the world and increases depreciation stocks of cereals. Consequently, *S. zeamais* constitutes an agro-economic and health problem. Yet, characterization of populations by molecular analysis is effective only for a minority of sub-Saharan populations.

\* Corresponding author. Tel.: +221 77 750 10 10; fax: +221 825 29 20.  
E-mail addresses: [mamanrackyndiaye@facat.educast.sn](mailto:mamanrackyndiaye@facat.educast.sn) (M.R. Ndiaye), [mbacke.sembene@facat.educast.sn](mailto:mbacke.sembene@facat.educast.sn) (M. Sembène).

<https://doi.org/10.1016/j.jSpr.2018.04.003>  
0022-474X/© 2018 Elsevier Ltd. All rights reserved.

**2°) POSTERS**

- **Ndiaye M. R.**, Ndong A., Bambou A. E., Dia C. A. K. M., Diome T., Thiaw C. and Sembene M., (2015) : Différenciation et Structure Génétique des Populations de *Sitophilus* spp (*Coleoptera, Curculionidae*), Ravageurs de Céréales Stockées en Afrique de l'Ouest. Collège doctoral (Caméroun).
- **Ndiaye M. R.** and Sembène M., (2015) : Phylogénie et Structure Génétique des Populations de *Sitophilus zeamais* (*Coleoptera, Curculionidae*), Ravageur de Céréales Stockées en Afrique de l'Ouest. Doctorial EDSEV.

**3°) COMMUNICATIONS**

- **Ndiaye M. R.**, Ndong A. and Sembene M., (2013) : Impact des zones agro-écologiques et Infrastructures de stockage sur la diversité et la structuration génétiques de *Sitophilus* spp (*Coleoptera, Curculionidae*) ravageurs du maïs stocké. Colloque International et Interdisciplinaire sur le thème : Plantes, Biodiversité et Ressources thérapeutique dans le Sahel. Du 10 au 12 Décembre 2013 à Dakar.
- **Ndiaye M. R.**, (2014) : Présentation du projet de la thèse, le 4 Novembre 2014 à Hammamet (Tunisie) au premier regroupement du Collège doctoral inter-régional en Biotechnologies végétales et Agroalimentaire de l'AUF.
- **Ndiaye M. R.**, (2015) : Différenciation et Structure Génétique des Populations de *Sitophilus* spp (*Coleoptera, Curculionidae*), Ravageurs de Céréales Stockées en Afrique de l'Ouest. Présenté le 22 Avril 2015 à Yaoundé (Caméroun) au deuxième regroupement du Collège doctoral inter-régional en Biotechnologies végétales et Agroalimentaire de l'AUF.
- **Ndiaye M. R.**, (2016) : Phylogénie and Génétique des Populations de *Sitophilus zeamais* (*Coleoptera, Curculionidae*), Ravageur de Céréales Stockées en Afrique de l'Ouest. Présenté le 16 Février 2016 à Abidjan (Côte d'Ivoire) au troisième regroupement du Collège doctoral inter-régional en Biotechnologies végétales et Agroalimentaire de l'AUF.
- **Ndiaye M. R.**, (2017) : Evolution phylogéographique and structure génétique des Populations de *Sitophilus zeamais* (*Coleoptera, Curculionidae*), ravageur de maïs stocké en Afrique de l'Ouest. Présenté le 22 Mars 2017 à Meknès (Maroc) au quatrième regroupement du Collège doctoral inter-régional en Biotechnologies végétales et Agroalimentaire de l'AUF.