

## TABLE DES MATIERES

LISTE DES ABREVIATIONS .....	i
GLOSSAIRE.....	ii
LISTE DES FIGURES.....	iii
LISTE DES TABLEAUX.....	iv
LISTE DES ANNEXES.....	v
I. INTRODUCTION .....	1
II. ETUDE BIBLIOGRAPHIQUE .....	2
A. Puces .....	2
1. Cycle de développement .....	2
2. Morphologie d'une puce adulte .....	4
3. Puces vectrices de pathogènes .....	5
B. La maladie peste.....	5
C. Généralités sur la phylogéographie.....	7
1. La phylogéographie.....	8
2. Intérêts de la phylogéographie en entomologie médicale .....	9
D. Les marqueurs moléculaires.....	9
III. MATERIELS ET METHODES.....	11
A. Les sites d'études .....	11
B. Identification morphologique des spécimens.....	11
C. Extraction d'ADN .....	11
1. Principe .....	11
2. Matériels et réactifs utilisés.....	12
3. Méthode.....	12
D. Mises au point des marqueurs moléculaires .....	13
1. Principe de la PCR .....	13
2. Les réactifs et matériels utilisés .....	15

3.	Mise au point par PCR du marqueur COI.....	17
4.	Mise au point par PCR du marqueur COI-tailed.....	17
5.	Mise au point par PCR du marqueur COII .....	18
6.	Mise au point par PCR du marqueur 28S .....	19
7.	Détection des amplifiats PCR .....	20
E.	Etudes préliminaires en phylogéographie de <i>X. cheopis</i> à Madagascar....	21
1.	Principe du séquençage d'ADN .....	21
2.	Amplification avec les marqueurs COI-tailed.....	22
3.	Séquençage.....	23
4.	Traitemet bioinformatique des données .....	23
IV.	RESULTATS .....	25
A.	Extraction d'ADN .....	25
B.	Les conditions optimales pour les différents marqueurs.....	25
1.	Mise au point par PCR du marqueur COI.....	25
2.	Mise au point par PCR du marqueur COI-tailed.....	26
3.	Mise au point par PCR du marqueur COII .....	26
4.	Mise au point par PCR du marqueur 28S .....	28
C.	Diversité génétique et phylogéographie de <i>X. cheopis</i> .....	29
1.	Amplification avec le marqueur COI-tailed.....	29
2.	Profils des électrophorégrammes après séquençage .....	30
3.	Traitemet bioinformatique des données .....	31
V.	DISCUSSION .....	36
VI.	CONCLUSION .....	39
VII.	REFERENCES BIBLIOGRAPHIQUES .....	40

## ANNEXES

### RESUME

### ABSTRACT

## **LISTE DES ABREVIATIONS**

ADN : Acide desoxyribonucléique

ARN: Acide ribonucléique

BET : Bromure d'éthidium

BSA : Sérum albumine bovine

COI : Cytochrome oxydase I

COII : Cytochrome oxydase II

dNTPs : Désoxyribonucléosides triphosphates

DG: Distance génétique

mtDNA : ADN mitochondrial

µl : microlitre

ng : nanogramme

PCR : Polymerase chain reaction

PBS : Phosphate buffer saline

pbs : Paires de bases

pH : Potentiel d'hydrogène

TBE : Tris Borate EDTA

Tm : melting temperature/ température de fusion

*X. cheopis* : *Xenopsylla cheopis*

## GLOSSAIRE

**Aptère:** qualifie les insectes dépourvus d'ailes.

**Distance génétique:** mesure de la divergence génétique entre deux populations ou individus ou du degré de parenté entre espèces en fonction des différences constatées entre les séquences de gène homologue. Elle est exprimée en centimorgan (cM).

**Electrophorèse :** technique de séparation des protéines déposées sur un gel selon leur vitesse de déplacement sous l'action d'un champ électrique.

**Evolution moléculaire** ou évolution du génome : ensemble des modifications d'un génome au cours de l'évolution des espèces.

**Génotype :** constitution génétique d'un organisme.

**Holométabole :** qualifie les insectes chez qui le passage de l'état de larve à l'état adulte se fait par la transition d'un état de nymphe. La morphologie et l'écologie des larves et des adultes sont généralement très différentes.

**Hygrophile :** organisme pour qui l'humidité est nécessaire à son bon développement.

**Mutation :** modification du matériel génétique (ADN ou ARN) par changement d'une ou plusieurs base(s). Cette modification peut être une insertion ou une délétion ou une substitution de bases.

**Phénotype :** manifestation apparente de la constitution du génome.

**Phototropisme :** capacité d'un organisme à s'orienter par rapport à la lumière.

**Similitude:** ressemblance entre deux séquences (généralement exprimée en degré de similitude relatif au % de substitutions entre les deux séquences).

## **LISTE DES FIGURES**

Figure 1: Cycle de développement d'une puce.....	2
Figure 2: Morphologie d'une puce femelle adulte ( <i>Ctenocephalides</i> ).....	4
Figure 3: Cycle de transmission de la peste .....	7
Figure 4: Etapes d'un cycle PCR.....	13
Figure 5: Cycle de température pour l'amplification avec le marqueur COI-tailed ..	18
Figure 6: Profil de l'électrophorèse après amplification avec le marqueur COI .....	25
Figure 7: Cycle de température optimale d'amplification avec le marqueur COI .....	26
Figure 8: Profil de l'électrophorèse après amplification avec le marqueur COI-tailed .....	26
Figure 9: Profils de l'électrophorèse après amplification avec le marqueur COII... ..	27
Figure 10: Cycle de température optimale d'amplification avec le marqueur COII .	27
Figure 11: Profil de l'électrophorèse après amplification avec le marqueur 28S.....	28
Figure 12: Cycle de température optimale d'amplification avec le marqueur 28S ....	29
Figure 13: Résultats de la PCR avec COI-tailed par site .....	30
Figure 14: Profils des électrophorégrammes après séquençage.....	31

## **LISTE DES TABLEAUX**

Tableau 1: Liste des marqueurs mis au point.....	16
Tableau 2: Mélange réactionnel pour l'amplification avec le marqueur COI .....	17
Tableau 3: Conditions thermiques d'amplification pour la mise au point du marqueur COI.....	17
Tableau 4: Mélange réactionnel pour l'amplification avec le marqueur COI-tailed..	18
Tableau 5: Mélange réactionnel pour l'amplification avec le marqueur COII.....	19
Tableau 6: Conditions thermiques d'amplification pour la mise au point du marqueur COII .....	19
Tableau 7: Mélange réactionnel pour l'amplification avec le marqueur 28S.....	20
Tableau 8: Conditions thermiques d'amplification pour la mise au point du marqueur 28S .....	20
Tableau 9: Composition du mélange réactionnel pour l'amplification avec les marqueurs COI-tailed.....	22
Tableau 10: Distribution des haplotypes par site .....	32
Tableau 11: Nombre d'individus par haplotype .....	33
Tableau 12: Distance génétique (DG) entre les représentants de chaque haplotype par site .....	35

## **LISTE DES ANNEXES**

**Annexe 1** : Liste des sites, leurs coordonnées GPS et le nombre d'échantillons par site

**Annexe 2** : Composition du *One Taq® 2X master mix*

**Annexe 3** : Préparation du gel d'agarose 1,5%

**Annexe 4** : Concentration des extraits d'ADN

## I. INTRODUCTION

Dans le monde, les puces qui sont des insectes parasites portent un intérêt médical et vétérinaire. Certaines espèces sont des vecteurs de pathogènes ; la plus connue étant *Xenopsylla cheopis*, une espèce vectrice de la bactérie de la maladie peste : *Yersinia pestis*. Cette maladie, bien qu'elle soit ancienne, continue de sévir dans certains pays du monde et sa déclaration est obligatoire (Duchemin et coll., 2006). Une des méthodes utilisées pour lutter contre cette maladie est la lutte anti-vectorielle (Zhou et coll., 2012).

A Madagascar, la plupart des études réalisées sur les puces ont concerné la systématique, la biologie et la résistance aux insecticides. Les études génétique, phylogénétique et phylogéographique n'ont jamais été effectuées. Pourtant, elles permettent de mieux connaître la diversité génétique de l'espèce, l'évolution moléculaire associée à la dispersion des populations de vecteurs et les aspects épidémiologiques de la maladie transmise.

Dans cette étude, nous nous sommes intéressés aux aspects diversité génétique et phylogéographie. Les objectifs de cette étude sont 1) de mettre au point les différents marqueurs moléculaires (nucléaire et mitochondriaux) potentiellement convenables pour les études phylogéographiques, 2) de déterminer la diversité des *X. cheopis* à Madagascar et 3) de réaliser une étude préliminaire sur la phylogéographie de cette puce vectrice dans l'île. Pour ce faire, l'extraction d'ADN suivie de l'amplification PCR pour les mises au point des marqueurs ont été effectuées. Une fois que les conditions optimales d'amplification ont été obtenues, l'amplification PCR suivie du séquençage partiel d'ADN ont été effectués dans le but d'obtenir des séquences nucléotidiques qui ont été ensuite analysées.

Le présent travail comporte quatre parties, la première partie est consacrée à l'étude bibliographique sur les puces, la maladie peste, la phylogéographie et les marqueurs moléculaires. La deuxième partie concerne les matériels et méthodes utilisés. La troisième partie regroupe les résultats obtenus et la dernière partie est consacrée à la discussion qui est suivie par une conclusion générale.

## II. ETUDE BIBLIOGRAPHIQUE

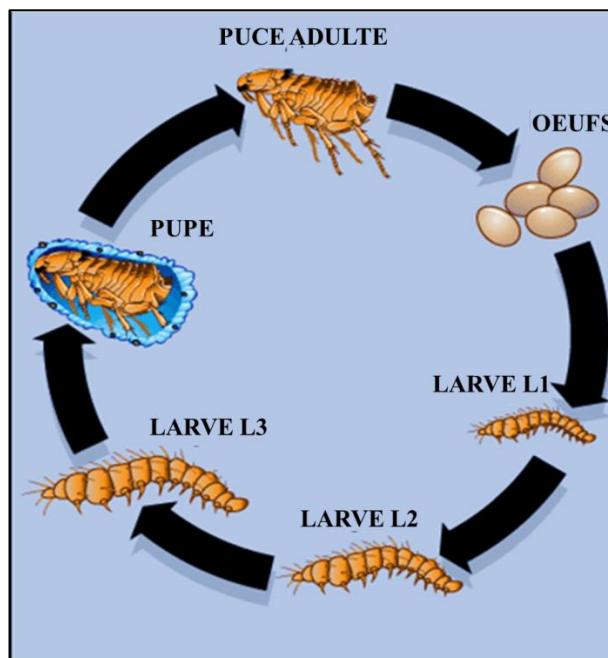
### A. Puces

Les puces sont des insectes holométaboles, aptères et suceurs-piqueurs appartenant à l'Ordre des Siphonaptères. Les Siphonaptères sont constitués de 16 familles, 246 genres et 2575 espèces (Whiting et coll., 2008). Ce sont des insectes parasitant les oiseaux et les mammifères tels que l'homme, les rongeurs et les chauve-souris. Parmi les espèces connues, on peut citer les puces de l'homme *Pulex irritans* (Linné, 1758) ; les puces de chien *Ctenocephalides felis* (Bouché, 1835) ; les puces de rats *X. cheopis* (Rothschild, 1903) ; les puces d'oiseaux *Xenopsylla gratiosa* (Jordan et Rothschild, 1923) ; les puces de gerbilles *Xenopsylla phloxera* (Hopkins, 1928) et les puces de chauves-souris *Ischnopsyllus octactenus* (Kolenati, 1856).

Les puces sont connues du point de vue médical et vétérinaire, étant donné que certaines espèces sont vectrices d'agents pathogènes infectant les hommes et les animaux, mais aussi, que leur piqûre peut entraîner des dermatites.

### 1. Cycle de développement

Etant des insectes holométaboles, les puces ont un cycle de développement comportant quatre stades bien définis (Figure 1).



**Figure 1:** Cycle de développement d'une puce.

### ➤ Le stade œuf

La puce femelle pond des œufs isolés ou par séries de 2 à 6 dans les litières de l'hôte, dans les poussières des habitations ou sur le pelage de l'hôte mais étant donné qu'ils n'y sont pas collés, ils sont souvent retrouvés au sol. Elle peut pondre une centaine d'œufs durant sa vie selon les espèces voire même des milliers (cas de *Tunga penetrans*). Ce stade dure environ une semaine (Rodhain et Perez, 1985).

### ➤ Le stade larvaire

Après leur éclosion, les œufs donnent des larves vermiformes de 0,5 mm de longueur et dépourvues d'yeux mais présentant une pièce buccale de type broyeur. Ce dispositif leur permet de se nourrir des détritus et des déjections de puces adultes (Franc, 1994) leur valant l'appellation de détritivores. Elles ont un phototropisme négatif et sont très hygrophiles (Rodhain et Perez, 1985). Ces larves se retrouvent dans les tapis, les moquettes ou les débris organiques et passent par 3 stades larvaires : L1, L2 et L3. La durée de chaque stade est variable et dépend de l'humidité, de la disponibilité de la nourriture et de différents facteurs environnementaux (Silverman et Rust, 1985). La différence entre ces stades repose surtout sur leur taille qui augmente à chaque stade.

### ➤ Le stade de pupe

Après une troisième mue donc le stade L3, la larve prend une forme en U et se fixe sur un support vertical afin de tisser un cocon qui la recouvre. Elle se développe en pupe 18 heures après la formation de ce cocon (Rust et Dryden, 1997). C'est un stade immobile durant lequel la larve L3 ne s'alimente pas mais subit une métamorphose. Il dure en moyenne 1 à 2 semaines. Toutefois, quand les conditions sont défavorables au développement, la durée du stade est variable et peut être longue (exemple : entre 14 et 182 jours pour *X. cheopis*) (Rodhain et Perez, 1985).

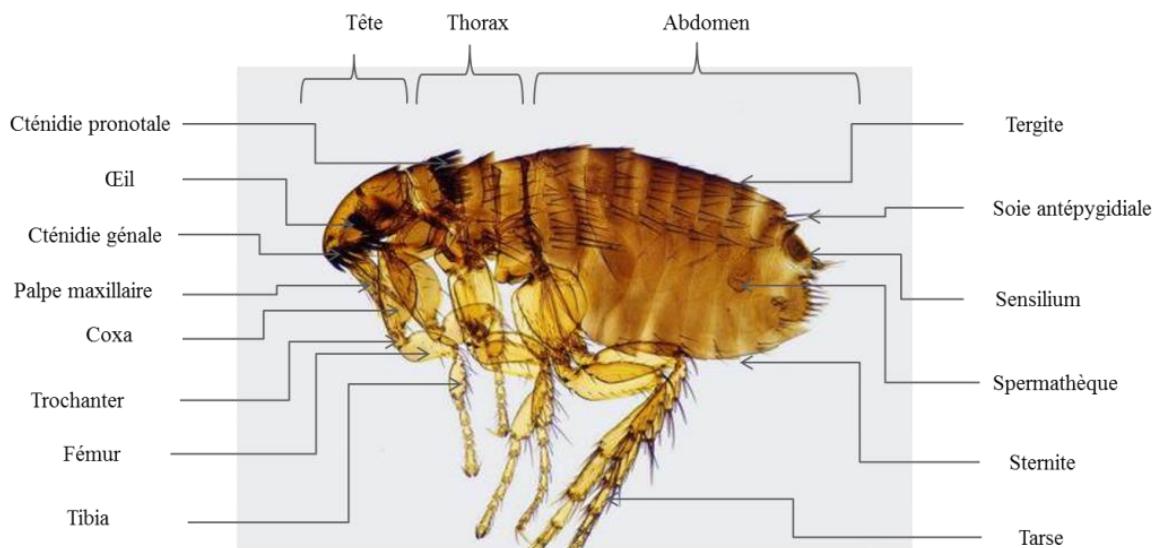
### ➤ Le stade adulte ou imago

Quand la puce émerge du cocon, elle pique immédiatement un hôte afin de se nourrir de son sang : les puces sont des insectes hématophages (Rodhain et Perez, 1985). Suivant les espèces, les puces adultes peuvent parasiter leurs hôtes sous trois formes.

- Les puces vivent en permanence sur leur hôte et ne le quittent que pour parasiter un autre hôte ; on parle de puces pilicoles ou puces de fourrures telles que *X. cheopis*, *Pulex irritans* et *Ctеноcephalides canis*.
- Les puces peuvent passer la plupart de leur vie dans les terriers ou les sols et n'entrent en contact avec leur hôte que lors des repas ; elles sont moins mobiles. Ces puces sont dénommées puces nidicoles ou puces de terrier. C'est le cas de *Ceratophyllus gallinae*.
- Certaines puces sont dites sédentaires et pénétrantes. Elles vivent fixées par les pièces buccales sur leur hôte ou s'enfoncent dans la peau de celui-ci et leur seule communication avec l'extérieur est leur orifice de ponte. C'est le cas de l'espèce *Echidnophaga gallinacea* qui se fixe autour des yeux des volailles et plus rarement des chiens, et aussi celui de *Tunga penetrans* qui s'enfonce dans la peau de l'homme.

## 2. Morphologie d'une puce adulte

La puce adulte ou imago est longue de 0,8 à 6,5 mm et possède un corps aplati et 3 paires de pattes lui permettant de faire des sauts (Rodhain et Perez, 1985). Son corps comporte trois parties dont la tête, le thorax et l'abdomen (Figure 2).



**Figure 2:** Morphologie d'une puce femelle adulte (*Ctenocephalides*).

### **3. Puces vectrices de pathogènes**

Par leur repas sanguin ou leurs fèces, certaines espèces de puces sont des vecteurs de maladies plus ou moins graves touchant les animaux et l'homme. C'est le cas de *X. cheopis*, vecteur de *Yersinia pestis* (bactérie responsable de la maladie peste), de *Rickettsia typhi* (bactérie responsable de la maladie typhus murin), et de *Ctenocephalides felis*, vecteur de *Bartonella henselae* (bactérie responsable de la maladie des griffes de chat) (Bitam et coll., 2010).

Dans le monde, environ 80 espèces de puces ont été trouvés naturellement infectés par la peste mais seule une petite proportion est considérée comme bons vecteurs (Gage et Kosoy, 2005 ; Eisen et Gage, 2012). A Madagascar, deux espèces sont connues comme vecteur: *X. cheopis* et *Synopsyllus fonquerneiei* (Andrianaivoarimanana et coll., 2013). Ces puces s'infectent par leur repas sanguin sur un animal infecté (Chanteau, 2006).

## **B. La maladie peste**

La maladie peste peut être transmise à l'homme soit par la piqûre de puces infectées soit par la transmission inter-humaine via l'inhalation d'aérosol infesté. C'est une maladie ancienne qui a causé trois pandémies majeures dans le monde. La première est appelée peste des Justiniens qui a débuté au VIème siècle et a touché les pays au pourtour de la mer méditerranéenne (Stenseth et coll., 2008). La seconde pandémie nommée peste noire est apparue en Europe au XIVème siècle. Des épidémies récurrentes se sont manifestées durant plus de 300 ans. La troisième pandémie a débuté vers le milieu du XIXème siècle en Chine puis, s'est étendue dans le monde (Stenseth et coll., 2008). C'est seulement lors de cette troisième pandémie que son agent étiologique *Yersinia pestis* a été découvert ainsi que ses vecteurs (Trubéa et coll., 2007). Depuis, elle est classée parmi les maladies réémergentes par l'Organisation Mondiale de la Santé (OMS) suite à l'augmentation du nombre de cas déclarés dans le monde entre 1987 et 2001 (36 876 cas dont 2847 mortels) (Chanteau, 2006).

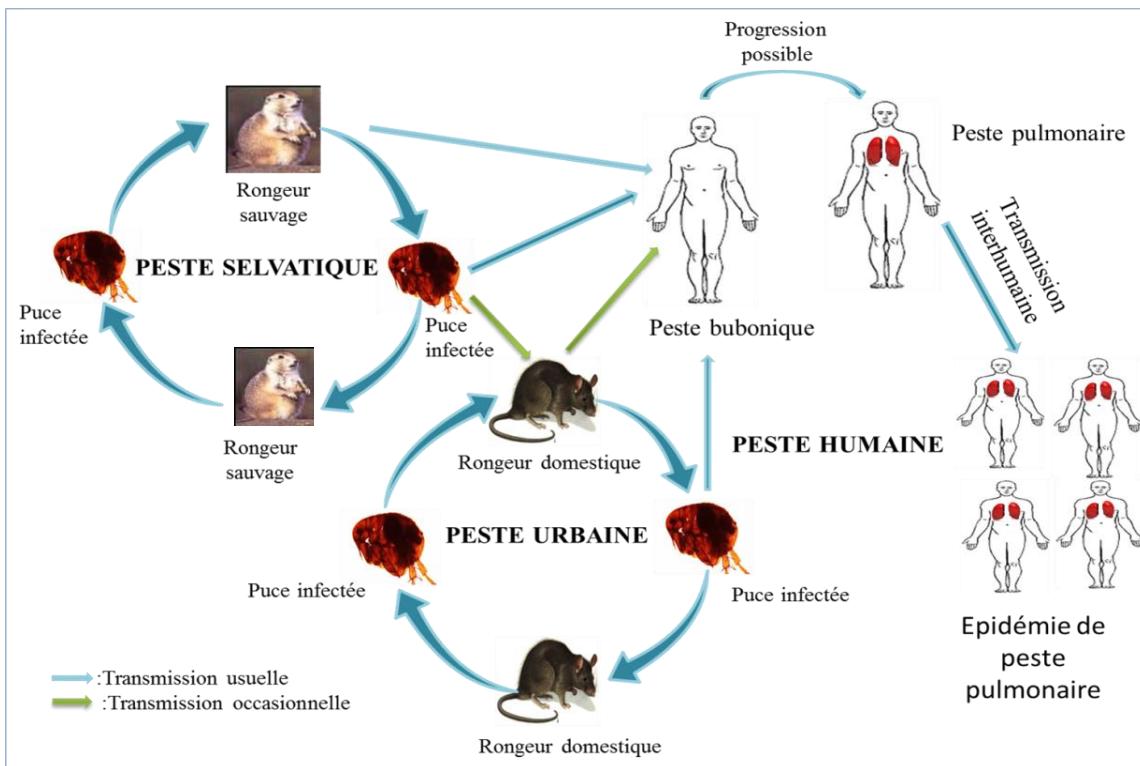
Actuellement, les foyers pestieux sont localisés dans quelques pays de l'Afrique, d'Asie et d'Amérique (Andrianaivoarimanana et coll., 2013). Entre 2004 et 2009, 12 503 cas de peste humaine ont été déclarés dans ces pays et ont causé 843 décès (World Health Organisation, 2010). Grace aux méthodes de lutte mises en

place, le nombre de cas a diminué en 2012 et est estimé à 427 dont 81 mortels (World Health Organisation, 2016).

Madagascar est le pays le plus touché par cette maladie. Elle y fut introduite en 1898 par bateau provenant d'Inde et s'est étendue à Antananarivo en 1921 (Chanteau et coll., 1998). Sa répartition dans toute l'île n'est pas homogène ; elle est généralement présente dans les régions ayant une altitude supérieure à 800 mètres (Chanteau et coll., 1998). La période épidémiologique s'étale de Septembre à Avril sur les Hautes Terres, tandis qu'à Mahajanga, elle se situe entre Juillet et Novembre (Chanteau et coll., 1998).

Deux cycles de transmission de cette maladie sont connus : le cycle selvatique et le cycle urbain (Figure 3). Le cycle selvatique implique les rongeurs sauvages qui sont les réservoirs, les puces vectrices et l'agent pathogène. Ce cycle se fait dans les milieux naturels tels que les milieux forestiers. Sa transmission se fait entre les rongeurs sauvages via la piqûre des puces infectées. Les espèces de puces connues étant impliquées dans le cycle selvatique appartiennent aux genres *Paractenopsyllus*, *Synopsyllus* et *Dinopsyllus* (Chanteau, 2006). Ce cycle continue jusqu'à la mort du rongeur sauvage ou quand la puce trouve une nouvelle source alimentaire comme les rats domestiques *Rattus rattus* (Lotfy, 2015).

Le cycle urbain implique les rats domestiques, les puces et accidentellement l'homme. Ce cycle s'effectue dans les zones d'habitation humaine. Les espèces de puces impliquées sont *X. cheopis* et *Synopsyllus fonquerneiei* (Andrianaivoarimanana et coll., 2013). Une troisième espèce, *Xenopsylla brasiliensis* a été trouvée associée à une épidémie de peste à Mandritsara entre septembre 2013 et Janvier 2014 (Miarinjara et coll., 2016). La peste urbaine peut produire des épidémies à la fois chez l'homme et chez les animaux comme les lapins ou les chats (Trubéa et coll., 2007). Chez l'homme, elle peut se présenter sous trois formes cliniques : la peste bubonique, la peste pulmonaire et la peste septicémique et peut être mortelle en absence de traitement.



**Figure 3:** Cycle de transmission de la peste (source : [www.researchgate.net/figure/6853227\\_fig2\\_Figure-2-Transmission-cycle-Illustration-courtesy-of-Neal-R-Chamberlain-PhD-AT](http://www.researchgate.net/figure/6853227_fig2_Figure-2-Transmission-cycle-Illustration-courtesy-of-Neal-R-Chamberlain-PhD-AT)).

### C. Généralités sur la phylogéographie

Depuis leur création, les êtres vivants sont soumis à des changements morphologiques et physiologiques, et de ce fait, ils évoluent. La première cause de cette évolution est la mutation des gènes qui produisent à leur tour de nouveaux caractères morphologiques et/ou physiologiques qui seront ensuite hérités par la descendance (Nei et Kumar, 2000). Le changement de séquence de ces gènes entraîne le polymorphisme génétique. Ainsi, les êtres vivants sont phénotypiquement et génotypiquement polymorphes. Ces variations de gènes peuvent déterminer l'apparition de nouveaux phénotypes voire même mener au cours de l'évolution à l'apparition de nouvelles espèces (Swyngedauw et Silvestre, 2008).

## *1. La phylogéographie*

Selon Avise (1998), la phylogéographie est définie comme la science qui étudie les principes et processus qui gouvernent la distribution géographique des lignées généalogiques, et en particulier, ceux mis en évidence au sein d'une espèce ou entre espèces étroitement apparentées. L'analyse et l'interprétation de cette distribution requièrent des données à partir de la génétique moléculaire, la génétique des populations, l'éthologie, la démographie, la phylogénie, la paléontologie, la géologie et la géographie historique (Avise, 2000). Le but de la phylogéographie est de comprendre comment les événements historiques ont aidé à comprendre la dispersion géographique des gènes, des populations et des espèces. En comparant les relations évolutives entre lignées avec leur localisation géographique, il est possible de comprendre les facteurs influençant la distribution des variations génétiques (Freeland, 2005).

L'approche phylogéographique la plus utilisée est la construction d'arbre phylogénétique à partir des séquences nucléotidiques en utilisant la méthode du plus proche voisin (Neighbour-joining) basée sur la distance, les méthodes de maximum de parcimonie, maximum de vraisemblance et l'inférence bayésienne ( Emerson et Hewitt, 2005).

Etant une contraction de l'expression «génotype haploïde», le terme haplotype se définit comme la constitution génétique d'un chromosome individuel. Il peut se référer à un ensemble de marqueurs (par exemple les polymorphismes d'un seul nucléotide) qui sont statistiquement associés sur un chromosome simple (FAO 2008). Les haplotypes reflètent les variabilités génétiques des individus ou des populations au sein d'une même espèce. Ainsi, les réseaux haplotypiques sont utilisés pour évaluer la variabilité génétique intraspécifique.

La distance génétique ou similarité nucléotidique est la mesure de la divergence génétique entre deux populations ou molécules ou du degré de parenté entre espèces en fonction des différences constatées entre les séquences de gène homologue. Ainsi, elle permet d'évaluer le polymorphisme génétique. La distance génétique est obtenue par comparaison par paires de deux séquences et en comptant le nombre de nucléotides différents entre elles. Elle est estimée suivant la formule :  $p=n_d/n$  où  $n_d$  et  $n$  représentent respectivement le nombre de nucléotides différents

entre les 2 séquences et le nombre total de nucléotides examinés ( Nei et Kumar, 2000). Mais cela ne renseigne pas sur la vraie estimation des mutations en un site donné au cours du temps. Il faut donc apporter des modifications aux valeurs de ces mesures : on parle alors de distance corrigée.

La phylogéographie est appliquée dans de nombreux domaines tels que l'entomologie ( Van der Mescht et coll., 2015), la virologie (Cuevas et coll.,2012), la biologie animale (Taberlet et Bouvet, 1994) et l'écologie (Newton et coll., 1999).

## ***2. Intérêts de la phylogéographie en entomologie médicale***

La phylogéographie est une discipline qui peut apporter beaucoup de connaissances en entomologie médicale (discipline étudiant les Arthropodes responsables de pathologies). Elle peut offrir des informations sur la variabilité génétique de l'espèce d'insecte vecteur et sur la distribution spatiale de ses populations. Aussi, elle permet de déterminer le mode de dispersion des populations d'insectes et par conséquent, d'inférer les risques épidémiologiques associés. Ces connaissances peuvent aider à développer une méthode de lutte anti-vectorielle plus efficace. En effet, le contrôle des populations de vecteurs est une arme privilégiée dans la lutte contre les maladies vectorielles. Il est évident que des recherches sur la systématique, la génétique, la biogéographie et l'écologie sont fondamentales pour améliorer les stratégies et les méthodes de lutte contre ces insectes (Aymès, 2009).

## **D. Les marqueurs moléculaires**

L'étude de la phylogéographie nécessite l'utilisation de marqueurs moléculaires pouvant révéler le polymorphisme génétique. Ces marqueurs peuvent être d'origine nucléaire (exemple : 28S, 18S, 16S.....) ou mitochondriale (mtCOI, mtCOII.....). Ils servent de repères pour suivre la transmission d'un segment de chromosome d'une génération à une autre (Biochard et coll., 1998). Ils permettent entre autres de quantifier la diversité génétique, de suivre les mouvements des individus, de mesurer la consanguinité, d'identifier les individus à partir des vestiges de spécimens, de caractériser les nouvelles espèces et de retracer les modèles historiques de dispersion (Freeland, 2005).

Les marqueurs nucléaires ciblent l'ADN nucléaire portant des gènes hérités des deux parents. Actuellement, différents marqueurs sont disponibles pour détecter le polymorphisme de l'ADN nucléaire. Ce sont les marqueurs pour les ARN ribosomiques et les marqueurs microsatellites. L'ADN codant pour l'ARN ribosomal est relativement bien conservé au cours de l'évolution. Ce qui en fait un outil tout à fait approprié pour une étude de phylogénie (Délandes, 1997). De plus, cette région conservée offre la possibilité de dessiner des amorces universelles utiles au séquençage d'ADN (Hillis et Dixon, 1991). Les marqueurs microsatellites sont des marqueurs détectant les variations de répétitions de l'ADN. Les régions microsatellites sont constituées par des répétitions en tandem de séquences nucléotidiques très courtes. Le polymorphisme de ces régions réside dans le nombre de ces répétitions.

Les marqueurs mitochondriaux ciblent l'ADN mitochondrial portant des gènes hérités uniquement de la mère (Avise, 1994). L'arrangement de leurs gènes apparaît très stable entre les différentes classes taxonomiques et l'évolution de la séquence de nucléotides est de 1 à 10 fois plus rapide que celle d'un ADN nucléaire (Avise et coll., 1987). Ce qui est un avantage pour les études intraspécifiques ou interspécifiques (Fotini, 2007). Du fait de leurs caractéristiques ainsi que leur abondance, leur structure simple, l'absence d'introns et de recombinaison qui facilite leur isolement, leur amplification et leur séquençage (Avise et coll., 1987), les données obtenues à partir de marqueurs mitochondriaux permettent de construire la phylogénie d'haplotypes pouvant être utilisée pour examiner la distribution géographique des lignées évolutives (Newton et coll., 1999).

### **III. MATERIELS ET METHODES**

#### **A. Les sites d'études**

*X. cheopis* a une distribution géographique très large à Madagascar et est rencontrée dans les différents bioclimats présents dans l'île. Par conséquent, le choix des sites d'échantillonnage a été basé sur la disponibilité des échantillons (spécimens et extraits d'ADN) au laboratoire de l'unité d'Entomologie Médicale de l'Institut Pasteur de Madagascar. Ces échantillons représentent au mieux les différentes régions de l'île. Des échantillons en provenance des îles voisines ont aussi été testés. Ainsi, des échantillons sur 24 sites dont 3 situés dans les îles voisines de Madagascar (Comores, Maurice et Mayotte (projet ARS-OI, 2014)) ont été utilisés dans le cadre de notre étude (Annexe 1).

#### **B. Identification morphologique des spécimens**

Les puces adultes ont été identifiées morphologiquement au laboratoire en utilisant des loupes. Les spécimens identifiés comme appartenant à l'espèce *X. cheopis* ont été utilisés pour la suite des expériences.

#### **C. Extraction d'ADN**

##### ***I. Principe***

L'extraction d'ADN est une étape permettant d'obtenir un extrait d'ADN à partir d'une cellule, d'un tissu, d'un organe ou d'un organisme entier.

Dans cette étude, l'ADN a été extrait par la méthode de broyage en présence de Chelex résine 100 à 5% qui est un puissant agent chélatant piégeant les enzymes dégradants (DNase) et les contaminants pouvant inhiber les réactions tout en laissant l'ADN en solution. Le broyage conduit à la lyse de la membrane cellulaire induisant la libération des contenus cellulaires vers l'extérieur qui sont ainsi mis en contact avec le réactif d'extraction.

## **2. Matériels et réactifs utilisés**

### **i. Les réactifs utilisés**

- *Chelex® 100 sodium form* (Sigma-Aldrich)
- Phosphate buffer saline (PBS) 1X

### **ii. Les matériels utilisés**

- Bille et tube stérile
- Tissus Lyser II (Retsch))
- *Nanodrop* (2000, Thermoscientific)
- Centrifugeur Rotina 38R (Hettich)
- Plaque chauffante (Stuart)
- Vortex (Fisher Scientific)
- Micropipettes

## **3. Méthode**

L'ADN a été extrait à partir des spécimens identifiés. La puce a été mise dans un tube contenant 20 µl de PBS 1X et une bille. Le spécimen a été broyé à 30 vibrations/seconde pendant 30 secondes à l'aide du Tissue Lyser II permettant ainsi la lyse cellulaire. Après l'ajout de 100µl de Chelex résine 100 5%, la solution a été vortexée pendant 15 secondes à 40Hz. Une centrifugation à 14000 tours/minute pendant 2 minutes suivie d'une incubation de 20 minutes dans un bain à sec à 100°C ont fait suite à cette étape. La bille a été enlevée et la solution a été vortexée à nouveau à 40Hz pendant 15 secondes puis, centrifugée à 14000 tours/minute pendant 2 minutes. A la fin de ces différentes étapes, le surnageant contenant l'extrait d'ADN a été récupéré tout en évitant de toucher au culot qui contenait des produits pouvant inhiber la réaction de polymérisation en chaîne.

Les extraits d'ADN récupérés ont été quantifiés par spectrophotométrie utilisant l'appareil *Nanodrop* à une longueur d'onde égale à 260nm.

## D. Mises au point des marqueurs moléculaires

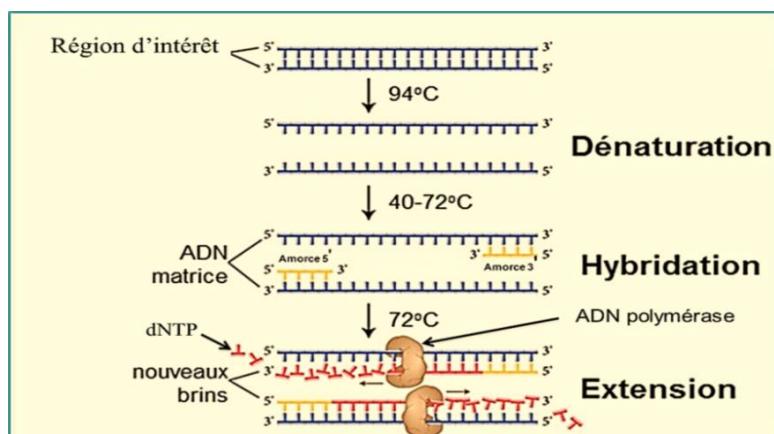
Les mises au point des marqueurs ont été effectuées par la méthode de la réaction de polymérisation en chaîne (*Polymerase Chain Reaction*, PCR).

### 1. Principe de la PCR

La PCR est une technique de la biologie moléculaire décrite par Kary Mullis en 1986 (Mullis et coll., 1986). C'est une technique permettant de générer plusieurs copies d'une région spécifique de la séquence d'ADN dans le but de la détecter et de l'étudier. Le principe est surtout basé sur une variation de température permettant la dénaturation de l'ADN, l'hybridation des amorces et l'amplification proprement dite.

#### ➤ Cycle de température

Une réaction est composée par plusieurs cycles comportant chacun 3 étapes (Figure 4) qui peut être répétées n fois ; n étant le nombre de cycles et donnant ainsi à la fin de la réaction  $2^n$  séquences.



**Figure 4:** Etapes d'un cycle PCR (<http://slideplayer.fr/slide/2648864/>).

La dénaturation est la première étape durant laquelle l'ADN double brin donne deux ADN simples brins. Ceci est dû à la rupture des liaisons hydrogènes entre les bases à une haute température qui est généralement de 95°C.

L'hybridation correspond à la fixation de l'amorce sur l'extrémité 5' de l'ADN cible dont elle est complémentaire. La spécificité de la réaction repose sur la qualité de cette étape. La température d'hybridation varie suivant le type d'amorce utilisé et

est principalement fonction de la température de fusion ou *melting temperature* (Tm). Cette Tm peut être calculée suivant la formule (Lazur, 1993):

$$Tm = 2^{\circ}C \times (\text{nombre de base A et T}) + 4^{\circ}C \times (\text{nombre de base C et G})$$

A, T, C et G représentent les bases.

La température d'hybridation est généralement inférieure d'environ 5°C à cette Tm calculée. Mais cette température d'hybridation dépend aussi d'autres facteurs tels que le pH ou la longueur des amorces. Ainsi, sa détermination se fait surtout lors des mises au point au laboratoire.

La dernière étape qui est l'élongation ou l'extension consiste à la formation de la séquence complémentaire par polymérisation des dNTPs. Cette étape nécessite la présence d'une enzyme : l'ADN polymérase appelée aussi la Taq polymérase. Elle se déroule généralement à une température égale à 72°C qui est la température optimale pour cette enzyme.

### ➤ Le mélange réactionnel ou mix

Le mélange est composé de plusieurs éléments dont :

- L'ADN matriciel : c'est le fragment d'ADN contenant la région cible et elle sert de matrice pour la formation de copies de cette région.

- Les amorces ou primers : ce sont des séquences composées de 10 à 30 nucléotides et sont complémentaires de la région à amplifier. Ils sont de 2 types les amorces *forward* et *reverse*.

- Les dNTPs sont composés du mélange de : dATP, dGTP, dCTP et dTTP qui sont les unités utilisées par la Taq polymérase lors de l'élongation.

- La Taq polymérase est une enzyme catalysant l'addition de dNTPs lors de l'élongation. Elle provient d'une bactérie *Thermus aquaticus* qui est résistante à une température de plus de 100°C. Elle a une activité 5' vers 3' et son activité maximale est atteinte à 72°C.

- La solution tampon permet de maintenir le pH stable. Elle contient plusieurs composants comme le Mg<sup>2+</sup> qui est un cofacteur de l'ADN polymérase.

## **2. *Les réactifs et matériels utilisés***

### **i. Les réactifs utilisés**

- *One Taq® 2X Master mix* (BioLabs, New England) un kit contenant les différents réactifs nécessaires pour la réaction PCR (Annexe 2)
- Eau stérile
- Sérum albumine bovine ou BSA (Sigma-Aldrich) 5%
- Marqueur de taille (Promega) : permet de connaître la taille des amplicons obtenus
- Bleu de charge (Promega) : permet de suivre la migration durant l'électrophorèse
- Bromure d'ethidium ou BET

### **ii. Les marqueurs mis au point par PCR**

Lors de cette étude, trois marqueurs mitochondriaux (COI, COI-tailed et COII) et un marqueur nucléaire (28S) (Sigma-Aldrich) ont été mis au point (Tableau 1). Le marqueur COI-tailed est composé d'un mélange d'amorces COI à queue M13 permettant d'amplifier une même région au niveau du gène codant pour le cytochrome oxydase 1 (COI). Elles sont utilisées en alternatives aux amorces COI classiques qui sont susceptibles de former des dimères pouvant être incorporées dans la réaction de séquençage. Ces dimères obstruent les 30-40 premières bases cibles de l'extrémité 5' des séquences (Ivanova et coll., 2007).

Le marqueur COII amplifie une région au niveau du gène codant pour la sous unité II du cytochrome oxydase. C'est un marqueur mitochondrial très utilisé en phylogénie et phylogéographie.

Le marqueur nucléaire 28S amplifie une région très conservée du gène codant l'ARN 28S et il est approprié pour faire une étude en phylogénie/phylogéographie (Hillis at Dixon, 1991, Mallatt et Sullivan, 1998).

**Tableau 1:** Liste des marqueurs mis au point

Marqueurs	Séquences 5'-3'	Tailles attendues	Références
		(pbs)	
<b>COI</b>	Forward : GGTCAACAAATCATAAAGATATTGG  Reverse : TAAACTTCAGGGTGACCAAAAAATCA	700	Folmer et coll., 1994
<b>COI-tailed</b>	Forward : - TGTAACGACGCCAGTTCAA CWAATCATAAAGATATTGG - TGTAACGACGCCAGTTCAA CTAACATARGATATYGG  Reverse : - CAGGAAACAGCTATGACTAAACTT CWGGRTGWCCAARAATCA - CAGGAAACAGCTATGACTAACYT CAGGATGACCAAAAAAYCA - CAGGAAACAGCTATGACTAACYT CDGGATGBCCAARAATCA	700	Données non publiées
<b>COII</b>	Forward : TCTAATATGGCAGATTAGTGC  Reverse : GAGACCAGTACTGCTTCAGTCATC	770	Whiting, 2002
<b>28S</b>	Forward : GACCCGTCTGAAGCACG  Reverse : GACTCCCTTACCTACAT	1470	Whiting, 2002

pbs : paires de bases

### iii. Les matériels utilisés

- Thermocycleur (Arktik Thermoscientific) : appareil permettant de faire varier la température suivant les conditions requises pour la réaction PCR.
- Centrifugeur Rotina 38 R (Hettich)
- Vortex (Fisher Scientific)
- Micropipettes

- Tubes individuels de 0,2 ml pour la PCR (ABgene)
- Cuve à électrophorèse
- *Gelscan* (Biorad laboratories)

### **3. Mise au point par PCR du marqueur COI**

Pour la PCR avec le marqueur COI, le tableau 2 suivant résume le volume de chaque composant du mélange réactionnel.

**Tableau 2:** Mélange réactionnel pour l'amplification avec le marqueur COI

Composants	Volumes (concentrations)
<b>Extrait d'ADN</b>	2µl ( 84,13 ng/µl)
<b>One Taq Master mix</b>	12,5µl
<b>Amorce forward</b>	1µl (10µM)
<b>Amorce reverse</b>	1µl (10µM)
<b>Eau</b>	8,5µl
<b>Volume total</b>	<b>25 µl</b>

Les conditions thermiques d'amplification PCR, faisant varier les température d'hybridation se résument comme suit (Tableau 3) :

**Tableau 3:** Conditions thermiques d'amplification pour la mise au point du marqueur COI

Etapes	Témpératures, durées	Cycles
<b>Dénaturation initiale</b>	95°C, 5 minutes	-
<b>Dénaturation</b>	95°C, 1 minute	
<b>Hybridation</b>	58 – 68°C, 1 minute	30
<b>Elongation</b>	72°C, 1 minute	
<b>Elongation finale</b>	72°C, 8 minutes	-

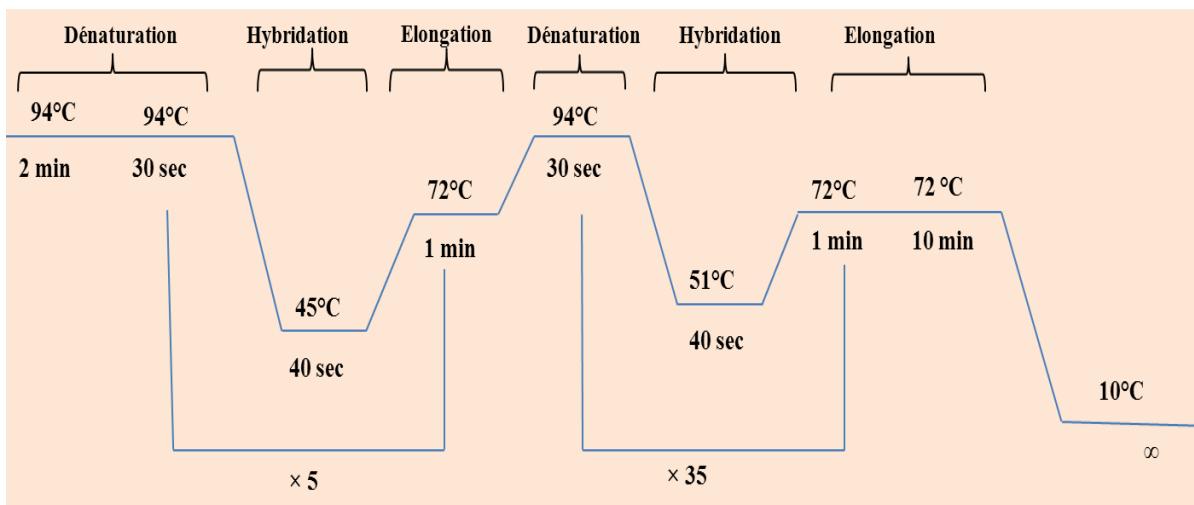
### **4. Mise au point par PCR du marqueur COI-tailed**

Les composants du mélange réactionnel sont présentés dans le tableau 4.

**Tableau 4:** Mélange réactionnel pour l'amplification avec le marqueur COI-tailed

Composants	Volumes (concentrations)
Extrait d'ADN	2µl ( 20,69 ng/µl)
One Taq Master mix	12,5µl
Amorces <i>forward</i>	0,4µl (10µM) par amorce
Amorces <i>reverse</i>	0,4µl (10µM) par amorce
Eau	8,5µl
Volume total	<b>25 µl</b>

Les conditions thermiques pour l'amplification sont composées de 2 cycles différents (Figure 5) afin de permettre aux amorces de mieux se fixer à l'ADN matriciel. Les amorces ainsi que les conditions de fonctionnement ont été acquises grâce à une collaboration avec une collègue de l'INRA Montpellier, le Dr Armelle CŒUR D'ACIER.



**Figure 5:** Cycle de température pour l'amplification avec le marqueur COI-tailed.

### 5. Mise au point par PCR du marqueur COII

Différentes conditions PCR ont été testées pour ce marqueur. Le tableau 5 suivant résume le volume de chaque composant du mélange réactionnel.

**Tableau 5:** Mélange réactionnel pour l'amplification avec le marqueur COII

Composants	Volumes (concentrations)
Extrait d'ADN	2µl ( 76,85ng/µl)
One Taq Master mix	12,5µl
Amorce <i>forward</i>	1µl (10µM)
Amorce <i>reverse</i>	1µl (10µM)
Eau	8,5µl
Volume total	25 µl

Le tableau 6 suivant présente les conditions thermiques pour l'amplification par PCR.

**Tableau 6:** Conditions thermiques d'amplification pour la mise au point du marqueur COII

Etapes	Témpératures, durées	Cycles
Dénaturation initiale	95°C, 5 minutes	-
Dénaturation	95°C, 1 minute	
Hybridation	53 – 64°C, 1 minute	30
Elongation	72°C, 1 minute	
Elongation finale	72°C, 8 minutes	-

Un ajout de 0,2µl de BSA (5%) a été effectué dans le mélange réactionnel pour certaine condition (Température d'hybridation = 62°C) dans le but d'éviter les hybridations non spécifiques.

## 6. *Mise au point par PCR du marqueur 28S*

Le tableau 7 suivant résume le volume pour chaque composant du mélange réactionnel utilisé lors de la mise au point.

**Tableau 7:** Mélange réactionnel pour l'amplification avec le marqueur 28S

Composants	Volumes (concentrations)
<b>Extrait d'ADN</b>	2µl ( 76,85 ng/µl)
<b>One Taq Master mix</b>	12,5µl
<b>Amorce forward</b>	1µl (10µM)
<b>Amorce reverse</b>	1µl (10µM)
<b>Eau</b>	8,5µl
<b>Volume total</b>	<b>25 µl</b>

Les conditions de température testées lors de l'amplification par PCR en faisant varier la température d'hybridation sont présentées dans le tableau 8 suivant.

**Tableau 8:** Conditions thermiques d'amplification pour la mise au point du marqueur 28S

Etapes	Témpératures, durées	Cycles
<b>Dénaturation initiale</b>	95°C, 5 minutes	-
<b>Dénaturation</b>	95°C, 1 minute	
<b>Hybridation</b>	52 – 62°C, 1 minute	30
<b>Elongation</b>	72°C, 2 minutes	
<b>Elongation finale</b>	72°C, 8 minutes	-

### 7. Détection des amplifiats PCR

La détection des amplifiats a été réalisée en faisant migrer les bandes par électrophorèse sur gel d'agarose puis, en révélant les bandes migrées par fluorescence à la lumière ultra-violet (UV).

Le gel d'agarose est un réseau complexe de polymères avec des pores permettant la séparation des molécules. L'électrophorèse sur ce gel permet de séparer les fragments d'ADN suivant leur taille selon le principe que les plus petites molécules migrent plus rapidement et plus loin que les grosses molécules. La molécule d'ADN est chargée négativement et sa soumission à un courant électrique entraîne sa migration vers le pôle positif. Dans toutes les manipulations, un gel d'agarose 1,5% a été utilisé (voir Annexe 3 pour la préparation). Quand le gel devient solide et est prêt à l'emploi, un dépôt d'amplicon de 5µl a été effectué pour

chaque échantillon (dont 2 µl de bleu de charge ajoutés au préalable) dans le puit du gel. Pour les marqueurs de taille, un dépôt de 2µl a été réalisé dans les puits situés à chaque extrémité du gel permettant ainsi de déterminer la taille des bandes pendant la visualisation à l'UV. La migration sur gel a été effectuée à 150V, 150mA pendant 1h30.

La visualisation des bandes au rayonnement UV a été effectuée à l'aide de l'appareil *Gelscan* et en utilisant le logiciel Quantity one (Biorad laboratories).

## **E. Etudes préliminaires en phylogéographie de *X. cheopis* à Madagascar**

Après les mises au point PCR des différents marqueurs, une étude préliminaire en phylogéographie a été effectuée. Ce genre d'étude nécessite l'acquisition des séquences nucléotidiques issues des différentes régions testées qui serviront de jeux de données pour les analyses bioinformatiques.

L'étude de la diversité génétique et phylogéographique de *X. cheopis* a été faite suivant 3 étapes : l'amplification PCR, le séquençage et le traitement de données de séquences.

### **1. Principe du séquençage d'ADN**

C'est un procédé permettant de déterminer l'agencement des acides nucléiques dans une séquence d'ADN. Deux techniques peuvent être utilisées pour ce séquençage : celle de Sanger (Sanger et coll., 1977) et la technique de Maxam et Gilbert (Maxam & Gilbert, 1977).

La méthode de Maxam et Gilbert découverte en 1977 est une méthode de séquençage chimique (Maxam & Gilbert, 1977). Elle présente des inconvénients du fait que les réactifs utilisés peuvent produire des effets néfastes sur la santé.

Le séquençage par la méthode de Sanger (1977) qui est une technique de séquençage enzymatique a été utilisé lors de cette étude. La méthode repose sur l'allongement par l'ADN polymérase d'un brin complémentaire à partir d'une amorce en utilisant un brin d'ADN matriciel. Elle est réalisée en présence de dNTPs et de son analogue qui est le didésoxyribonucléoside triphosphate (ou ddNTP)

jouant le rôle de terminateur de la chaîne et empêchant ainsi l'élongation de se poursuivre. Cela donne alors un mélange de fragments de diverses tailles se terminant par des ddNTPs permettant ainsi de déterminer la position de chaque base dans la séquence. Ces fragments sont ensuite visualisés par électrophorèse sur gel de polyacrylamide ou sur un séquenceur capillaire (Mauger, 2012 ; Sanger et coll., 1977).

## 2. Amplification avec les marqueurs COI-tailed

Parmi les marqueurs mis au point, nous avons choisi d'utiliser ce marqueur COI-tailed pour plusieurs raisons. Premièrement, ce marqueur est un mélange de plusieurs amores qui ciblent la même région du COI. Son utilisation permet d'anticiper les variations nucléotidiques possibles dans cette région et par conséquent, assure l'obtention d'amplifiats pour le séquençage. Deuxièmement, ce marqueur étant actuellement testé par d'autres collaborateurs internationaux, son utilisation permet d'obtenir des données de séquences qui seront combinables avec celles des autres pays collaborateurs pour des études phylogéographiques plus poussées et à l'échelle mondiale. La composition du mélange réactionnel est présentée dans le tableau 9 suivant.

**Tableau 9:** Composition du mélange réactionnel pour l'amplification avec les marqueurs COI-tailed

Composants	Volumes (concentrations)
<b>Extrait d'ADN</b>	4 à 6µl
<b>One Taq Master mix</b>	25µl
<b>Amorce forward</b>	0,8µl (10µM) par amorce
<b>Amorce reverse</b>	0,8µl (10µM) par amorce
<b>Eau</b>	15 à 17µl
<b>Volume total</b>	<b>50 µl</b>

Le cycle de température utilisé est le cycle présenté dans le paragraphe D.4. A la fin de la PCR, 5µl d'amplicon ont été prélevés et déposés dans un gel d'agarose

pour la migration. Après la visualisation sous rayon UV, les amplicons présentant des bandes claires à 700 pbs ont été sélectionnés et envoyés pour faire le séquençage.

### **3. Séquençage**

Ainsi, 2 plaques de 96 échantillons par plaque (environ 40 $\mu$ l d'amplicon par échantillon) ont été envoyées pour séquençage (par capillarité de Sanger) au laboratoire Macrogen Inc. (Corée) en utilisant les amorces universelles M13F/R.

### **4. Traitement bioinformatique des données**

Le séquençage a permis d'obtenir des séquences nucléotidiques qui ont été utilisées pour les analyses bioinformatiques.

➤ Le nettoyage et l'assemblage des séquences ont été réalisés en utilisant le logiciel STADEN PACKAGE v2.0.0b (Medical Research Council Laboratory of Molecular Biology, Cambridge, UK). Ce logiciel comporte 5 composants (Prégap4, Gap4, Spin, Trev et Gap5). Pour notre étude, nous avons utilisé les composants Prégap4 et Gap4. Prégap4 a permis de changer le format des fichiers « .abi » obtenu après séquençage en format « .ztr » (le format « ztr » est le format pris en charge par Gap4) et de faire le nettoyage des séquences. Après cette étape, les séquences *reverse* et *forward* ont été assemblées et éditées avec Gap4 afin d'obtenir le contigs ou *consensus* c'est-à-dire la séquence résultant de l'assemblage des 2 séquences *reverse* et *forward*.

La séquence partielle de COI obtenue dans la banque de données GenBank (<https://www.ncbi.nlm.nih.gov/genbank>) dont le numéro d'acquisition est KF437545.1 a été utilisée comme séquence référente.

➤ L'alignement des séquences a été effectué en utilisant le logiciel MEGA 6 (Tamura et coll., 2013) après obtention des contigs. L'algorithme MUSCLE a permis de faire l'alignement des séquences en utilisant les options par défaut. Quand nécessaire, la vérification manuelle a été effectuée.

➤ Ce même logiciel MEGA 6 a permis de déterminer la distance génétique ou la divergence génétique entre 2 individus après l'alignement. L'identification des haplotypes ou l'ensemble des variations de l'ADN ou

polymorphisme présent dans les populations a été réalisée avec le logiciel DNASP v5 (Librado et Rozas, 2009). Pour chaque site, les différents haplotypes présents ont été sélectionnés pour la mesure de la distance génétique.

## IV. RESULTATS

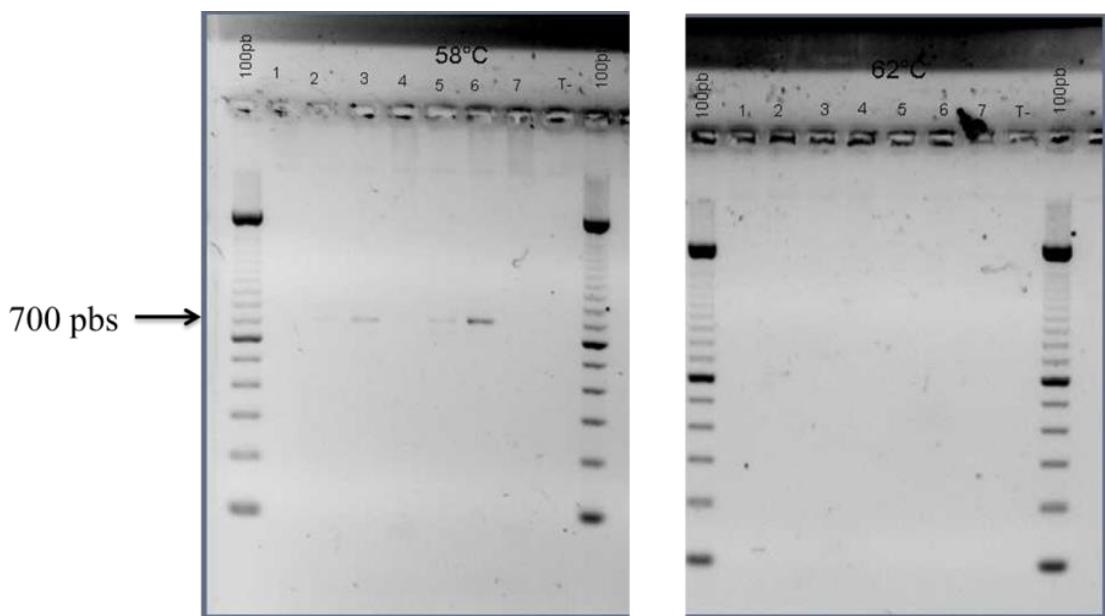
### A. Extraction d'ADN

L'extraction d'ADN par broyage en présence de Chelex résine 100 à 5%, a permis d'obtenir en moyenne 80 $\mu$ l d'ADN dont la concentration est variable allant de 1,6ng/ $\mu$ l à 80,25ng/ $\mu$ l avec une moyenne de 14,51ng/ $\mu$ l (Annexe 4).

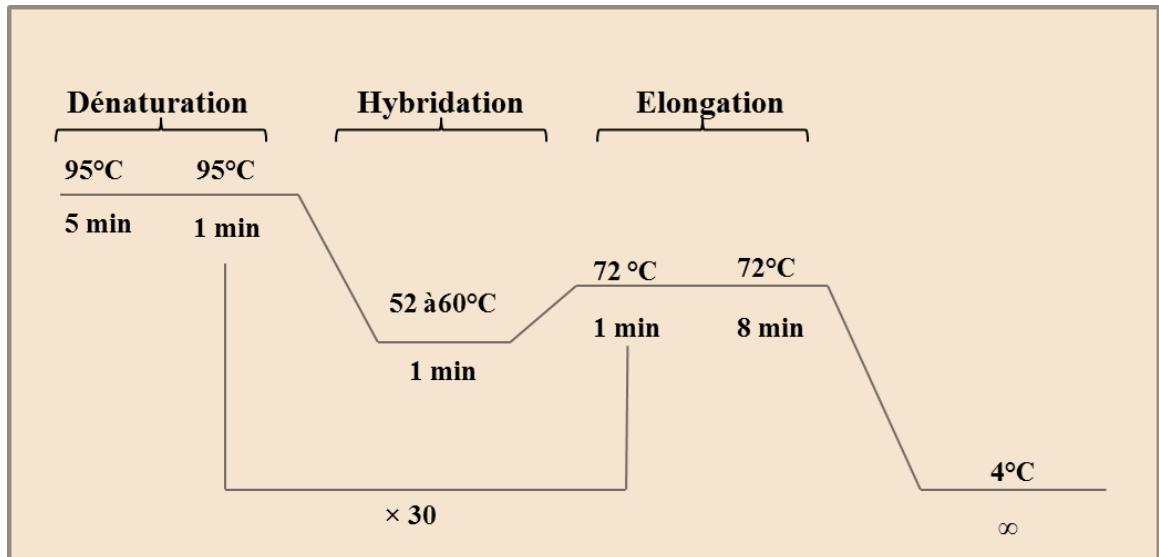
### B. Les conditions optimales pour les différents marqueurs

#### 1. *Mise au point par PCR du marqueur COI*

La lecture des gels après électrophorèse a permis d'observer la présence de bandes vers 700 pbs à des températures d'hybridation comprises entre 52 et 60°C avec une bande forte à 60°C. La figure 6 ci-dessous montre le profil du gel obtenu à température d'hybridation égale à 58°C. Mais à des températures d'hybridation élevées, aucune bande n'est observée. La figure 7 montre les conditions avec lesquelles des bandes ont été obtenues.



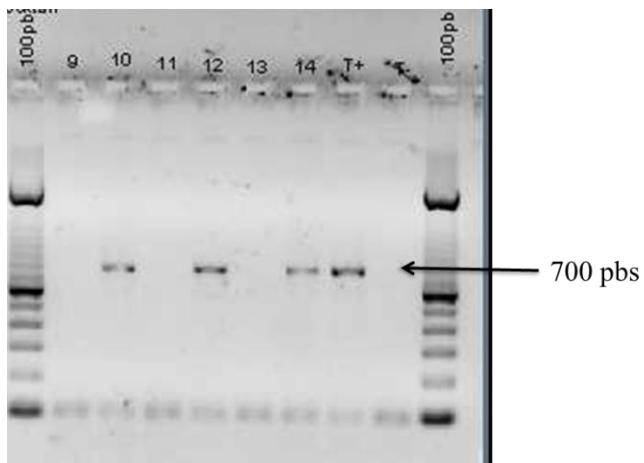
**Figure 6:** Profil de l'électrophorèse après amplification avec le marqueur COI. Température d'hybridation égale à 58°C (gauche) et à 62°C (droite). 100pb : marqueur de taille, 1 à 7 : numéros des extraits d'ADN de puce testés, T- : témoin négatif.



**Figure 7:** Cycle de température optimale d'amplification avec le marqueur COI.

### 2. Mise au point par PCR du marqueur COI-tailed

Après la lecture du gel contenant les amplicons obtenus avec le marqueur COI-tailed, des bandes vers 700 pbs sont observées (Figure 8). Ainsi, la condition d'amplification contenant 2 cycles est gardée pour les PCR avec ce marqueur.



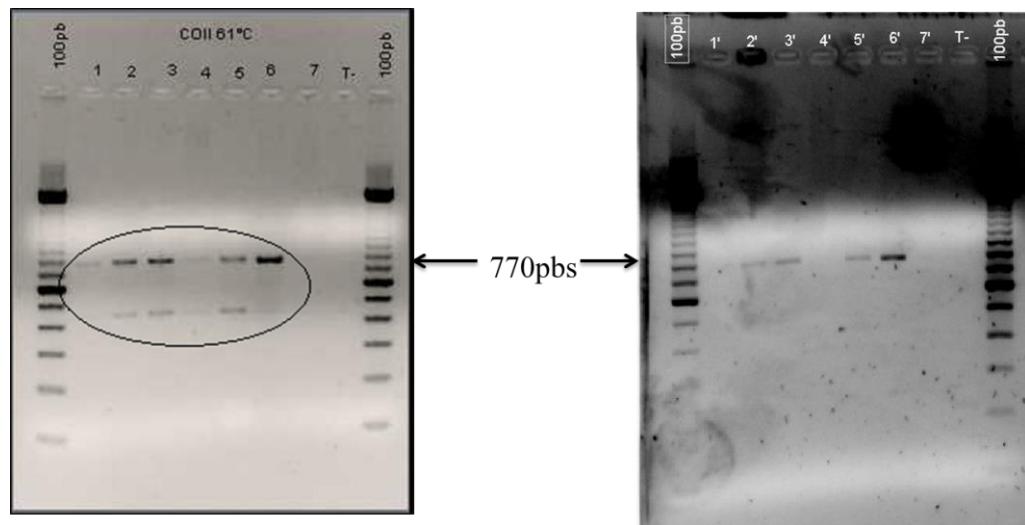
**Figure 8:** Profil de l'électrophorèse après amplification avec le marqueur COI-tailed.

100pb : marqueur de taille, 9 à 14 : numéros des extraits d'ADN de puce testés, T+ : témoin positif, T- : témoin négatif.

### 3. Mise au point par PCR du marqueur COII

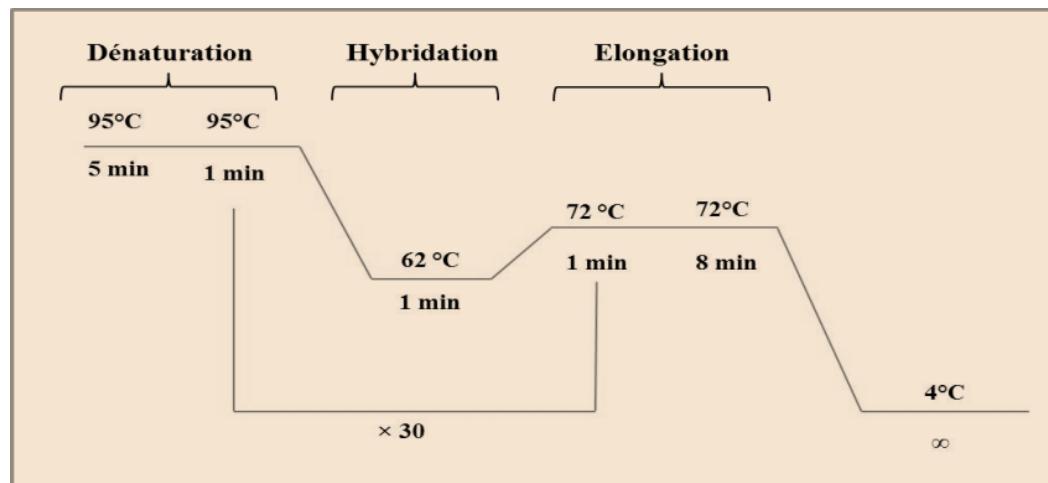
Après la lecture des gels contenant les amplicons obtenus avec les différentes températures d'hybridation, il est observé qu'à une température d'hybridation basse

(inférieur ou égale à 61°C), le gel présente des bandes vers 770 pbs et des bandes non spécifiques à environ 400 pbs (Figure 9 à gauche). Ces bandes parasites disparaissent à une température d'hybridation égale à 62°C (Figure 9 à droite). L'addition du BSA n'améliore pas ces résultats, on constate même une diminution du nombre des résultats positifs.



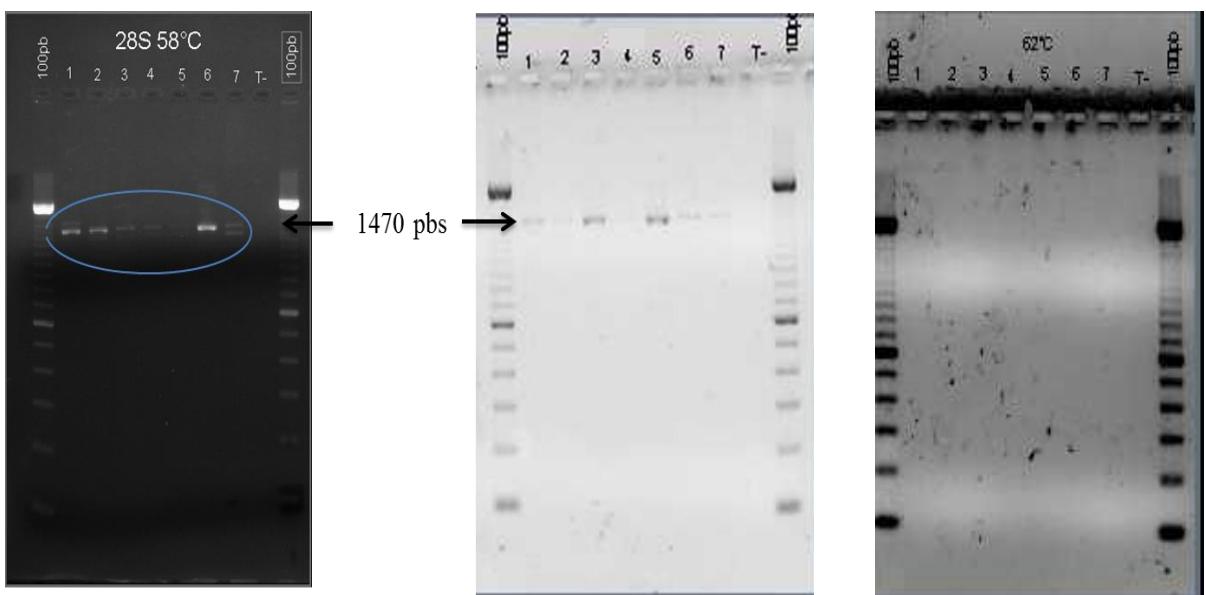
**Figure 9:** Profils de l'électrophorèse après amplification avec le marqueur COII. Température d'hybridation égale à 61°C (à gauche) et à 62°C (à droite). 1 à 7 et 1' à 7' : numéros des extraits d'ADN de puce testés, T- : témoin négatif.

Ces résultats suggèrent que la température d'hybridation optimale pour le marqueur COII est donc 62°C. La figure suivante (Figure 10) montre le cycle de température optimale pour le marqueur COII.



#### 4. Mise au point par PCR du marqueur 28S

A des températures d'hybridation comprises entre 52 et 60°C, des bandes sont observées aux environs de 1470 pbs. Des bandes aspécifiques ont été observées pour certaines températures d'hybridation (56 et 58 °C) (Figure 11). A la température d'hybridation égale à 59°C des bandes bien distinctes sont observées à 1470pb et à 60°C des bandes sont aussi observées mais elles sont floues. Quand la température d'hybridation est augmentée à 62°C, aucune bande n'est observée. Ce qui signifie que la région cible n'a pas pu être amplifiée à haute température.

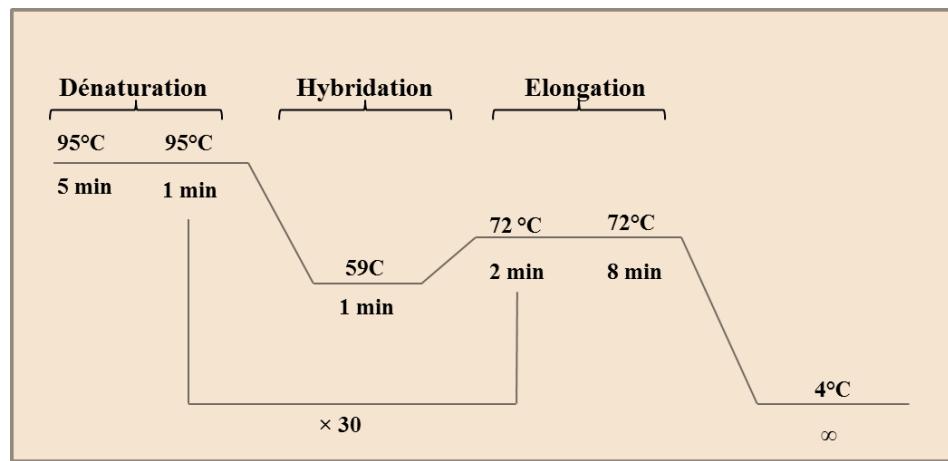


**Figure 11:** Profil de l'électrophorèse après amplification avec le marqueur 28S.

Température d'hybridation égale à 58°C (gauche), à 59°C (milieu) et à 62°C (droite).

1 à 7 : numéros des extraits d'ADN de puce testés, T- : témoin négatif.

Le cycle de température donnant une amplification optimale pour le marqueur 28S est résumé dans la figure suivante (Figure 12).



**Figure 12:** Cycle de température optimale d'amplification avec le marqueur 28S.

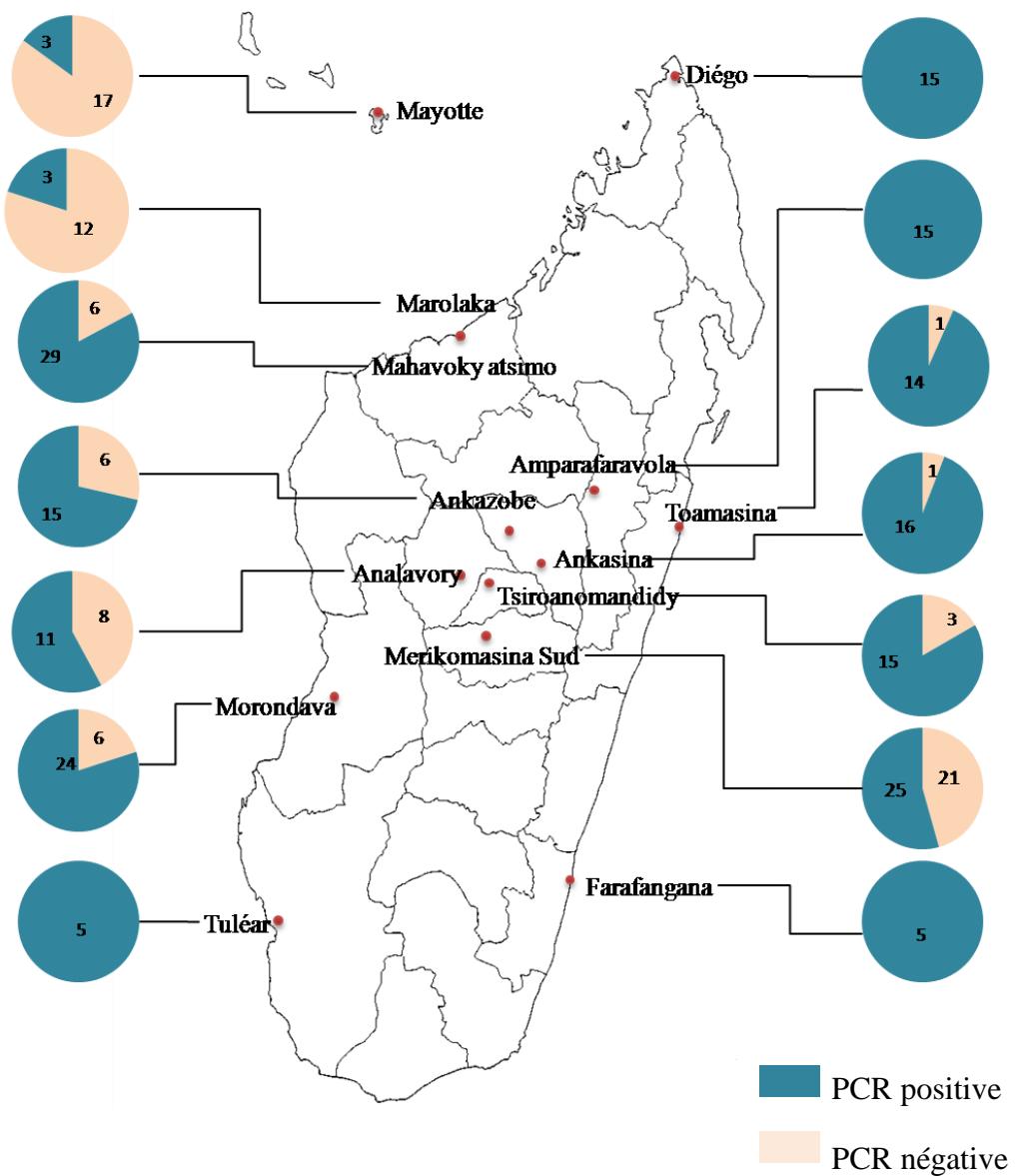
### C. Diversité génétique et phylogéographie de *X. cheopis*

#### 1. Amplification avec le marqueur COI-tailed

Sur les 24 sites étudiés, seul 14 ont donné des résultats positifs. Le nombre de résultats positifs et négatifs pour ces 14 sites est donné sur la figure 13. Les extraits d'ADN provenant des sites Antohomadinika, Analakely, Antananarivo, Andapa, Zazafotsy, Antanamalaza, Hiazofotsy, Ivory, Maurice et Comores n'ont pu être amplifiés.

Du fait du peu de résultats positifs pour les échantillons du site Mahajanga Marolaka ( $N=3$ ), ils n'ont pas été envoyés à séquencer. Ainsi, les amplicons des échantillons provenant de 12 sites situés à Madagascar ont été envoyés pour le séquençage. Pour les sites situés hors de Madagascar, seul Mayotte a présenté des résultats PCR positifs bien que le nombre soit faible ( $N= 3$ ), ces amplicons ont été envoyés à séquencer. Au total, 192 échantillons ont été envoyés au séquençage.

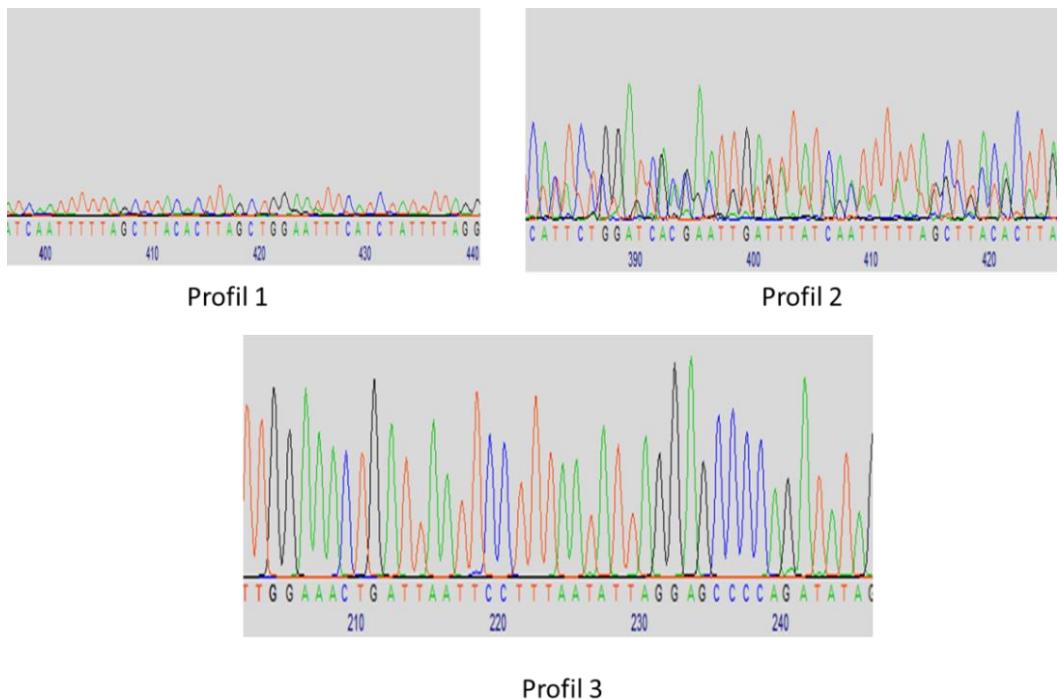
En outre, l'obtention des amplicons après PCR traduit le fait que le marqueur COI-tailed, conçu pour amplifier les *X. cheopis* de l'Afrique de l'Ouest a permis d'amplifier les *X. cheopis* de Madagascar.



**Figure 13:** Résultats de la PCR avec COI-tailed par site.

## 2. Profils des électrophorégrammes après séquençage

Les résultats reçus sont présentés en 4 formats : .abi, .phd, .pdf et .txt. Le fichier .pdf contient l'électrophorégramme (Figure 15) dont les pics de différentes couleurs représentent les bases (rouge pour T, verte pour A, bleu pour C et noir pour G). L'observation des électrophorégrammes a conduit à l'identification de 3 profils différents (Figure 14).



**Figure 14:** Profils des électrophorégrammes après séquençage.

Le profil 1 montre des pics faibles qui pourraient être dus à de faibles quantités d'ADN. Le profil 2 montre des pics qui se chevauchent qui pourraient être traduits par le fait que l'ADN n'est pas pur ou qu'il y a eu mélange des extraits. Quant au profil 3, les pics observés sont bien visibles et bien individualisés signifiant que l'ADN est pur et que sa concentration est optimale pour le séquençage.

### 3. Traitement bioinformatique des données

- **Obtention de contigs après assemblage des séquences *forward* et *reverse***

L'assemblage des séquences *forward* et *reverse* a permis d'obtenir 152 contigs sur 192 échantillons envoyés. La taille des contigs est située entre 599 et 659 pbs. Certaines séquences sont de mauvaise qualité, ce qui n'a pas permis l'assemblage.

Les 152 contigs sont formés par 151 séquences de Madagascar et une séquence de Mayotte. Ainsi, les 151 séquences ont été alignées pour travailler sur la phylogéographie de Madagascar.

### ➤ Variabilité génétique à Madagascar

Après alignement des 151 contigs obtenus, l'identification des haplotypes a montré la présence de 19 haplotypes pour *X. cheopis* à Madagascar. La répartition des haplotypes H1, H2 jusqu'à H19 dans les sites est représentée dans le tableau suivant (Tableau 10).

Il est observé que l'haplotype H2 est présent dans tous les sites sauf à Farafangana. Le plus grand nombre d'haplotypes a été retrouvé dans la population d'Analavory.

**Tableau 10:** Distribution des haplotypes par site

Site	Nombre d'individu par site	Haplotype (H)
Majunga	27	H1-H2
Ankasina	12	H2-H3
Amparafaravola	11	H2-H4
Ankazobe	15	H2-H5-H6-H7
Tsiroanomandidy	10	H2-H8-H9
Tuléar	3	H2-H10
Farafangana	1	H11
Analavory	8	H2-H12-H13-H14-H15-H16
Toamasina	12	H2-H17
Diégo	15	H2-H18
Morondava	18	H2-H19
Merikomasina	19	H2-H8

Il est aussi constaté que l'haplotype H2 regroupe le plus grand nombre d'individus avec 128 échantillons tandis que les autres haplotypes ne regroupent qu'un à 4 échantillons (Tableau 11). Ainsi l'haplotype H2 est l'haplotype majeur.

**Tableau 11:** Nombre d'individus par haplotype

Haplotype	Nombre d'individu
H1	1
H2	128
H3	1
H4	1
H5	1
H6	1
H7	1
H8	2
H9	2
H10	1
H11	1
H12	1
H13	1
H14	1
H15	1
H16	1
H17	1
H18	1
H19	4

Dans chaque site, un échantillon par haplotype qui y est présent a été choisi pour la mesure de la distance génétique. Elle montre des valeurs situées entre 0,00 et 0,34 (Tableau 12). La similarité entre les échantillons est donc comprise entre 66% et 100% dont la majorité tend vers 100%. Il est observé que l'individu n°347

d'Ankazobe et le n°366 de Tsiroanomandidy sont les plus différents avec une distance génétique égale à 0,34. Il est constaté que les individus n°700 de Majunga, n°308 d'Ankasina, n°323 d'Amparafaravola, n°436 de Tuléar, n°446 de Farafangana, n°688 de Morondava et n°737 de Merikomasina sont similaires à 100% (distance génétique égale à 0). Cette forte similarité nucléotidique est observée dans la plupart des individus provenant des différents sites. Bien que certains échantillons appartiennent à des sites éloignés l'un de l'autre, ils ont été trouvés fortement similaires. C'est le cas de deux individus en provenance de Diego (n°664) et de Tsiroanomandidy (n°369) avec une valeur de similarité égale à 100%. Pour d'autres individus, ils sont géographiquement proches mais très distants génétiquement. C'est le cas de l'individu en provenance d'Analavory (n°503) et de celui de Tsiroanomandidy (n°366) avec une valeur de similarité égale à 70%.

**Tableau 12:** Distance génétique (DG) entre les représentants de chaque haplotype par site

DG	1		2		3		4				5			6		7		8						9		10		11				12			
N° échantillon	508	700	308	318	323	324	346	347	355	356	365	366	369	435	436	446	492	495	497	499	503	505	630	632	654	664	671	685	687	688	689	725	737		
1	508	0,00																																	
	700	0,03	0,00																																
2	308	0,03	0,00	0,00																															
	318	0,04	0,00	0,00	0,00																														
3	323	0,03	0,00	0,00	0,00	0,00																													
	324	0,04	0,01	0,01	0,01	0,01	0,00																												
4	346	0,03	0,00	0,00	0,00	0,00	0,01	0,00																											
	347	0,17	0,15	0,15	0,15	0,15	0,15	0,00																											
	355	0,04	0,00	0,00	0,00	0,00	0,01	0,00	0,15	0,00																									
	356	0,04	0,00	0,00	0,01	0,00	0,01	0,00	0,15	0,00	0,00																								
5	365	0,32	0,29	0,29	0,29	0,29	0,29	0,29	0,33	0,29	0,29	0,00																							
	366	0,34	0,30	0,30	0,30	0,30	0,30	0,30	0,34	0,30	0,30	0,03	0,00																						
	369	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00																						
6	435	0,04	0,00	0,00	0,01	0,00	0,01	0,00	0,15	0,01	0,01	0,29	0,30	0,00	0,00																				
	436	0,04	0,00	0,00	0,00	0,01	0,00	0,00	0,15	0,01	0,01	0,29	0,30	0,00	0,00	0,00																			
7	446	0,04	0,00	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,01	0,29	0,30	0,00	0,00	0,00	0,00																		
8	492	0,04	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,29	0,00	0,00	0,00	0,00	0,00																		
	495	0,05	0,01	0,01	0,01	0,02	0,01	0,15	0,01	0,01	0,30	0,30	0,01	0,01	0,01	0,01	0,01	0,00																	
	497	0,04	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,28	0,29	0,00	0,00	0,00	0,00	0,00	0,01	0,00																
	499	0,04	0,01	0,01	0,01	0,01	0,01	0,16	0,01	0,01	0,29	0,29	0,01	0,01	0,01	0,01	0,01	0,01	0,00																
	503	0,04	0,01	0,01	0,01	0,01	0,01	0,15	0,01	0,01	0,29	0,30	0,01	0,01	0,01	0,01	0,01	0,01	0,00																
9	505	0,05	0,01	0,01	0,01	0,01	0,02	0,01	0,16	0,01	0,01	0,29	0,30	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,00														
	630	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,00												
10	632	0,04	0,01	0,01	0,01	0,01	0,01	0,15	0,01	0,01	0,29	0,30	0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,01	0,01	0,01	0,00											
	654	0,03	0,00	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00											
11	664	0,03	0,00	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,00	0,00											
	671	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00	0,00											
	685	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00	0,00											
	687	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00	0,00											
	688	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00	0,00											
12	689	0,03	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00	0,00											
	725	0,32	0,29	0,29	0,29	0,29	0,33	0,29	0,29	0,00	0,03	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29	0,29											
	737	0,03	0,00	0,00	0,00	0,00	0,01	0,00	0,15	0,00	0,00	0,29	0,30	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,01	0,01	0,01	0,00											

1 : Majunga ; 2 : Ankasina ; 3 : Amparafaravola ; 4 : Ankazobe ; 5 : Tsiroanomandidy ; 6 : Tuléar ; 7 : Farafangana ; 8 : Analavory ;

9 : Toamasina ; 10 : Diégo ; 11 : Morondava ; 12 : Merikomasina

## V. DISCUSSION

Les résultats obtenus pendant les mises au point des marqueurs ont permis de constater que la température d'hybridation des amores est une variable importante conditionnant la bonne mise en marche de l'amplification PCR (Lorenz, 2012 ; Wittwer et Garling, 1991). En effet, en diminuant cette température, les amores se fixent sur des régions non ciblées entraînant ainsi la présence de bandes indésirables. Une température élevée par rapport à la température d'hybridation optimale peut dégrader les amores empêchant ainsi leur fixation sur la séquence cible.

Bien que testés avec la même température d'hybridation, certains échantillons n'ont pas pu être amplifiés suggérant que d'autres variables sont à considérer pour avoir une amplification à la bonne taille. Une de ces variables est la qualité de l'ADN, un extrait de bonne qualité étant un extrait comportant une quantité d'ADN suffisante et des ADN non dégradés. Pour s'assurer de cette qualité, il est nécessaire d'utiliser des produits d'extraction qui ne contiennent pas trop d'inhibiteurs de PCR (inhibiteur de l'enzyme Taq polymérase) qui peut conduire à l'échec de la réaction (Orlando et Hänni, 2000). Il est aussi nécessaire de voir si la quantité d'ADN dans l'extrait est suffisante. Pour cela, il faut vérifier la présence de l'ADN dans l'extrait et quantifier sa concentration. Aussi, il est important de conserver les extraits d'ADN dans les conditions adéquates (-20°C) et faire des aliquotes de la solution d'ADN mère à petits volumes. Cela évite la dégradation et la possible contamination de l'extrait entier. Une autre explication de la non-amplification PCR pour certains échantillons serait aussi que les amores ne s'étaient pas fixées à la région cible (amores pas entièrement complémentaires à la séquence d'ADN).

Cette étude a permis d'obtenir les conditions optimales d'amplification pour chaque marqueur potentiellement convenable pour une étude phylogénétique et phylogéographique de *X. cheopis* à Madagascar. Rappelons qu'un marqueur convenable est un marqueur permettant de révéler le polymorphisme génétique et donnant ainsi des informations exploitables pour les analyses phylogéographiques. Une fois les conditions optimales d'amplification PCR identifiées pour ces différents marqueurs, le processus de séquençage peut être effectué en vue d'identifier les marqueurs donnant le plus de polymorphisme possible.

Le séquençage des amplicons a montré que le succès de cette réaction dépend de plusieurs facteurs tels que la pureté de l'ADN et sa concentration. Un mauvais séquençage rend difficile l'assemblage des séquences *reverse* et *forward* ou ne permet pas du tout de faire cet assemblage.

L'étude de la diversité génétique avec les amorces COI-tailed nous montre la présence de 19 haplotypes de *X. cheopis* dans les sites étudiés à Madagascar traduisant l'existence d'une diversité haplotypique au sein de l'espèce. Ces différents haplotypes sont répartis dans chaque site même si la plupart des échantillons sont regroupés dans l'haplotype 2 (Tableau 11). Pour certains sites, le nombre d'individus est faible (ex. Farafangana ; n=1), ce qui n'a pas permis de révéler le polymorphisme génétique présent dans ce site. Le fait que certains sites ayant un nombre d'échantillons élevé (comme Majunga, Merokomasina et Morondava) mais ne présentant que 2 haplotypes suggère que la diversité génétique y est assez faible. Il est aussi observé que le site Analavory présente le plus grand nombre d'haplotypes indiquant que ce site contient une grande diversité génétique intraspécifique. En résumé, *X. cheopis* varie génétiquement d'un site à un autre.

L'analyse de la distance génétique a suggéré que la majorité des individus sont génétiquement similaires à 100% bien qu'il y ait quand même une certaine diversité génétique. Aussi, il a été montré que certains échantillons appartenant à des sites éloignés de l'un à l'autre peuvent être fortement similaires. Pour d'autres individus, ils sont géographiquement proches mais très distants génétiquement. Ces variations de similarité en fonction du site géographique pourraient être associées au mouvement des hôtes ; à noter qu'étant ectoparasite, l'espèce *X. cheopis* (qui est une puce de fourrure) est étroitement liée à son hôte.

Cette étude montre que la portion du gène mitochondrial COI, largement utilisé en phylogéographie, peut révéler le polymorphisme génétique de *X. cheopis*. Aussi, il a été démontré que la diversité intraspécifique se répartit différemment selon les sites. Etant préliminaire, cette étude a donné des informations importantes pour des études plus approfondies sur la détermination du mode de dispersion de cet insecte vecteur à l'échelle nationale. La connaissance du mode de dispersion du vecteur de la maladie peste pourrait contribuer à déterminer le mode de propagation de cette maladie.

Une des perspectives de ce travail est de faire une étude de la diversité génétique de *X. cheopis* en utilisant les autres marqueurs qui ont été mis au point lors de cette étude. Il est alors possible de comparer ces informations nucléotidiques pour voir lesquels des marqueurs donnent des signaux plus polymorphes et aussi s'il y a cohérence, après analyses, entre les résultats des différents marqueurs. Il est possible également de faire une étude phylogéographique de cette espèce à l'échelle mondiale utilisant des spécimens en provenance de différents pays.

## **VI. CONCLUSION**

*X. cheopis* est une espèce de puce connue comme vectrice majeure de la maladie peste humaine. Ainsi, il est nécessaire tout d'abord de l'identifier, mais aussi d'étudier sa variabilité génétique et son mode de dispersion spatiale à Madagascar. Les objectifs de cette étude ont été de mettre au point les différents marqueurs moléculaires (COI, COI-tailed, COII et 28S) potentiellement convenables pour les études phylogéographiques, de déterminer la diversité intraspécifique de *X. cheopis* à Madagascar et de réaliser une étude préliminaire sur sa phylogéographie. La technique d'amplification PCR a été utilisée pour déterminer les conditions de température adéquates pour chaque marqueur. Ainsi, les différentes conditions optimales d'amplification PCR pour chaque marqueur ont été identifiées. Un des marqueurs, le COI-tailed a été choisi pour faire du séquençage partiel du gène mtCOI afin de déterminer le polymorphisme génétique pour chaque site géographique testé. Le séquençage utilisant le marqueur COI-tailed a permis d'identifier la présence d'une diversité haplotypique au sein des différentes populations géographiques de l'espèce. Dix-neuf haplotypes ont été identifiés. Le nombre et le type d'haplotypes présents varient en fonction du site. La comparaison par paires des séquences nucléotidiques a donné des valeurs de similarité entre 66% et 100%. Même si tous les marqueurs mis au point n'ont pas été utilisés pour l'étude préliminaire en phylogéographie, cette étude a démontré le fait que le marqueur COI-tailed est approprié pour une telle étude. Ces données préliminaires amorcent l'étude approfondie de la phylogéographie de cet insecte vecteur à l'échelle nationale.

## VII. REFERENCES BIBLIOGRAPHIQUES

- Andrianaivoarimanana, V., Kreppel, K., Elissa, N., Duplantier J.C., Carniel, E., Rajerison, M., Jambou, R., 2013. Understanding the persistence of plague foci in Madagascar. *PLoS Neglected Tropical Diseases*, 7(11), pp.1–8.
- Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lambt, T., Neigel, J.E., Reeb,C.A., Saunders, N.C., 1987. Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18(1), pp.489–522.
- Avise, J.C., 1994. Molecular markers, natural history and evolution. Chapman & Hall (Ed.), New York, 511 p.
- Avise, J.C., 1998. The history and purview of phylogeography. *Molecular Ecology*, (7), pp.371–379.
- Avise, J.C., 2000. Phylogeography the history and formation of species. Harvard University press (Ed.), Londres, 445 p.
- Aymes, N.D.E., 2009. Contribution à la phylogénie du genre *Stomoxys* (Diptera, Muscidae) et à la phylogéographie de *Stomoxys calcitrans* (L. 1758). [Thèse de Doctorat : Biologie des Populations et Ecologie]. Montpellier. Universite Paul Valery - Montpellier III , 146 p.
- Biochard, D., Le Roy, P., Levziel, H., Elsen, J.M., 1998. Utilisation des marqueurs moléculaires en génétique animale. *INRA Productions Animales*, 11(1), pp.67–80.
- Bitam, I., Dittmar, K., Parola, P., Whiting, M.F., Raoult, D., 2010. Fleas and flea-borne diseases. *International Journal of Infectious Diseases*, 14(8), pp.667–676.
- Chanteau, S., 2006. Atlas de la peste à Madagascar. IRD (Ed.), Paris, 94 p.
- Chanteau, S., Ratsifasoamanana, L., Rasoamanana, B., Rahalison, L., Randriambelosoa, J., Roux, J., Rabeson, D., 1998. Plague, a reemerging disease in Madagascar. *Emerging Infectious Diseases*, 4(1), pp.101–104.

Cuevas, J.M., Delaunay, A., Visser, J.C., Bellstedt, D.U., Jacquot, E., Elena, S.F., 2012. Phylogeography and molecular evolution of Potato virus Y. *PLOS One*. 7(5), pp 37853.

Delandes, H., 1997. Phylogénie moléculaire des huîtres du genre *Crassostrea* approchée par les techniques de restriction enzymatique et de séquençage. [Mémoire de Maîtrise: Biochimie]. Université des Sciences de la Rochelle, 37 p.

Duchemin, J.B., Fournier, P.E., Parola, P., 2006. Les puces et les maladies transmises à l'Homme. *Medecine Tropicale*, 66(1), pp.21-29.

Eisen, R.J., Gage, K.L., 2012. Transmission of flea-borne zoonotic agents. *Annual Review of Entomology*, 57, pp.61-82.

Emerson, B.C., Hewitt, G.M., 2005. Phylogeography. *Current biology*, 15(10), pp.367-371.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), pp.294-299.

FAO, 2008. L'état des ressources zoogénétiques pour l'alimentation et l'agriculture dans le monde. FAO (Ed.), Rome, 367 p.

Fotini, K., 2007. Biologie et phylogéographie de Monochamus Galloprovincialis (Coleoptera, Cerambycidae) vecteur du nématode du pin en Europe. [Thèse de Doctorat : Physiologie et Biologie des Organismes et des Populations–Interactions]. Orléans ; Université d'Orléans, 186 p.

Franc, M., 1994. Puces et méthodes de lutte. *Revue scientifique et technique (International Office of Epizootics)*, 13(4), pp.1019–1037.

Freeland, J.R., 2005. Molecular ecology. John Wiley & Sons (Ed.), Chichester, 388 p.

Gage, K.L., Kosoy, M.Y., 2005. Natural history of plague: perspectives from more than a century of research. *Annual Review of Entomology*, 50, pp.505-528.

Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology*, 66(4), pp.411–453.

Ivanova, N.V., Zemlak, T.S., Hanner, R.H., Hebert, P.D.N., 2007. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, 7(4), pp.544–548.

Lazur, D., 1993. Un procédé de réPLICATION in vitro Génie génétique. Lavoisier (Ed.), Paris, 388 p.

Librado, P., Rozas, J., 2009. DNSPA v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, pp.1451-1452.

Lorenz, T.C., 2012. Polymerase Chain Reaction: basic protocol plus troubleshooting and optimization strategies. *Journal of Visualized Experiments*, 63, pp.3998.

Lotfy, W.M., 2015. Plague in Egypt: disease biology, history and contemporary analysis: A minireview. *Journal of Advanced Research*, 6(4), pp.549–554.

Mallatt, J., Sullivan, J., 1998. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. *Molecular Biology and Evolution*, 15(12), pp.1706–1718.

Mauger, F., 2012. Développement de méthodes d'analyse de l'ADN par clivage d'une chimère ARN/ADN et par spectrométrie de masse MALDI-TOF. [Thèse de Doctorat : Chimie Moléculaire]. Université Pierre et Marie Curie, 345 p.

Maxam, A. M., Gilbert, W., 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(2), pp.560–564.

Miarinjara, A., Rogier, C., Harimalala, M., Ramihangihajason, T.R., Boyer, S., 2016. *Xenopsylla brasiliensis* fleas in plague focus areas, Madagascar. *Emerging Infectious Diseases*, 22(12), pp.2207–2208.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., Elrich, H., 1986. Specific enzymatic amplification of DNA in vitro: The Polymerase Chain Reaction. *Cold Spring Harbor Symposia on Quantitative Biology.*, 51, pp.263-273.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics 1st ed. Oxford University press (Ed.), New York, 333 p.

Newton, A.C., Allnutt, T.R., Gillies A.C.M., Lowe, A.J., Ennos, R.A., 1999. Molecular phylogeography, intraspecific variation and the conservation of tree species. *Tree*, 14(4), pp.140–145.

Orlando, L., Hänni, C., 2000. Du nouveau pour l'ADN ancien. *Société Française de Génétique*, 8-9, pp.1-16.

Rodhain, F., Perez, C., 1985. Précis d'entomologie médicale et vétérinaire. Maloine (Ed.), Paris, 458 p.

Rust, M.K., Dryden, M.W., 1997. The biology, ecology and management of the cat flea. *Annual Review of Entomology*, 42(1), pp.451–473.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp.5463–5467.

Silverman, J., Rust, M.K., 1985. Extended longevity of the pre-emerged adult cat flea (Siphonaptera, Pulicidae) and factors stimulating emergence from the pupal cocoon. *Annals of the Entomological Society of America*, 78(6), pp.763–768.

Stenseth, N.C., Atshabar, B.B., Begon, M., Belmain, S.R., Bertherat, E., Carniel, E., Gage, K.L., Leirs, H., Rahalison, L., 2008. Plague: past, present, and future. *Plos Medicine Neglected diseases*, 5(1), pp.9–13.

Swynghedauw, B., Silvestre, J.S., 2008. Biologie et génétique moléculaires. 3rd ed. Dunod (Ed.), Paris, 174 p.

Taberlet, P., Bouvet, J., 1994. Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Biological Sciences*, 255(1344), pp.195-200.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), pp.2725–2729.

Trubéa, F., Hernandez, E., Cavallo, J.D., 2007. La peste une zoonose ré-émergente selon l'Organisation Mondiale de la Santé. *Médecine et Armées*, 35(2), pp.127–134.

Van der Mescht, L., Matthee, S., Matthee, C.A., 2015. Comparative phylogeography between two generalist flea species reveal a complex interaction between parasite life history and host vicariance: parasite-host association matters. *BMC Evolutionary Biology*, 15, pp. 105-120.

Whiting, M.F., 2002. Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. *Zoological Scripta*, 31, pp. 93-104.

Whiting, M.F., Whiting, A.S., Hastriterb, M.W., Dittmar, K., 2008. A molecular phylogeny of fleas (Insecta: Siphonaptera): Origins and host associations. *Cladistics*, 24(5), pp.677–707.

Wittwer, C.T., Garling, D.J., 1991. Rapid cycle DNA amplification: Time and temperature optimization. *Biotechniques*, 10(1), pp.76–83.

World Health Organisation, 2010. Human plague: review of regional morbidity and mortality, 2004–2009. *Weekly Epidemiological Record*, 85(6), pp.40–48.

World Health Organisation, 2016. Plague around the world, 2010-2015. *The Weekly Epidemiological Record*, 91(8), pp.89–104.

Zhou, W., Russell, C.W., Johnson, K.L., Mortensen, R.D., Erickson, D.L., 2012. Gene expression analysis of *X. cheopis* (Siphonaptera: Pulicidae) suggests a role for reactive oxygen species in response to *Yersinia pestis* infection. *Journal of Medical Entomology*, 49(2), pp.364–370.

## **ANNEXES**

**Annexe 1 :** Liste des sites, leurs coordonnées GPS et le nombre d'échantillons par site

Pays	Sites	Latitude	longitude	Altitude (m)	Nombre d'échantillons
Mayotte	Village Longoni	-12,733300°	45,161767°	13	20
Comores	Mtsangani	-11,703333°	43,253333°	16	20
Maurice	Port Louis	-20,165000°	57,496667°	10	20
Madagascar	Analavory	-18,953333°	46,671667°	1059	19
	Ankasina	-18,900617°	47,505617°	1265	17
	Amparafaravola	-17,585967°	48,221050°	1880	15
	Ankazobe	-18,433950°	47,168200°	1257	21
	Tsiroanomandidy	-18,936067°	46,314317°	898	20
	Mahavoky atsimo	-15,706633°	46,317617°	7	45
	Tolinary	-23,354314°	43,667525°	8	5
	Farafangana	-22,816687°	47,816666°	9	5
	Toamasina	-18,123850°	49,382767°	14	15
	Diégo	-12,279400°	49,305600°	37	15
	Morondava	-20,296864°	44,295719°	8	30
	Merinkomasina sud	-19,651076°	46,6197247°	1228	46
	Antanamalaza	ND	ND	ND	18
	Analakely	-18,906667°	47,526667°	1276	22
	Marolaka	-15,725883°	46,317717°	7	15
	Antananarivo	ND	ND	ND	30
	Antohomadinika	-18,903333°	47,513333°	1250	15
	Hiazofootsy	-20,601667°	46,571667°	1630	15
	Zazafotsy	-22,205940°	46,363822°	762	20
	Andapa	-14,649993°	49,650138°	483	15
	Ivory	-21,566667°	46,916667°	516	30

ND : Non Determiné

**Annexe 2** : Composition du *One Taq® 2X master mix*

20 mM Tris-HCl (pH 8,9 à 25°C)

1,8 Mm MgCl<sub>2</sub>

22 mM NH<sub>4</sub>Cl

22mM KCl

0,2 mM dNTPs

5% glycérole

0,06% IGEPAL® CA-630

0,05% Tween® 20

25 units/ml *One Taq DNA Polymerase*

### **Annexe 3:** Préparation du gel d'agarose 1,5%

- Réactifs

- Poudre d'Agarose (Sigma-Aldrich)
- Tris Borate EDTA ou TBE 0,5X : un tampon ayant un pH égal à 8. il est utilisé comme tampon lors de l'électrophorèse.
- Bromure d'ethidium ou BET : un agent intercalant qui devient fluorescent lors de son exposition aux rayons ultra-violet (UV). Cette intensité augmente lorsqu'il est lié à un fragment d'ADN.
- Bleu de charge (Promega) : permet de suivre la migration durant l'électrophorèse.
- Marqueur de taille (Promega) : permet de connaître la taille des amplicons obtenus.

- Préparation du gel d'Agarose 1,5%

Quatre grammes et demi de poudre d'agarose et 300ml de TBE 0,5X ont été mis dans un Erlenmeyer. Le mélange a été chauffé à haute température durant 5 minutes afin que le gel fonde. Après la cuisson, le gel a été refroidi puis 10µl de BET a été ajouté. Le mélange a ensuite été versé dans une moule contenant une peigne et laissé se solidifier. Le gel a été ensuite mis dans un cuve pour migration contenant du TBE 0,5X dans lequel il est complètement immergé. Le peigne a été ensuite enlevé donnant ainsi les trous dans lesquels les amplicons vont être déposés.

**Annexe 4** : Concentration des extraits d'ADN

Site : Marolaka

Echantillon N°	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
Quantité d'ADN (ng/µl)	29,3	8,5	21	15,25	46,7	15,9	17,6	21,3	33,45	14,55	16,45	16,4	15,96	25,5	54,5

Site : Antohomadinika

Echantillon N°	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483
Quantité d'ADN (ng/µl)	3,65	4,3	4,9	7,4	11,5	6,9	10,55	4,5	4,1	29,35	8,05	8,55	16,85	12,85	12,45

Site : Analavory

Echantillon N°	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507
Quantité d'ADN (ng/µl)	2,45	3,1	3,7	5,3	2,35	5,1	1,95	2,4	7,33	2,35	6,15	3,76	3,05	8	4,45	2,33	4,85	4,75	7,9	31,35	5,95

Site : Mahavoky atsimo

Echantillon N°	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526
Quantité d'ADN (ng/µl)	5,75	11,1	9,9	4,25	10,35	9,05	19,15	6,95	9,5	5,15	5,75	5,15	3,15	5,85	16,6	19,15	15,95	18,4	11,95

Echantillon N°	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713
Quantité d'ADN (ng/µl)	2,65	4,45	3,55	12,6	11,2	23,2	9,3	14,65	13,6	80,25	13,45	10,3	32,35	8	45,95	41,55

Echantillon N°	714	715	716	717	718	719	720	721	722	723	724	771	772	773
Quantité d'ADN (ng/µl)	46,15	46,5	23,55	16,1	21,3	48,9	10,7	54,6	45,35	18	25,55	2,95	2,3	18,25

Site : Comores

Echantillon N°	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546
Quantité d'ADN (ng/µl)	6,75	4,3	6	5,55	6,9	7,5	6,6	10,25	4	6,1	7,75	5,85	9,8	4,85	10,65	10,35	10,8	14,15	10,3	10,95

Site : Maurice

Echantillon N°	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566
Quantité d'ADN (ng/µl)	6,55	5,7	6,2	5,4	5,5	8,95	5,6	6,2	8,2	6,15	8,3	5,5	7,8	6,95	9,55	11,2	8,2	15,1	9,4	7,9

Site : Hiazfotsy

Echantillon N°	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581
Quantité d'ADN (ng/μl)	7,2	17,8	6,4	7,6	9,35	21,55	20	20,16	19,8	5,1	16,95	10,05	31,76	20,35	20,8

Site : Zazafotsy

Echantillon N°	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607
Quantité d'ADN (ng/μl)	5,2	5,65	7,15	9	5,45	7,05	5,45	8	7,6	7,6	12,75	7,3	10,55	13,55	11,6	7,15	17,4	16,3	27,85	6,9

Site : Andapa

Echantillon N°	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624
Quantité d'ADN (ng/μl)	15,6	12,75	4,05	2,6	1,6	3,5	2,6	4,4	4,5	3,5	10,55	3,7	12,2	5	9,4	9,3	6,5

Site : Toamasina

Echantillon N°	628	629	630	631	632	633	639	640	641	642	643	644	645	646	647
Quantité d'ADN (ng/μl)	3,75	8,6	4,75	4,6	4,2	4,2	2,35	6,95	5,63	7,75	2,9	37,5	4,9	4,3	6,6

Site : Diégo

Echantillon N°	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664
Quantité d'ADN (ng/µl)	3,43	3,4	2,95	2,9	2,7	1,95	3,25	5	9,45	7,5	5,7	11,9	4,6	4	4,25

Site : Morondava

Echantillon N°	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684
Quantité d'ADN (ng/µl)	4,55	22,95	4,7	7,75	6	5,6	4,13	4,05	5,4	3,2	19,25	17,4	5,9	25,05	17,25	24,5	11,5

Echantillon N°	685	686	687	688	689	690	691	692	693	694	695	696	697
Quantité d'ADN (ng/µl)	15,7	22,85	2,65	37,615	55	75,75	18,85	44,45	47,05	30,75	39,75	70,85	18,05

Site : Antanamalaza

Echantillon N°	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791
Quantité d'ADN (ng/µl)	13,3	24,45	15,05	16	20,65	14,05	14,65	14,85	18,2	14,4	18,4	65,2	58,25	82,75	41,65	72,15	88,7	54,45

Site : Analakely

Echantillon N°	792	793	794	795	796	797	798	799	800	801	802
Quantité d'ADN (ng/µl)	16,8	15,9	15,6	20,4	14	29,35	24,3	14,1	13	16,95	11,9

Echantillon N°	803	804	805	806	807	808	809	810	811	812	813
Quantité d'ADN (ng/µl)	22,85	24,4	17,85	18,5	40,15	22,2	25,85	17,5	15,25	28,15	24,7

Site : Ivory

Echantillon N°	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829
Quantité d'ADN (ng/µl)	6,1	45,45	18,55	17,6	12,05	13,9	22,25	40,3	17,1	24,1	45,1	29,15	33,2	32,1	35,5	14,85

Echantillon N°	830	831	832	833	834	835	836	837	838	839	840	841	842	843
Quantité d'ADN (ng/µl)	21,05	11	30,6	12,8	14,1	37,55	24,25	17,5	12,85	12,25	19,3	16,4	42,65	38,6

## **RESUME**

**Titre** : Identification de marqueurs moléculaires pour l'étude préliminaire en phylogéographie de *Xenopsylla cheopis* (Rothschild, 1903) à Madagascar

**Auteur** : RAKOTOBE HARIMANANA Ravo Niaina

*Xenopsylla cheopis* est une espèce de puce connue comme étant le vecteur majeur de la maladie peste à Madagascar. Une étude phylogéographique sur cette espèce assure l'acquisition des connaissances sur sa diversité génétique, la distribution géographique de ses populations et plus loin, son mode de dispersion à l'échelle nationale. Dans cette étude, les objectifs principaux ont été de mettre au point par PCR les différents marqueurs moléculaires (nucléaire et mitochondriaux) potentiellement convenables pour des études de diversité génétique et de réaliser une étude préliminaire sur la phylogéographie de cette puce vectrice dans l'île. Pour ce faire, l'extraction de l'ADN a été réalisée sur les spécimens. Trois marqueurs mitochondriaux (COI, COII et COI-tailed) et un marqueur nucléaire (28S) ont été mis au point par amplification PCR. Une fois que les conditions optimales ont été identifiées, des échantillons en provenance de 24 sites (21 à Madagascar et 3 dans les îles voisines) ont été amplifiés utilisant le marqueur COI-tailed puis, les amplifiats obtenus ont été séquencés. Suite au séquençage et au nettoyage des séquences, 153 séquences nucléotidiques ont été obtenues. Les analyses des données de séquences ont permis de révéler la présence de 19 haplotypes dans les sites échantillonnés avec l'haplotype H2 qui est présent dans tous les sites sauf à Farafangana. La répartition géographique de ces haplotypes est hétérogène. Globalement, la similarité nucléotidique varie entre 66 - 100% toutefois, la majorité des valeurs de similarité se trouve à environ 100%. Ceci concluant la présence de diversité génétique inter- et intrapopulation de *Xenopsylla cheopis* dans ces sites d'étude, plus généralement à Madagascar.

Mots clés : *Xenopsylla cheopis*, peste, Madagascar, marqueurs moléculaires, Phylogéographie.

## **ABSTRACT**

**Title:** Identification of molecular markers for the preliminary phylogeographic study of *Xenopsylla cheopis* (Rothschild, 1903) in Madagascar

**Author:** RAKOTOBE HARIMANANA Ravo Niaina

*Xenopsylla cheopis* is a flea species known as the major vector of plague in Madagascar. Phylogeographic studies on this species ensure acquiring knowledge on its genetic diversity, on geographic distribution of its populations and on its dispersal pattern at a national scale. In this study, the main objectives were to set up PCR-based optimal conditions for molecular marker use (nuclear and mitochondrial markers); markers that are potentially suitable for genetic diversity and phylogeographic studies. To do this, DNA extraction has been performed using flea specimens. Three mitochondrial markers (mtCOI, mtCOII and mtCOI-tailed) and one nuclear marker (28S) have been set up by PCR. Once optimal conditions identified, flea samples from 24 sites (21 sites in Madagascar and three sites in the neighboring islands) have been amplified using COI-tailed then, obtained amplicons have been sent to sequencing. One hundred and fifty three sequences have been obtained. Data analysis using these sequences allowed revealing the presence of 19 haplotypes in the Malagasy sampled sites; with the haplotype H2 being present in almost all sites (except Farafangana). Geographic distribution of these haplotypes is heterogeneous. Globally, nucleotide similarity varies between 66 - 100% whereas; the majority of similarity values are about 100%. This concludes the presence of intra- and interpopulation genetic diversity for *Xenopsylla cheopis* at these sites, more generally in Madagascar.

**Keywords :** *Xenopsylla cheopis*, plague, Madagascar, molecular markers, Phylogeography.

## **RESUME**

**Titre** : Identification de marqueurs moléculaires pour l'étude préliminaire en phylogéographie de *Xenopsylla cheopis* (Rothschild, 1903) à Madagascar

**Auteur** : RAKOTOBE HARIMANANA Ravo Niaina

*Xenopsylla cheopis* est une espèce de puce connue comme étant le vecteur majeur de la maladie peste à Madagascar. Une étude phylogéographique sur cette espèce assure l'acquisition des connaissances sur sa diversité génétique, la distribution géographique de ses populations et plus loin, son mode de dispersion à l'échelle nationale. Dans cette étude, les objectifs principaux ont été de mettre au point par PCR les différents marqueurs moléculaires (nucléaire et mitochondriaux) potentiellement convenables pour des études de diversité génétique et de réaliser une étude préliminaire sur la phylogéographie de cette puce vectrice dans l'île. Pour ce faire, l'extraction de l'ADN a été réalisée sur les spécimens. Trois marqueurs mitochondriaux (COI, COII et COI-tailed) et un marqueur nucléaire (28S) ont été mis au point par amplification PCR. Une fois que les conditions optimales ont été identifiées, des échantillons en provenance de 24 sites (21 à Madagascar et 3 dans les îles voisines) ont été amplifiés utilisant le marqueur COI-tailed puis, les amplifiats obtenus ont été séquencés. Suite au séquençage et au nettoyage des séquences, 153 séquences nucléotidiques ont été obtenues. Les analyses des données de séquences ont permis de révéler la présence de 19 haplotypes dans les sites échantillonnés avec l'haplotype H2 qui est présent dans tous les sites sauf à Farafangana. La répartition géographique de ces haplotypes est hétérogène. Globalement, la similarité nucléotidique varie entre 66 - 100% toutefois, la majorité des valeurs de similarité se trouve à environ 100%. Ceci concluant la présence de diversité génétique inter- et intrapopulation de *Xenopsylla cheopis* dans ces sites d'étude, plus généralement à Madagascar.

Mots clés : *Xenopsylla cheopis*, peste, Madagascar, marqueurs moléculaires, Phylogéographie.