



**HAL**  
open science

# Génétique, génomique et histoire évolutive d'*Hermetia illucens*

Joseph Guilliet

► **To cite this version:**

Joseph Guilliet. Génétique, génomique et histoire évolutive d'*Hermetia illucens*. Génétique des populations [q-bio.PE]. Université Paris-Saclay, 2022. Français. NNT : 2022UPASL065 . tel-03935345

**HAL Id: tel-03935345**

**<https://theses.hal.science/tel-03935345v1>**

Submitted on 11 Jan 2023

**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.

# Génétique, génomique et histoire évolutive d'*Hermetia illucens*

*Genetics, genomics, and evolutionary history of Hermetia illucens*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 577, Structure et Dynamique des Systèmes Vivants (SDSV)  
Spécialité de doctorat : Génétique  
Graduate School : Science de la vie et santé. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **EGCE** (Évolution, Génomes, Comportement, Ecologie - Université Paris-Saclay, CNRS, IRD), sous la direction de  
**Jonathan FILEE**, Chargé de Recherche.

Thèse soutenue à Paris-Saclay, le 17 novembre 2022, par

**Joseph GUILLIET**

## Composition du Jury

<b>Myriam HARRY</b> Professeure des universités, Paris Saclay	Présidente
<b>Philippe LOPEZ</b> Professeur, Sorbonne Université, Paris	Rapporteur & Examineur
<b>Nathalie CASSE</b> Maîtresse de conférences, Le Mans université, HDR	Rapporteuse & Examinatrice
<b>Elisabeth HERNIOU</b> Directrice de recherche, Université de Tours, CNRS	Examinatrice
<b>Mohammed-Amin MADOU</b> Chargé de recherche, CEA Fontenay-aux-Roses	Examineur
<b>Jonathan FILEE</b> Chargé de recherche, CNRS, Paris Saclay	Directeur de thèse



**Titre :** Génétique, génomique et histoire évolutive d'*Hermetia illucens*.

**Mots clés :** Génétique ; *Hermetia illucens* ; évolution ; agro-industrie ; écologie ; bio-informatique.

**Résumé :** *Hermetia illucens*, la mouche soldat noire, est un diptère de la famille des Stratiomyidae. Cet insecte détritivore est de plus en plus utilisé en élevage de masse pour réduire les biodéchets et produire de la nourriture à destination des élevages aquacoles et avicoles. Elle est originaire du nord de l'Amérique Latine et s'est répandue sur une grande partie de la planète. Depuis une dizaine d'années, un nombre croissant d'entreprises élèvent et transforment des larves dans une optique d'économie circulaire. Cette thèse CIFRE a pour but d'améliorer les connaissances génétiques et génomiques sur la mouche soldat noire pour l'amélioration des performances et de la sélection. Nous nous sommes concentrés, dans un premier temps, sur l'obtention et l'analyse d'un génome de référence, pour l'espèce en utilisant du séquençage hybride. Nous avons également travaillé sur la diversité génétique et l'histoire invasive de la BSF. Nous l'avons fait au moyen d'un échantillonnage

mondial en utilisant à la fois le gène CO1 et l'assemblage d'une soixantaine de génomes mitochondriaux complets. Nous avons montré la complexité des scénarios d'invasion de la mouche soldat noire combinant à la fois un processus ancien de dispersion naturelle sur le continent Américain ainsi que, l'importance des introductions humaines, notamment en raison des élevages industriels. Enfin nous nous sommes également intéressés aux variations structurales des génomes, induites par les éléments transposables dans les populations naturelles d'*Hermetia illucens*, afin de déterminer les éléments les plus actifs dans les génomes ainsi que leurs dynamiques évolutives. Ce travail a révélé une étonnante et inattendue diversité génétique et génomique au sein de l'espèce. Il ouvre de nouvelles perspectives vers des expérimentations de sélection et d'amélioration sur la mouche soldat noire pour l'industrie entomocole.

**Title:** Genetic, genomic, and evolutionary history of *Hermetia illucens*.

**Keywords:** Genetics; *Hermetia illucens*; evolution; agro-industry; ecology; bioinformatics.

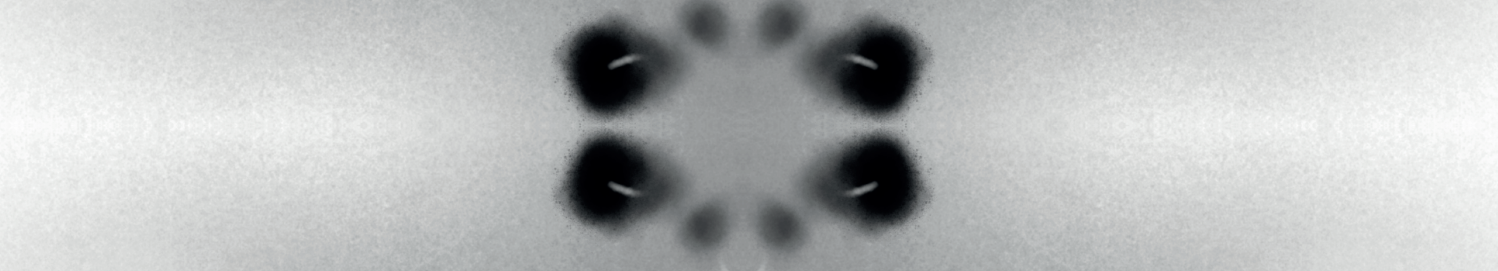
**Abstract:** *Hermetia illucens*, the black soldier fly, is a dipteran of the family Stratiomyidae. This detritivorous insect is increasingly used in mass production to reduce bio-waste and produce food for aquaculture and poultry farms. It originates from the north of Latin America and has spread over much of the planet. For the last ten years, many companies have been breeding and processing larvae from a circular economy perspective. This CIFRE thesis aims to improve genetic and genomic knowledge on the black soldier fly for performance improvement and selection. We focused, in a first step, on obtaining and analyzing a reference genome for the species using hybrid sequencing. We also worked on the genetic diversity and invasive history of BSF. We did this through global sampling using the CO1 gene and

assembling about 60 complete mitochondrial genomes. We have shown the complexity of the black soldier fly's invasion scenarios combining an ancient process of natural dispersal on the American continent and the importance of human introductions, notably due to industrial breeding. Finally, we also looked at the structural variations of the genomes induced by transposable elements in natural populations of *Hermetia illucens* to determine the most active elements in the genomes and their evolutionary dynamics. This work revealed the species' surprising and unexpected genetic and genomic diversity. It opens new perspectives towards selection and improvement experiments on the black soldier fly for the entomoculture industry.

---



Be.vtk



Be.vtk

A mon papa,  
Avec qui j'aurais aimé avoir le temps de partager ce moment de ma vie

# Sommaire général

Avant-propos	6
Remerciements	7
Introduction	11
Partie 1 : Du général au particulier	
Partie 2 : <i>Hermetia illucens</i> et <i>Homo sapiens</i>	
Partie 3 : Génétique	
Chapitre 1 :	112
Un génome pour les gouverner tous	
Chapitre 2 :	124
Ce que les génomes mitochondriaux complets nous apprennent sur l'histoire évolutive de la mouche soldat noire, <i>Hermetia illucens</i> .	
Chapitre 3 :	162
Une diversité nucléaire pleine de surprises	
Discussion générale	183
Conclusion	196
Bibliographie	198
Annexes	219

# Où il est question de comment ?

---

*La science, il est vrai, ne progresse qu'en remplaçant partout  
le pourquoi par le comment.  
André Gide*

---

D'aussi loin que je me souviens, j'ai toujours été fasciné par le fonctionnement des choses qui m'entourent, peut-être est-ce pour cela que j'ai décidé de devenir ingénieur. Je ne peux pas vraiment dire que je sois tombé dans la biologie directement, mon parcours n'a pas choisi le chemin le plus direct. Mais deux questions m'ont toujours animé : « comment cela fonctionne-t-il ? » et « comment l'ont-ils découvert ? ».

Ces deux questions sont à la fois complémentaires et différentes, si la première m'a permis de toujours vouloir étudier, la seconde m'a fait découvrir tous les génies derrière ces découvertes et toute l'ingéniosité des hommes et des femmes.

C'est à force de me poser ces questions que j'ai fini par embrasser ma thèse. J'ai appliqué ces deux mêmes questions : Comment est répartie la diversité génétique de la mouche soldat, comment s'en servir dans le cadre de sélection et surtout comment vais-je faire pour répondre moi-même à ces questions ? (Pour cette dernière, c'est, entre autres, en étant très bien entouré)

Et c'est en rentrant dans cet univers de la recherche que j'ai fait mienne la citation de Blaise Pascal : « si j'ai vu plus loin, c'est en me tenant sur les épaules de géants ». Car ce n'est finalement que cela, la recherche, nous apportons tous notre pierre à l'édifice pour essayer de voir toujours plus loin.

Outre la biologie, j'ai également appliqué mes deux questions favorites à tous les domaines que je croisais dans ma vie, affamé de comprendre et curieux de tout, j'aspire à mieux cerner le monde qui m'entoure. C'est ainsi qu'en parallèle de ma thèse j'ai entrepris l'écriture d'un roman, que je me suis pris de passion pour l'économie et la finance, et que j'ai toujours cette même curiosité pour les mathématiques, l'histoire, la physique, l'art et la géographie.

Voilà donc entre vos mains, l'une de mes fameuses questions que j'ai pris le temps de pousser un peu plus loin que les autres.

J'espère que la lecture de ce manuscrit vous apportera des connaissances, et des réponses à des questions que vous ne vous seriez jamais posées.



# Remerciements

---

*Fluctuat nec mergitur*

---

Merci pour les heures, les jours, les années, les discussions, les beaux moments. Merci pour cette thèse. Merci pour la complicité que nous avons partagée. J'ai aimé chaque instant. Merci pour l'aide, et les méninges en fusion, pour les projets qui ont vu le jour et ceux qui viendront. Merci d'avoir été là quand j'en avais besoin. Merci de m'avoir fait rêver. Merci de continuer.

Jonathan, merci pour cet encadrement, pour les échanges qui m'ont permis de passer d'un simple ingénieur 😊 à un docteur. Je suis très fier de t'avoir eu comme directeur de thèse. Merci pour la confiance et la liberté que tu m'as accordées, je ne crois pas que j'aurai pu espérer mieux.

Nicolas, merci du fond du cœur pour l'aide que tu m'as apportée tant au niveau humain que scientifique. Merci pour ta rigueur et tes conseils. Ils me sont précieux et le seront encore.

Merci maman, pour les relectures, les conseils, ton temps à m'écouter et les fautes corrigées ! Merci à mes frères et sœurs Estelle, Lise-Marie, Éric, Laurie-Anne et Aurélien ! J'ai toujours pu compter sur vous, quelques soient mes difficultés et pour ça vous êtes mes héros.

Merci à l'équipe de Cycle Farms, vos discussions et votre aide m'ont été précieuses. Floran et Marc-Antoine en premier lieu pour m'avoir fait confiance, Guillaume pour les conversations toujours enrichissantes, Adrien et Arnaud pour m'en avoir appris énormément sur notre mouche.

Merci aux honorables membres du jury, Myriam, Elisabeth, Amin, Nathalie Philippe et Armel d'avoir accepté d'évaluer mon travail.

Merci à Jean-Philippe et Éric pour les comités de suivi et les inestimables conseils que vous avez sus me donner.

Merci à tous mes amis (les pompiers en tête), et à ceux que j'ai rencontré le temps d'un voyage, d'un congrès ou d'une soirée. Vous raconter mon



sujet de thèse, encore et encore, m'a apporté une aide considérable lorsqu'il a fallu rédiger ce manuscrit ! A force de raconter l'histoire j'avais, avant même de commencer, une bonne idée de la façon dont j'allais m'y prendre !

Marie et Baptiste, vous avez été plus que des co-bureaux, je suis extrêmement heureux de vous avoir rencontrés. Merci pour les rires et les partages.

Merci Maurice (c'est mon ordinateur), de ne pas m'avoir lâché même si j'ai dû te réanimer tellement souvent que désormais c'est de l'acharnement thérapeutique.

Merci aux anciens, Cécile, Andréas, Romain, Julia, Maxime, Perrine et Vincent de m'avoir accueilli au labo et pour les moments conviviaux (oui c'est un synonyme de bières). Merci et bon courage aux futurs (Marie et Baptiste ça va la pression ?) Maléna, Etienne, Samuel, Héloïse, Erina, Siddharth, Zhen, Ksenia, Apolline et Jérémy (team BSF rpz) !

Merci aux équipes du labo et en particulier à Frédéric, David, Aurélie, Clément et Jean-Bernard. Chacun de vous m'a beaucoup apporté.

Pouik à Gorphe et Forge !


Merci à Beink pour les dessins, et je suis sûr que cette entreprise fera des merveilles ! Merci, encore, Jeanne. Merci Caty, Yves et Victor.


Enfin, merci à toi, que je connais ou non, pour prendre le temps de lire ces quelques mots et peut être même de parcourir les quelques pages de ce manuscrit. J'y ai mis du cœur et ma personnalité, j'y ai mis ma sueur et des heures de ma vie, j'y ai mis mes neurones et ceux des autres. Si une thèse est un achèvement pour certain, pour moi c'est une étape. J'ai ouvert plus de questions, que j'en ai résolues, comptez sur moi pour essayer d'y répondre.


Une thèse est une aventure sociale solitaire, j'ai adoré la mienne ! Et quitte à paraître fleur bleue, je vous aime.

Joseph


# Hermetia illucens

 mouche soldat noire


 Lalat askar hitam


 mosca-soldado-negro

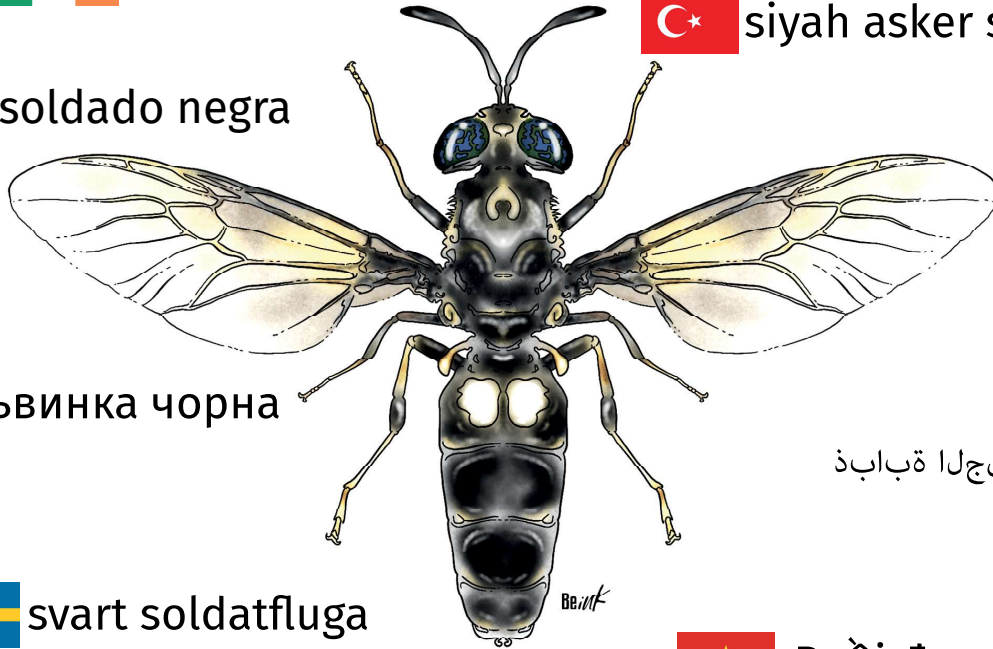
 black soldier fly

 schwarze soldatenfliege

 saighdiúir cuile dubh


 siyah asker sineği

 mosca soldado negra




 黑水虻


 львинка чорна


 دوسألا يدنجالا ةبابذ


 svart soldatfluga

 Ruồi đen

 mosca soldat negra


 zwarte soldaatvlieg


 उडणारा काळा सैनिकि


 mosca soldato nera

 רוחשה לייחה בובז

 Lalat tentara hitam

 Чёрная львинка

 czarny żołnierz

 アメリカミズアブ

 mustasotilaskärpänen

# Introduction

## Table des matières de l'introduction

Table des matières .....	11
Table des illustrations .....	12
Définitions .....	15
Partie 1 : Du général au particulier .....	17
Les insectes .....	18
Phylogénie générale des insectes .....	26
Diptères et mouches soldats .....	36
<i>Hermetia illucens</i> .....	45
Description des différents stades d' <i>Hermetia illucens</i> et Cycle de vie .....	46
Répartition de la mouche soldat noire .....	54
Partie 2 : <i>Hermetia illucens</i> et <i>Homo sapiens</i> .....	56
Entomologie médico-légale .....	57
<i>Quid</i> des déchets ? .....	59
Microbiologie et digestion .....	66
Utilisation comme source de protéine .....	70
Utilisation comme source de lipide .....	74
Autres utilisations et aspects réglementaires .....	77
Aspect financier .....	81
Partie 3 : Génétique .....	86
Structure des génomes et séquençage .....	87
Structure des génomes .....	87
Le séquençage .....	94
Génétique des populations .....	100
La diversité génétique et son étude .....	100

<i>Hermetia illucens</i> , comme espèce d'étude en génétique des populations.....	102
Sélection et amélioration .....	107
Problématiques.....	111

## Table des illustrations de l'introduction

Figure 1 : Echelle des temps géologiques.....	18
Figure 2 : Carte de la distribution de la croûte continentale, des bassins océaniques et des limites de plaques il y a 500 millions d'années. Issus de Merdith et al 2021.....	19
Figure 3 : Diversité des hexapodes - proportion des espèces décrites. Figure issue de Grimaldi et Engel – 2005 – Evolution of the insects.	20
Figure 4 : Biomasse estimée par groupe d'animaux. Adaptée de Eggleton 2020.....	21
Figure 5 : Estimation de la richesse mondiale en espèce au niveau de l'embranchement et des ordres.....	21
Figure 6 : Cladogramme représentant les premières subdivisions des insectes.....	23
Figure 7 : Cladogramme des tétrapodes représentant un groupe paraphylétique : les reptiles.....	24
Figure 8 : Cladogramme des tétrapodes représentant un groupe polyphylétique : les animaux à sang chaud.....	25
Figure 9 : Spéciation allopatrique.....	30
Figure 10 : Spéciation péripatrique (A), parapatricque (B) et sympatrique (C).....	31
Figure 11 : Cladogramme des animaux aux insectes.....	33
Figure 12 : Cladogramme des insectes vers les ordres.....	35
Figure 13 : Ailes en haltères d'un Tipule (vue dorsale).....	36
Figure 14 : Echelle des temps géologiques.....	37
Figure 15 : <i>Ornidia obesa</i> .....	38
Figure 16 : <i>Clogmia albipunctata</i> .....	40
Figure 17 : Cladogramme des diptères au Stratiomyidae.....	41
Figure 18 : Dessin d'observation d'un Stratiomyidae.....	42
Figure 19 : <i>Stratiomys chamaeleon</i> .....	42
Figure 20 : <i>Odontomyia ornata</i> .....	43
Figure 21 : <i>Hermetia illucens</i> .....	45

Figure 22 : Description des différentes parties d'une <i>Hermetia illucens</i> adulte.....	48
Figure 23 : Dimorphisme sexuel chez <i>Hermetia illucens</i> .....	48
Figure 24 : Reproduction chez <i>Hermetia illucens</i> .....	50
Figure 25 : Œufs d' <i>Hermetia illucens</i> .....	50
Figure 26 : Description des différentes parties d'une larve d' <i>Hermetia illucens</i> .....	51
Figure 27 : Evolution des larves d' <i>Hermetia illucens</i> .....	52
Figure 28 : Résumé du cycle de vie d' <i>Hermetia illucens</i> .....	53
Figure 29 : Ensemble des rangs taxonomiques de la mouche soldat noire.....	55
Figure 30 : Datation de l'intervalle post-mortem – reconstruction ...	58
Figure 31 : quantité de déchet par habitant et par jour.....	60
Figure 32 : Déchet généré par région du monde.....	61
Figure 33 : Composition mondiale des déchets.....	62
Figure 34 : <i>Musca domestica</i> - Dessin observation (Beink).....	64
Figure 35 : Pyramide du traitement des déchets.....	65
Figure 36 : Parties et pH d'un intestin d' <i>Hermetia illucens</i> .....	66
Figure 37 : Analyse Nonmetric Multidimensional Scaling (NMDS) du microbiote de larves d' <i>Hermetia illucens</i> nourris avec différents substrats.....	67
Figure 38 : Résumé schématique du système digestif des larves de BSF.....	69
Figure 39 : Analyse de l'inclusion de protéines issu de mouche soldat noires dans la croissance des truites européennes – <i>Salmo trutta</i> . 72	
Figure 40 : Comparaison des quantités de biocarburant produites en fonction des matières premières utilisées.....	74
Figure 41 : Résumé schématique de la production de biocarburant par l'action combiné de <i>Tenebrio molitor</i> et <i>Hermetia illucens</i> à partir de fanes de maïs.....	75
Figure 42 : Résumé schématique du processus de transformation de biodéchet par l'utilisation de mouches soldat noires.....	77
Figure 43 : Augmentation du nombre de publications traitant d' <i>Hermetia illucens</i> entre 1955 et 2021.....	79
Figure 44 : Mots clefs et relations entre ces mots clefs dans les publications traitant d' <i>Hermetia illucens</i> .....	80
Figure 45 : Croissance attendu du marché des insectes comestibles.....	82

Figure 46 : Evolution du nombre d'entreprises d'élevages d'insectes .....	83
Figure 47 : Quantités comparées de nourriture d'espace et d'eau ....	83
Figure 48 : Génome mitochondrial d' <i>Hermetia illucens</i> .....	88
Figure 49 : Chromosomes d' <i>Hermetia illucens</i> .....	89
Figure 50 : Couverture Log2 (mâle/femelle) des données de reséquençage du génome entier.....	90
Figure 51 : Comparaison des tailles de génomes d'arthropodes .....	90
Figure 52 : Classification simplifiée des éléments transposables .....	92
Figure 53 : Proportion des différentes familles de TE chez les diptères .....	93
Figure 54 : MinION développé par Oxford Nanopore Technologie .....	95
Figure 55 : Détail d'une molécule d'ADN séquençé dans l'un des pores moléculaires du MinION .....	96
Figure 56 : Exemple d'assemblage OLC.....	98
Figure 57 : Résumé d'un assemblage avec une étape de scaffolding .....	99
Figure 58 : Effets comparés de souches génétiquement différentes sur différents substrats.....	102
Figure 59 : Réseau d'haplotype CO1 d' <i>Hermetia illucens</i> en Corée du Sud .....	103
Figure 60 : Réseaux d'haplotypes CO1 après un échantillonnage mondial sur 7 zones biogéographiques différentes.....	104
Figure 61 : Dynamique de la distribution mondiale déduite des groupes génétiques d' <i>Hermetia illucens</i> déterminés via des marqueurs microsatellites.....	106
Figure 62 : Comparaison des mutants PTTH et sauvages.....	107
Figure 63 : Comparaison des mutants sans ailes et sauvage .....	108
Figure 64 : Evolution du nombre d'œufs collectés entre une souche d'élevage (bleu) et une souche consanguine issue de la souche d'élevage (rouge).....	109
Figure 65 : Comparaison du poids des larves sous sélection massale .....	110

## Définitions

**Gène :** En terme moléculaire, un gène se définit comme la séquence complète d'acide nucléique nécessaire à la synthèse d'un polypeptide ou d'un ARN fonctionnel.

**Génome :** Il s'agit de la séquence nucléotidique en ADN (ou ARN pour certains virus) qui contient toutes les informations d'un organisme. Le génome se décompose en gènes codant pour les protéines ou les gènes non codant (ARN de transfert ou ribosomique par exemple), et l'ADN non codant.

**Génétique :** La génétique est l'étude des gènes, des variations génétiques et de l'hérédité dans les organismes. Selon ce qui est étudié, on parlera de génétique quantitative, qui étudie les caractères quantitatifs (taille, poids, vitesse de croissance ...); de génétique des populations qui étudie les forces évolutives et leurs effets qui vont influencer la diversité génétique des populations.

**Génomique :** La génomique est l'étude des génomes. Elle s'intéresse tout particulièrement aux structures et à l'évolution de ceux-ci. On pourra retrouver dans cette discipline, la génomique structurale, qui a pour but de séquencer, d'assembler et d'annoter les génomes entiers et la génomique fonctionnelle qui elle, cherchera d'avantage à déterminer la fonction des gènes.

**Histoire évolutive :** On entend par histoire évolutive d'une espèce, l'ensemble des évènements qui ont conduit à avoir une structure de population telle qu'on la retrouve aujourd'hui. L'histoire évolutive s'intéresse d'une part à la diversité génétique (en datant des séparations de populations par exemple) mais également aux mécanismes qui ont conduit aux observables présents (routes d'invasions, évènements de colonisations...).

*Hermetia illucens*: Il s'agit de la nomenclature binomiale de la mouche soldat noire (illustré ci-dessous). Souvent abrégée en **BSF** (Black Soldier Fly – en anglais), *Hermetia illucens* est un diptère de la famille des Stratiomyidae.

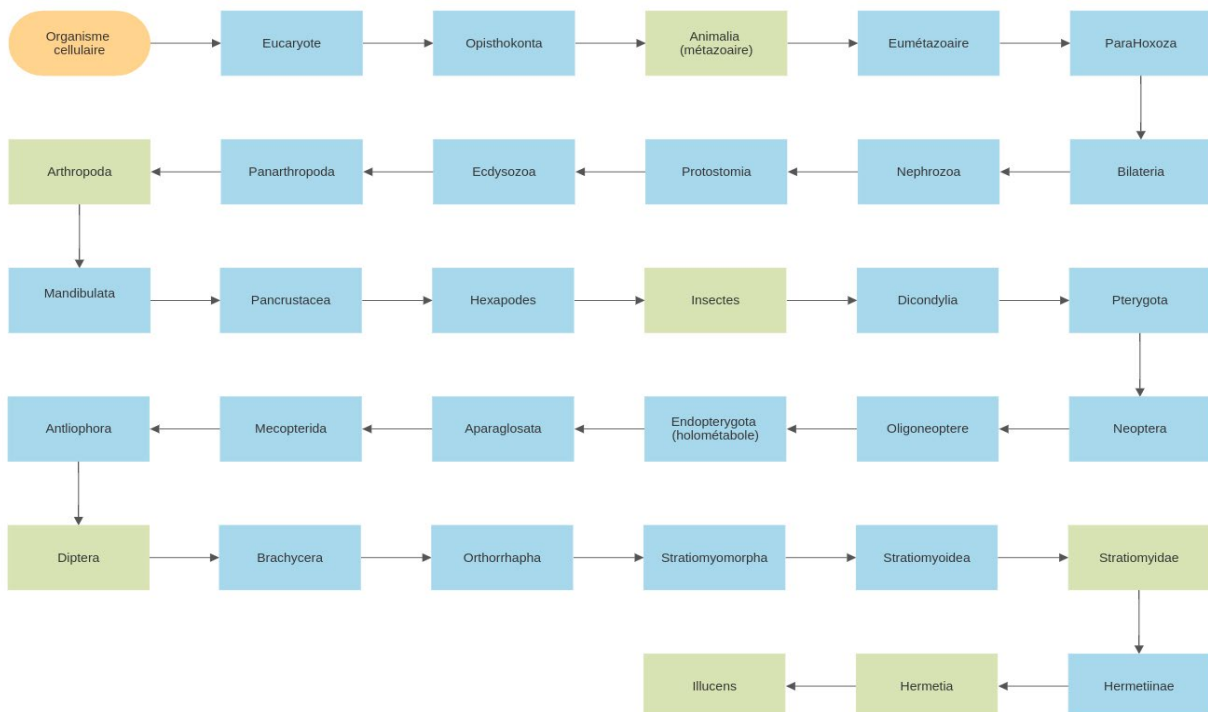




# Partie 1 : Du général au particulier

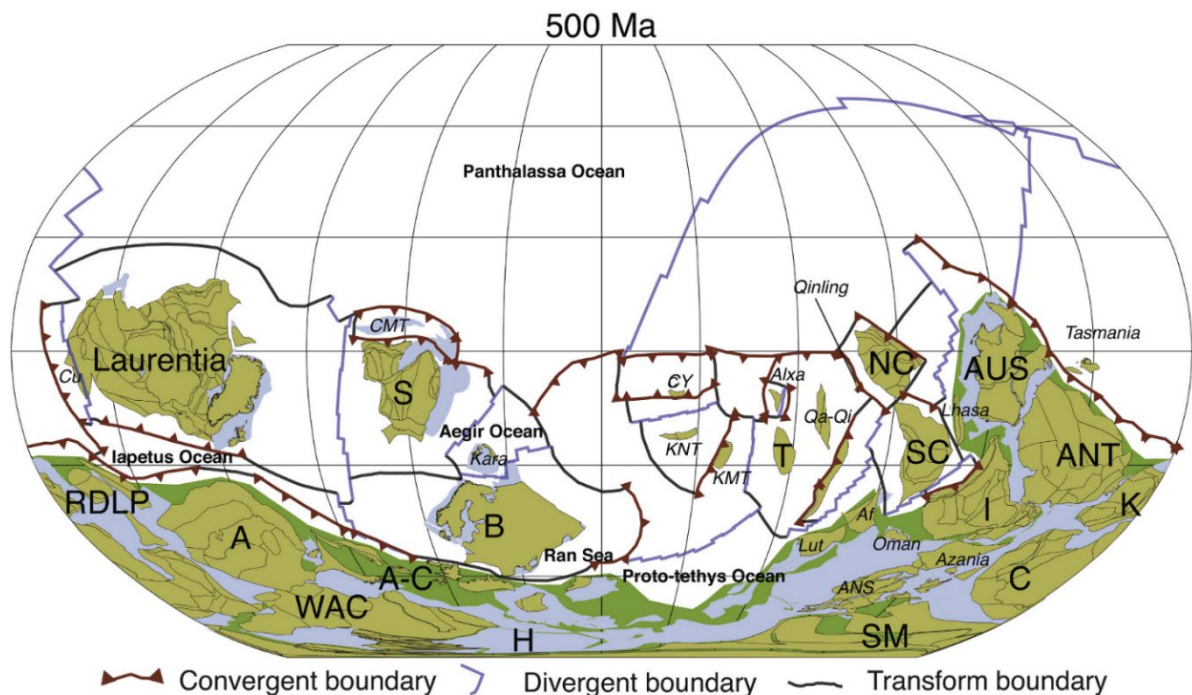
*Au cours de cette première partie, je vous propose un voyage. Mais pas n'importe quel voyage, ceux qui peuvent vous faire rêver, vous faire toucher du doigt un autre monde. Un monde si proche de nous que nous n'en avons même plus conscience, à l'exception peut-être de nos soirées d'été où l'on entend un bourdonnement dans nos oreilles. Le monde des insectes.*

*Les insectes sont partout présents autour de nous, ils nous font vivre autant qu'ils prennent nos vies. Ils nous rendent le monde plus beau autant qu'ils le détériorent. Alors finalement qui sont-ils ? Comment l'être humain les a-t-il classés, rangés, ordonnés ? Et surtout qui est cette mouche soldat noire ?*



Cette figure n'aura plus de secret pour vous à la fin de cette partie.





Laurentia : Amérique du nord; A : Amérique du sud; B Europe; NC & SC : Chine  
WAC & SM : Afrique du Nord; S : Sibérie ; C : Congo; K : Afrique du Sud

Figure 2 : Carte de la distribution de la croûte continentale, des bassins océaniques et des limites de plaques il y a 500 millions d'années. Issus de Merdith et al 2021.

Environ quarante millions d'années après les premiers hexapodes, l'ancêtre commun des insectes apparaît, et encore trente millions d'années après, apparaissent les ptérygotes, c'est-à-dire les premiers insectes volants (Montagna et al 2019).

Pendant 150 à 200 millions d'années, les insectes furent les seuls animaux sur terre à pouvoir voler, ils ont ainsi conquis une grande partie de la planète, en effet les premiers vertébrés volants sont apparus au cours du Trias, il y a environ 215 millions d'années (Britt et al 2018).

Les insectes représentent la majeure partie de la biodiversité animale actuelle. En 2020, sur 1 659 420 espèces animales décrites (Zhang 2013), 1,05 millions sont des insectes (The IUCN Red List of Threatened Species). Les insectes font partie d'un groupe un peu plus large que sont les hexapodes dont ils représentent l'écrasante

majorité (figure 3). Cependant, une grande partie n'a pas encore été décrite, et on estime le nombre total d'espèces à 5,5 millions (Stork 2018). Ceci représenterait plus de la moitié du nombre d'espèces animales estimées (Mora et al 2011).

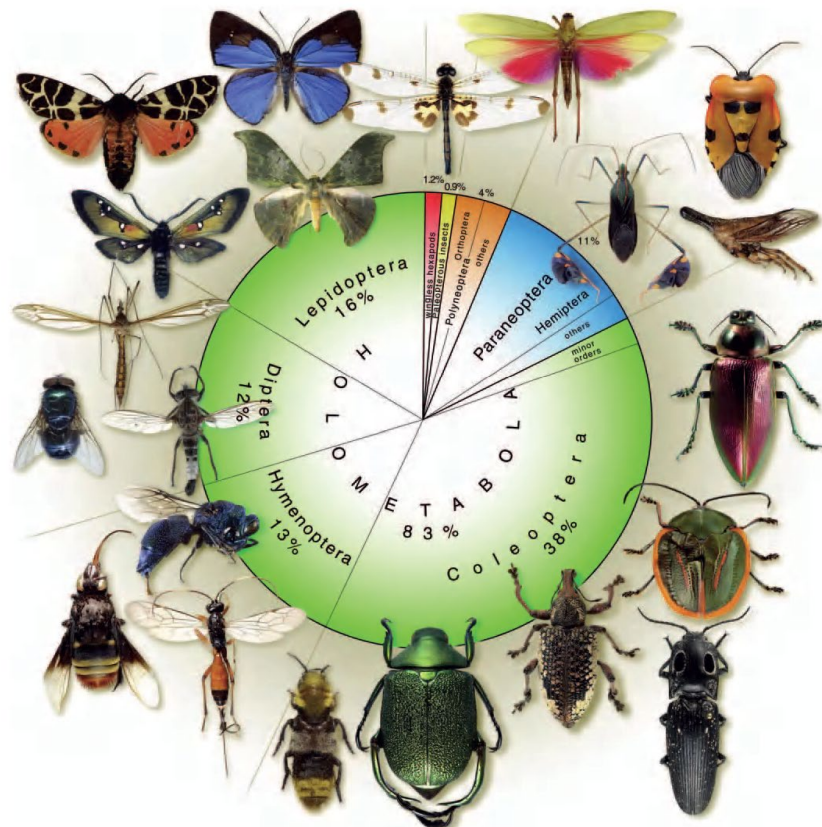


Figure 3 : Diversité des hexapodes - proportion des espèces décrites. Figure issue de Grimaldi et Engel - 2005 - Evolution of the insects

Donnons quand même quelques chiffres supplémentaires, pour se donner davantage le vertige. L'entomologiste Edward Osborne Wilson estime qu'il y a  $10^{19}$  insectes vivant en même temps (Hall 2008).

Quant à la biomasse, des arthropodes terrestres, dont 80% des espèces estimées sont des insectes (les 20% restant étant composées des myriapodes et des arachnides), elle est évaluée à 200 millions de tonnes (Figure 4) soit 3 fois celle de tous les humains combinés (Eggleton 2020).

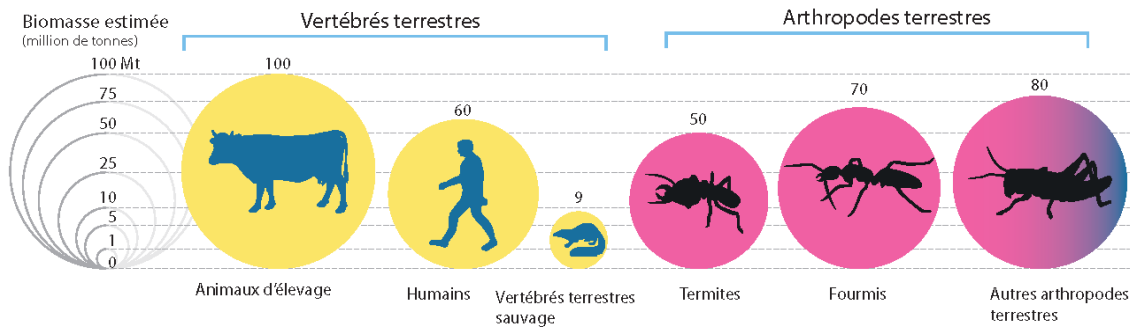


Figure 4 : Biomasse estimée par groupe d'animaux. Adaptée de Eggleton 2020

Parmi les insectes qui ont été décrits, on peut citer les Diptères (mouches ; moustiques ->160 000 espèces décrites), les lépidoptères (papillons et autres ->158 000 espèces décrites) les coléoptères (scarabées, coccinelles ... ->392 000 espèces décrites). Toutes les données proviennent de [Eggleton 2020](#) (figure 5).

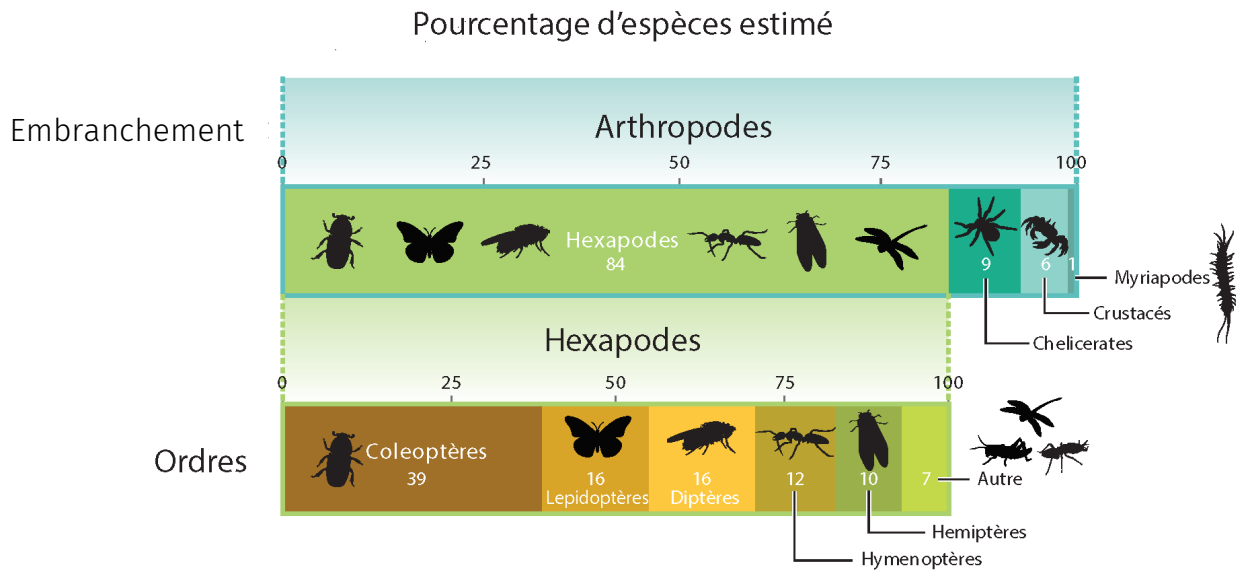


Figure 5 : Estimation de la richesse mondiale en espèce au niveau de l'embranchement et des ordres.

84% des arthropodes sont des hexapodes. L'écrasante majorité des hexapodes sont des insectes. Adaptée de Eggleton et al 2020.

A propos des scarabées, la diversité réellement hallucinante au sein de ce groupe, fera dire au généticien John Burdon Sanderson Haldane “ Dieu a une affection démesurée pour les coléoptères” ; en effet il s’agit du groupe le plus représenté chez les insectes (étant eux-mêmes les plus nombreux parmi les animaux).

Tous les insectes ont des points communs, tels que, comme nous l’avons déjà vu, un corps métamérisé et segmenté en trois parties, que l’on appelle tagmes, mais également un exosquelette composé de chitine, une respiration qui s’effectue à l’aide de trachées respiratoires et des pièces buccales externes.

Pour ce qui est de leur développement, les insectes passent par différents stades, différentes mues. Ici il va falloir commencer à rentrer un peu dans le détail en les séparant en deux clades (figure 6) :

- Les Aptérygotes (Brauer, 1885) :

Il s’agit d’insectes qui au cours de leur évolution n’ont jamais eu d’ailes. Chez les aptères, le développement est qualifié d’amétabole, c'est-à-dire que les jeunes insectes ressemblent à l’adulte en version réduite.

- Les Ptérygotes (Lang, 1888) :

Il s’agit ici de la majorité des insectes, et qui ont des ailes (ou qui les ont perdues au cours de leur histoire évolutive). Ici à nouveau on va distinguer deux types de développements au cours du cycle de vie :

- Les Hétérométaboles : Qui n’ont pas de stade immobile entre la larve et l’adulte et dont le stade juvénile ne se distingue du stade adulte, que par l’absence d’appareil génital et d’ailes (par exemple les punaises).

- Les Holométaboles : Qui eux, possèdent la plus part du temps un stade immobile entre la larve et l’adulte, la nymphe ; de plus, chez les holométaboles les larves sont très différentes des adultes avec des milieux de vie qui peuvent être très différents (par exemple : les mouches ou les papillons).

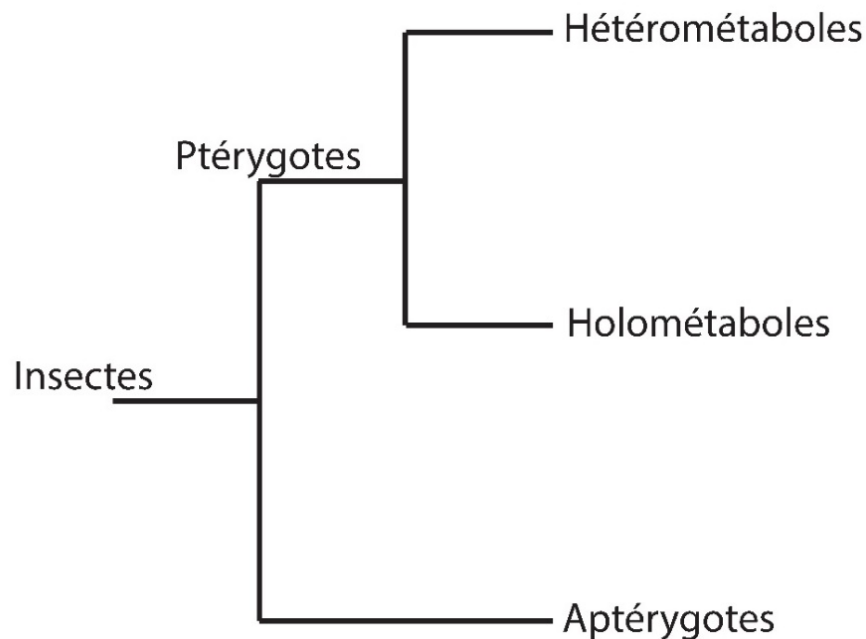


Figure 6 : Cladogramme représentant les premières subdivisions des insectes.

*Un cladogramme est une représentation schématique des liens de parentés entre les différents groupes du vivant. Les longueurs de branches ne sont pas représentatives d'un temps d'évolution. Ici on voit que tous les holométaboles sont des ptérygotes et eux-mêmes sont des insectes*

Comme vous pouvez le constater nous avons déjà commencé à ranger, à classer, les insectes dans des cases. Nous avons regroupé nos clades dans d'autres clades, etc. Donc avant de passer à une partie qui sera consacrée au rangement des êtres vivants, il est temps de donner quelques définitions.

Un taxon est un regroupement d'êtres vivants selon des caractéristiques communes.

Ce groupe, s'il contient tous les descendants d'un ancêtre commun, sera qualifié de monophylétique, par exemple les félins.

Si ce groupe exclut l'un des descendants de l'ancêtre commun, on qualifie alors ce groupe de paraphylétique, c'est le cas des reptiles (en rouge sur la figure 7) par exemple, qui exclue les oiseaux (en bleu).

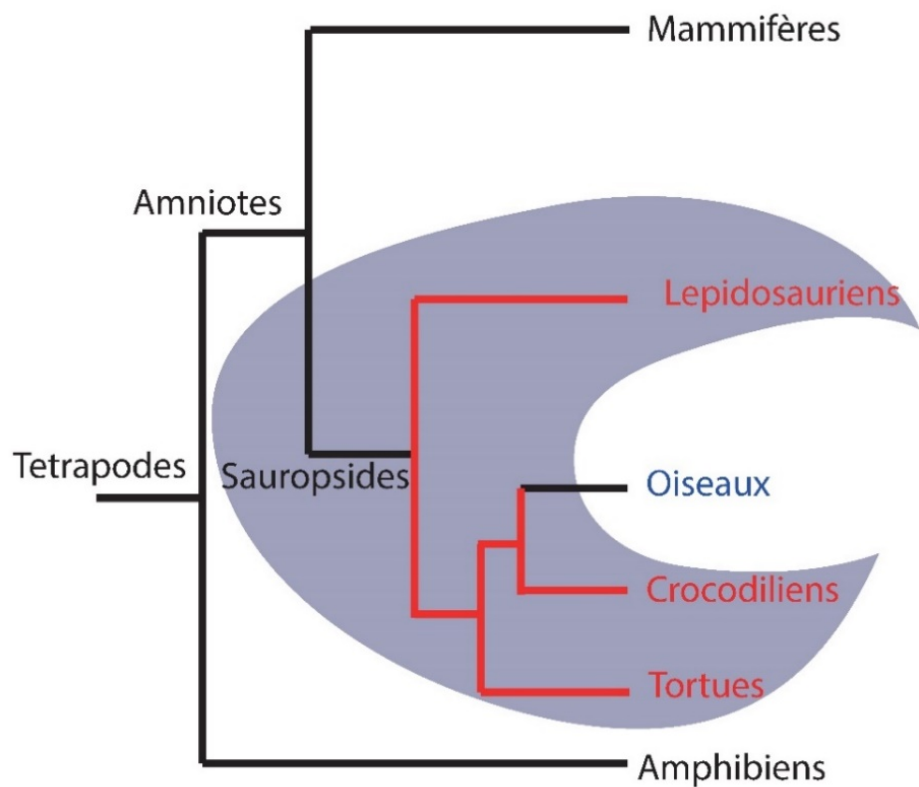


Figure 7 : Cladogramme des tétrapodes représentant un groupe paraphylétique : les reptiles

Si le critère de regroupement est un caractère qui n'a pas été hérité d'un ancêtre commun, alors on qualifie ce taxon de polyphylétique. C'est le cas par exemple des animaux à sang chaud (figure 8).



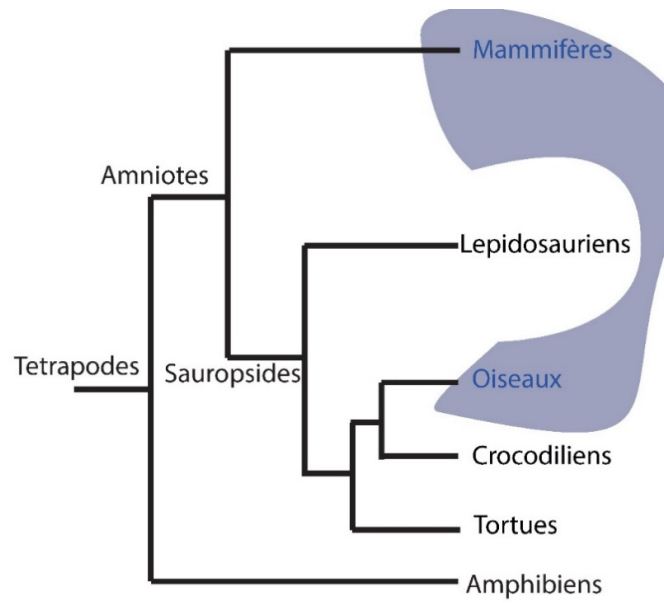


Figure 8 : Cladogramme des tétrapodes représentant un groupe polyphylétique : les animaux à sang chaud.

## Phylogénie générale des insectes.

---

*“L’immense tentative humaine d’expliquer l’univers n’est autre qu’une admirable et multiple illusion : celle qui tente de satisfaire une aspiration métaphysique et une nécessité logique dans une quête toujours renouvelée de nommer, d’expliquer et d’ordonner les objets et les phénomènes présents et agissants sur cette terre, afin de mieux les maîtriser et de réussir l’entreprise infiniment complexe et nécessaire de vivre en société. » - Nicole Revel (Fleurs de paroles : histoire naturelle Palawan, tome 1. Les dons de Nagsalad)*

---

Ah ! Ranger, classer, ordonner ... nommer, mais avant tout observer. Comme le dit excellemment bien l’anthropologue et directrice de recherche émérite du CNRS Françoise Grenand dans son article intitulé : “Nommer son univers : Pourquoi ? Comment ? Exemples parmi des sociétés amazoniennes” :

Observer pour décrire,  
Décrire pour nommer,  
Nommer pour connaître,  
Connaître non seulement pour utiliser,  
Mais aussi pour penser.

Car il va de soi que nous pensons avec nos mots, avec notre vocabulaire, alors nommer c’est se donner des outils pour mieux comprendre le monde qui nous entoure.

Rangeons donc.

L’idée du “qui se ressemble s’assemble” était dans les esprits des naturalistes et autres biologistes depuis bien longtemps. Aristote, déjà, séparait les êtres vivants selon des caractéristiques qu’il avait définies. Puis Linné, créera la classification moderne et la nomenclature binomiale avec un esprit fixiste. Puis viendra Lamarck qui proposera une idée d’évolution, (“l’usage entraîne la

transformation" / "le caractère acquis des parents se transmet aux descendants"). Et enfin Wallace et Darwin, ("Les ressemblances et différences entre organismes résultent de l'évolution et de leur interaction avec le milieu").

La systématique est donc l'art de décrire, de trier, de ranger, d'ordonner le monde vivant qui nous entoure. Aussi nous faut-il des critères pour faire ce tri ;

D'un point de vue morphologique, le "qui se ressemble s'assemble" peut nous amener à faire des erreurs, en effet, comme le souligne Darwin, les ressemblances peuvent provenir de l'interaction avec le milieu sous forme de convergence évolutive. C'est peut-être pour cela que l'on a autant d'espèces qui ressemblent à des crabes ? (McLaughlin & Lemaître 1997)

D'un point de vue génétique, en revanche, plus deux séquences héritées d'un ancêtre commun seront ressemblantes plus les individus seront apparentés, exception faite des transferts horizontaux entre espèces. Et c'est cela la phylogénie moléculaire, nous ne sommes plus dans une classification mettant en relation les ancêtres et leurs descendant (généalogie) ; mais dans la recherche des plus proches parents entre les individus ou entre les espèces. Mais d'ailleurs voilà deux termes que nous allons avoir besoin de définir avant de continuer.

Commençons donc par l'individu, chacun pouvant se figurer ce qu'est un individu, surtout lorsque l'on parle d'animaux, je ne m'attarderai pas sur cette notion, je me contenterai de donner la définition proposée par le petit Robert un individu est un "corps organisé vivant d'une existence propre et qui ne saurait être divisé sans être détruit (plante, animal...)."

Une fois accepté la notion d'individu, nous arrivons très vite à la notion d'espèce.

La biologie nous offre, pour l'espèce, pas moins de 22 concepts différents (Oui, là est l'embarras - Shakespeare ; Hamlet acte III, scène 1), alors il faut faire un choix, car sans une base forte et une notion acceptée, il est difficile de définir le reste, vous en conviendrez.

Aussi lorsque je parlerai d'espèce désormais, me référerai-je à la définition la plus couramment acceptée dans les milieux

scientifiques, celle d'Ernst Mayr de 1942 : *“ une espèce est une population ou un ensemble de populations dont les individus peuvent effectivement ou potentiellement se reproduire entre eux et engendrer une descendance viable et féconde, dans des conditions naturelles. Ainsi, l'espèce est la plus grande unité de population, au sein de laquelle le flux génétique est possible. Les individus d'une même espèce sont donc génétiquement isolés d'autres ensembles équivalents, du point de vue reproductif.”*

Maintenant que nous sommes armés de ces définitions, demandons-nous comment construire des phylogénies ? Comme nous l'avons vu, avant l'avènement de la génétique, des classifications du vivant existaient déjà, elles se basaient sur des caractères morphologiques, comportementaux ou encore des milieux de vie commun. Avoir des caractères communs était donc le signe d'un ancêtre commun. Désormais, et ce grâce à l'apport de la génétique, nous pouvons affiner nos classifications, en comparant des séquences protéiques ou nucléotidiques (ADN, ARN). En effet les séquences génétiques homologues sont héritées d'un ancêtre commun et accumulent des mutations au fur et à mesure des générations.

Partant de ce constat, il devient alors possible de réaliser des phylogénies, basées sur un gène, un ensemble de gènes, voire des génomes complets. Elles nous permettront d'une part de connaître les relations de parentés entre les différents individus et d'autre part de dater le temps de coalescence de ces séquences. Ce dernier est un proxy du temps auquel vivait l'ancêtre commun des individus étudiés. Pour le dire autrement, que l'on parle de spéciation ou de séparation de populations, la séparation des individus au regard d'un gène donné peut se situer avant la séparation effective des populations. L'apport de la génétique dans la phylogénie a été déterminant, tant au niveau de la justesse des relations phylogénétiques des individus, qu'au niveau de la datation des ancêtres communs et de l'évolution de la structure des génomes (Lopez et al 2002).

Ainsi l'apport de la génétique a permis de résoudre des ambiguïtés de longue date, mais également de replacer les espèces les unes par rapport aux autres dans l'histoire évolutive. C'est le cas notamment des papillons dont certains genres ont été récemment replacés au sein de la phylogénie générale des lépidoptères (Kawahara and

Breinholt, 2014). Cependant toutes les incertitudes n'ont pas encore été résolues à l'image de la monophylie des nématocères qui n'est pas encore établie (Yeates and Wiegmann, 1999).

Pour faire ces phylogénies, il faut partir d'un alignement de séquence nucléotidique ou protéique puis via des méthodes bio-informatiques, réaliser l'arbre étant celui qui correspond à l'histoire évolutive de toutes les séquences. Ces méthodes se basent, soit sur des distances génétiques, c'est-à-dire, sur les similarités entre les paires de séquences, par exemple la méthode du Neighbour Joining (Saitou et Nei 1987) ; soit sur des études probabilistes (pour chaque acide aminé ou nucléotide d'un alignement de séquences on calcul la probabilité qui va optimiser les paramètres définis). C'est le cas de la méthode du Maximum Likelihood (maximum de vraisemblance) qui est une méthode d'inférence de la phylogénie.

Cette méthode évalue une hypothèse sur l'histoire évolutive en termes de probabilité que le modèle proposé et l'histoire hypothétique donnent lieu à l'ensemble des données observées. La supposition est, qu'une histoire avec une probabilité plus élevée d'atteindre l'état observé est préférée à une histoire avec une probabilité plus faible (Nixon 2001). La méthode recherche l'arbre avec la probabilité ou la vraisemblance la plus élevée. Cette méthode se base sur des modèles, définis à l'avance, qui appliqueront différentes probabilités aux paramètres génétiques tel que par exemple le taux de substitution ou la fréquence nucléotidique (Delsuc et Douzery 2004). Une autre méthode probabiliste, similaire au maximum de vraisemblance et qui diffère par l'utilisation d'une distribution a priori, est plus rapide et permet de traiter davantage de données : c'est l'inférence bayésienne (Delsuc et Douzery 2004).

Nous voilà donc équipé pour classer les espèces. Mais posons la question de l'apparition de ces dernières. L'apparition d'une nouvelle espèce est appelée spéciation, et regardons de plus près comment cela arrive.

Il existe deux principaux "moteurs de l'évolution" : le premier est la dérive génétique (ce sont les événements aléatoires, qui surviennent au sein d'une population et qui ont pour effet de modifier la fréquence d'un allèle ou d'un génotype), le second est la sélection naturelle (il s'agit ici des avantages ou désavantages reproductifs

procurés par la présence ou l'absence de variation génétique). Ces moteurs agissent en permanence au sein des populations.

Si l'on vient y ajouter les mutations (modifications lors de réplication des gènes, qui peuvent être transmises à la descendance, si elles interviennent dans la lignée germinale) et la migration (augmentation de la variabilité génétique, par de nouveaux arrivants d'une population éloignée), nous avons ici les quatre forces évolutives qui font et défont les espèces.

Aussi, au cours des générations qui se succèdent, des individus d'une population peuvent se retrouver isolés des autres, par la formation d'une barrière physique (cours d'eau, montagne ...) et ainsi, évoluer indépendamment des premiers individus. Au fur et à mesure des générations, et à cause des forces évolutives précédemment citées (particulièrement la sélection naturelle et la dérive génétique), un isolement reproductif entre les individus peut conduire à l'apparition d'une nouvelle espèce. Il s'agit ici d'une spéciation allopatrique (figure 9), qui est celle, retrouvée majoritairement chez les animaux. C'est le cas du coléoptère *Phoberus capensis* d'Afrique du Sud qui a connu une spéciation allopatrique induite par le climat. Ces insectes incapables de voler ont évolué différemment entre les forêts tempérées d'Afrique du Sud et les hauts sommets montagneux (Strümpher *et al.*, 2016).

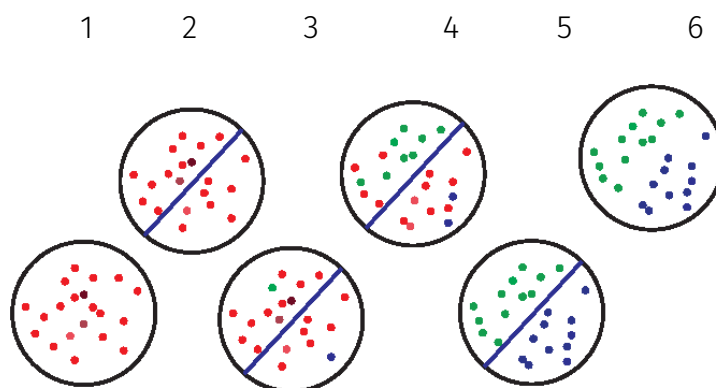


Figure 9 : Spéciation allopatrique

Une espèce vit sur un territoire donné (1). Une barrière naturelle apparaît, montagne ou rivière (2). Au fil des générations des mutations apparaissent (3) et s'accumulent (4) dans les deux sous-populations formées. Jusqu'à ce que l'isolement reproductif soit complet (5), même si la barrière naturelle disparaît (6).

D'autres modes de spéciation existent, la colonisation d'une nouvelle niche écologique par la population de base. Que cette niche soit isolée ou non, peut conduire à l'apparition d'une nouvelle espèce (respectivement péripatrique et parapatricque). Un exemple de spéciation péripatrique est la colonisation des îles de l'archipel d'Hawaï par le genre *Drosophilae* (Kaneshiro and Boake, 1987), au Japon un complexe d'espèces de coléoptère *Chrysolina angusticollis* vivant en parapatricque présente un isolement sexuel post-zygotique conduisant à une spéciation (Katakura *et al.*, 1996)

Enfin un dernier mode, qui reste controversé, est la spéciation sympatrique, où la nouvelle espèce apparaît, au sein d'une population alors qu'il n'y a pas d'isolement géographique (Berlocher and Feder, 2002 - figure 10).

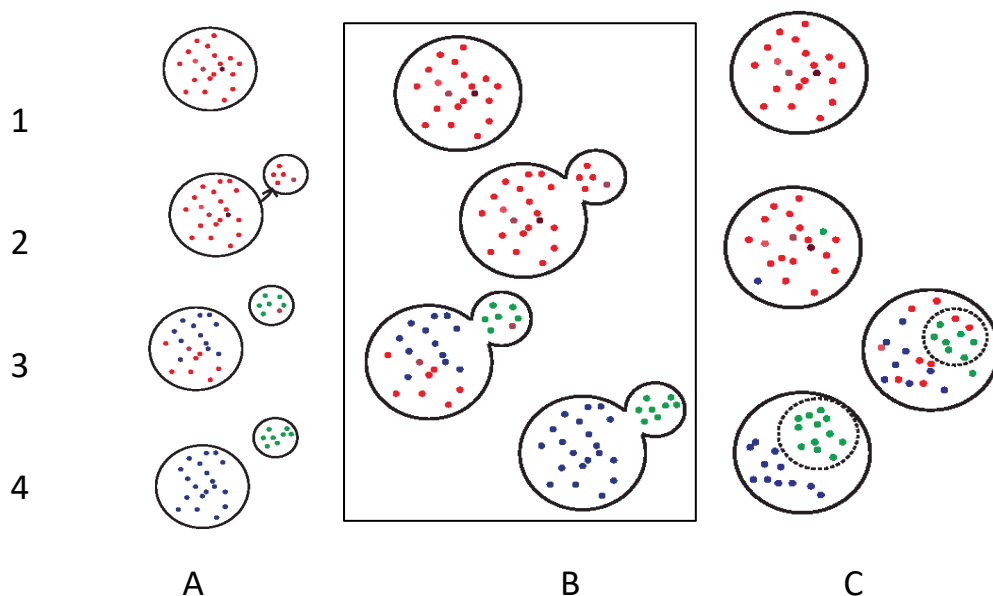


Figure 10 : Spéciation péripatrique (A), parapatricque (B) et sympatrique (C)

- A : Spéciation péripatrique : Une espèce vit sur un territoire donné (1). Une sous population colonise un territoire isolé du premier (2). Les mutations s'accumulent au sein des deux populations formées (3) jusqu'à ce que l'isolement reproductif soit total (4)
- B : Spéciation parapatricque : Une espèce vit sur un territoire donné (1). Une sous population colonise un territoire connecté du premier (2). Les mutations s'accumulent au sein des deux populations formées (3) jusqu'à ce que l'isolement reproductif soit total (4)
- C : Spéciation sympatrique : Une espèce vit sur un territoire donné (1). Des mutations commencent à s'accumuler (2). Un isolement reproductif arrive tandis que les deux populations sont encore pleinement en contact (3) jusqu'à ce que l'isolement reproductif soit total (4)

Maintenant que la notion d'espèces est acquise et que l'apparition de nouvelles espèces est définie, nous pouvons accéder au premier regroupement d'espèces que l'on appelle le genre, ce qu'il faut directement dire ici, c'est que le genre est une construction humaine basée sur des critères plus ou moins arbitraires. Et que selon les critères choisis, le taxon peut s'avérer être paraphylétique, comme nous l'avons vu précédemment.

Puis, par la suite les genres sont regroupés en famille, les familles en ordre, les ordres en classe, les classes en embranchement, les embranchements en règne. Cette classification, encore utilisée aujourd'hui a été définie par Linné au 18ème siècle.

L'ensemble de ces taxons peut également avoir des rangs supplémentaires, dit intercalaire, afin d'essayer de faire des regroupements toujours plus précis du monde vivant.

Vous l'avez compris, bien que les critères employés soient nombreux et variés, les regroupements du vivant ne sont qu'une simple construction de l'humain. D'ailleurs pour satisfaire notre anthropocentrisme voici sa classification en 7 rangs :

Règne : Animalia  
Embranchement : Chordata  
Classe : Mammalia  
Ordre : Primates  
Famille : Hominidae  
Genre : *Homo*  
Espèce : *Sapiens*

Mais revenons à nos insectes, seul le règne est partagé avec *Homo sapiens*, dès l'embranchement les insectes partent dans une autre boîte (en réalité c'est même à l'infra-règne - rang taxonomique intermédiaire du règne et l'embranchement - que les deux divergent, l'un étant un protostomien, l'autre un deutérostomien).

Les insectes sont donc des animaux (règne), ils sont pluricellulaires et hétérotrophes, ce qui en fait des eumétazoaires, ils sont bilatériens, protostomiens et évoluent par mues successives, ce qui



en fait des ecdysozoaires. Ils font partie des arthropodes (embranchement).

Les insectes possèdent des appendices caractéristiques au niveau de la tête des mandibules. En ajoutant les informations que nous avons déjà, à savoir le corps en trois parties et chitinisé, la respiration au moyen de trachées respiratoires ainsi que des pièces buccales externes nous avons une description complète des insectes (Classe) (figure 11).

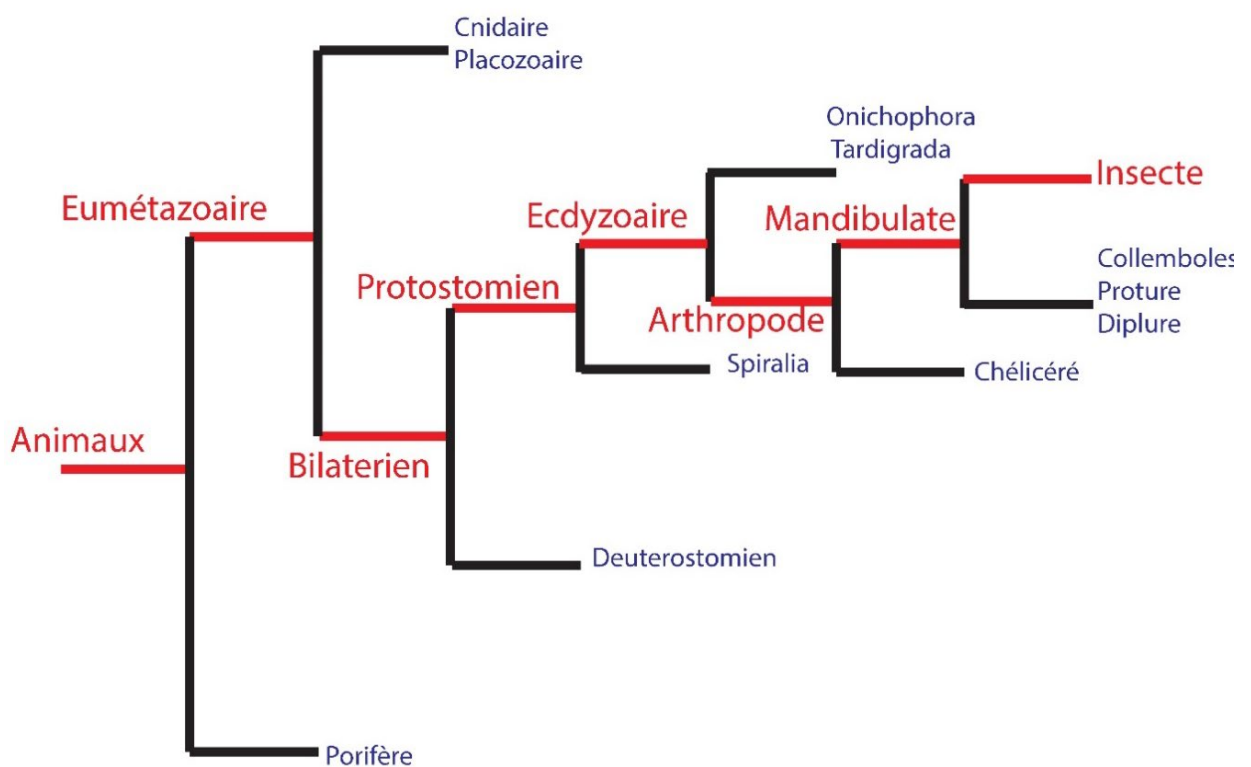


Figure 11 : Cladogramme des animaux aux insectes.

Les insectes sont des animaux et plus précisément des eumétazoaires qui sont dit animaux vrai, ces derniers se différencient des porifères (éponges) en ayant plusieurs types de tissus. Parmi les eumétazoaires, on retrouve les bilatériens qui ont une symétrie bilatérale contrairement aux cnidaires (méduses). Les insectes sont également protostomien, c'est-à-dire qu'ils ont la bouche qui se forme avant l'anus dans leur développement, ce qui est l'inverse pour les deutérostomiens (les mammifères par exemple). Ils évoluent par mue successive ce qui en fait des ecdysozoaires. Les insectes possèdent des appendices articulés (ce qui exclut les tardigrades), ils ont des mandibules et non des chélicères (comme les araignées). Enfin, contrairement aux collemboles et apparentés, les insectes ont des pièces buccales externe.

Nous avons ici parcouru les caractéristiques allant du plus général au plus particulier. C'est cela que nous apporte la phylogénie, des caractères qui permettent de dire quels individus sont les plus apparentés, en fonction de la présence ou de l'absence de ces caractéristiques.

Tous les animaux ayant une respiration via des trachées respiratoires sont plus apparentés que tous ceux qui présentent un côté droit et un côté gauche (bilatériens).

Poursuivons maintenant dans cette phylogénie, et intéressons-nous aux taxons de rangs inférieurs (figure 12). Le premier taxon que nous laissons de côté sont les aptérygotes, tel que *Lepisma saccharina*.

Ces derniers s'opposent aux ptérygotes, insectes présentant un état adulte ailé (deux paires d'ailes). Certaines espèces ont pu perdre leurs ailes au cours de l'évolution, tel que les poux (*Phthirus pubis* par exemple) ou seulement une partie des représentants de l'espèce, tel que les cochenilles mâles ou les fourmis ouvrières. Ils se peut aussi qu'ils aient une paire d'ailes atrophiée tels que les diptères (mouches, moustiques)

Parmi les ptérygotes, ceux dont les ailes ne peuvent pas se replier en arrière au repos (telles que les libellules) sont qualifiés de paléoptères, tandis que les autres insectes sont des néoptères.

Les néoptères forment l'immense majorité des insectes que l'on rencontre aujourd'hui. Parmi eux nous retrouvons ceux dont nous avons eu l'occasion de parler précédemment, à savoir ceux qui ont un développement de type holométabole. (Nous laissons donc de côté désormais les hétérométaboles, où nous pouvons retrouver par exemple les mantes religieuses et les punaises).

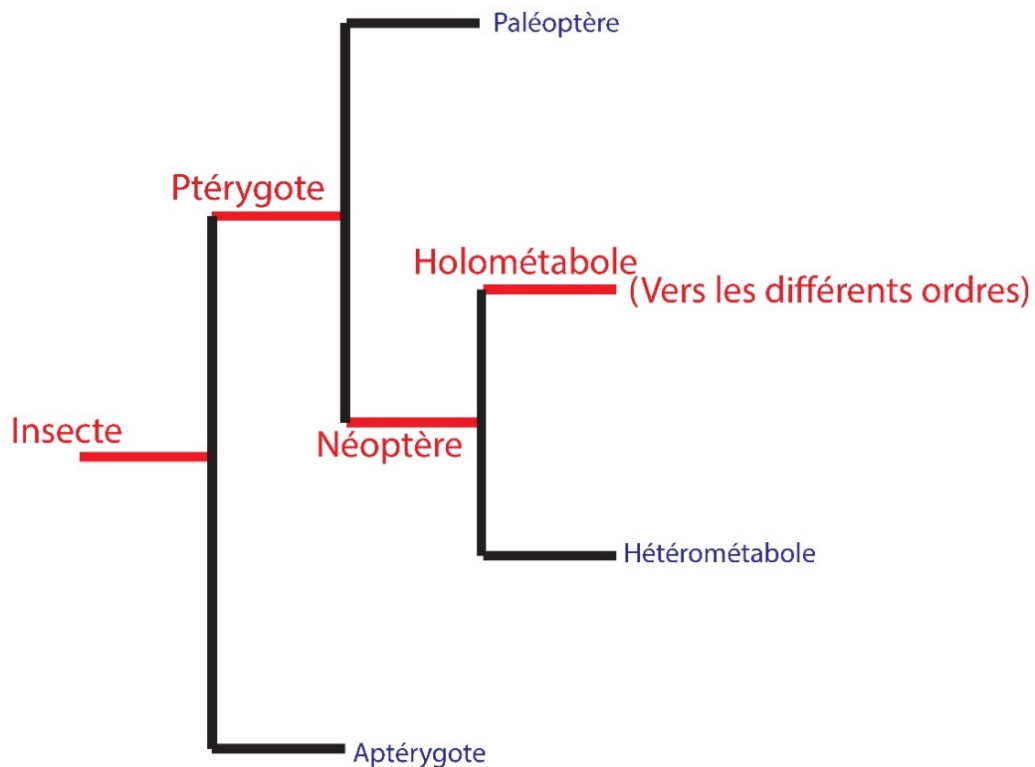


Figure 12 : Cladogramme des insectes vers les ordres.

*Les insectes peuvent ne pas avoir d'ailes (les aptérygotes) mais la majorité en possède. Si ces ailes peuvent se replier en arrière au repos, ils sont des néoptères, contrairement aux paléoptères (comme les libellules) et ils peuvent être holométales ou hétérométabole.*

C'est dans ce taxon-là, les holométales, que nous allons voir quelques noms qui parleront au plus grand nombre, tel que les hyménoptères (abeilles, fourmis, guêpes), les lépidoptères (papillons) les diptères (mouches moustiques et syrphes), les coléoptères (hannetons, coccinelles). Ces noms de taxons correspondent tous à des Ordres d'insectes.

Voici donc pour les principaux ordres des Insectes, et également une petite observation du besoin impérieux d'*Homo sapiens*, de créer toujours plus de boîtes, pour y ranger ses colocataires.

## Diptères et mouches soldats

---

*Les mouches bourdonnaient sur ce ventre putride,  
D'où sortaient de noirs bataillons  
De larves, qui coulaient comme un épais liquide  
Le long de ces vivants haillons.*

*Baudelaire - Les fleurs du mal - Une charogne.*

---

Maintenant que nous avons vu la classification du vivant, au cours du chapitre précédent, nous allons nous arrêter sur l'un des ordres en particulier, celui des Diptères. L'origine du mot diptère, vient du grec (pteron : aile) et signifie deux ailes, cependant comme tous les ptérygotes, les diptères ont bien quatre ailes. Chez ces derniers l'une des deux paires, la postérieure, est atrophiée. Elles sont transformées en un balancier (ailes en haltères), appelées licols, assurant la stabilité au cours du vol (figure 13).



Figure 13 : Ailes en haltères d'un Tipule (vue dorsale).

*Les deux ailes en haltères d'un Tipule sont pointées avec les flèche rouge, elles sont essentielles dans la stabilisation du vol chez les diptères (Crédit photo : André Lequet - 2013).*





Figure 15 : *Ornidia obesa*.

Gail Hampshire CC BY 2.0

Ils sont également les principaux insectes coprophages, détritiphages et nécrophages.

Par ailleurs, ces animaux sont aussi responsables de transmission de maladies à l'homme (Les moustiques sont vecteurs de la dengue, du virus du Zika, ou encore du paludisme) ou aux animaux (la mouche charbonneuse, *Stomoxys calcitrans*, transmet le virus de l'anémie infectieuse équine). Ils peuvent également être des ravageurs de cultures (*Psila rosae*, la mouche de la carotte par exemple). La diversité de l'ordre des diptères reste malgré tout, peu étudiée si l'on excepte un nombre limité d'espèces modèles comme la drosophile ainsi que des espèces transmettant des maladies ou ravageant les cultures.

Sur le plan écologique, les diptères représentent le taxon le plus diversifié et compte tenu de ce que nous avons dit juste avant, ils ont probablement un impact économique plus fort sur l'être humain que tous les autres ordres d'insectes (Kongyeli Badii, 2020). Ils

représentent 10 à 15% des espèces animales connues, répartis dans environ 17 000 genres, 190 familles, 28 à 39 superfamilles et 12 à 15 infra ordres. Par ailleurs, quelques 3100 espèces fossiles ont été décrites (Yeates and Wiegmann, 2017).

Les Diptères sont le deuxième ordre d'insectes, en nombre d'espèces connues et décrites, après les Coléoptères, et juste devant les Lépidoptères et les Hyménoptères (Eggleton 2020).

Le régime alimentaire des diptères est très varié, il consiste essentiellement en un large éventail d'hématophagie et de parasitisme (à la fois endo- et ectoparasitisme) de vertébrés et d'invertébrés. Également, plusieurs formes de mycétophages, saprophages et phytophages. La plupart des espèces de cet ordre sont connues pour être présentes dans presque toutes les zones zoogéographiques du monde. Elles se sont avérées bien adaptables à une grande diversité d'habitats. Elles peuvent survivre dans presque tous les environnements habitables de la terre, à l'exception des profondeurs océaniques (Grimaldi and Engel, 2005).

Nous avons déjà eu l'occasion de parler un peu avant de la monophylie des diptères, mais au sein de cet ordre, plusieurs regroupement traditionnellement admis s'avèrent être paraphylétique comme pour les nématocères. Cependant le caractère monophylétique des brachycères est quant à lui bien établi (Yeates et Wiegmann 1999).

Cette séparation arrive donc au niveau de l'infra-ordre, et c'est justement au cœur des brachycères que nous allons retrouver une famille en particulier : les *Stratiomyidae*.

Exit donc les moustiques, syrphes, drosophiles, taons et autres mignonneries tel que le moucheron des évier : *Clogmia albipunctata* (figure 16), et bonjour aux mouches soldats (figure 17).





Figure 16 : *Clogmia albipunctata*.

Ghent, Belgium CC BY SA 3.0

Peut-être n'êtes-vous pas bilingue en grec, mais "*Stratiomyidae*" vient du grec et signifie mouches soldat, décrite pour la première fois par un entomologiste Français, Pierre-André Latreille, au début du 19<sup>ème</sup> siècle ([Latreille et al., 1802](#)), cette famille de diptères contient quelques 2900 espèces réparties dans environ 400 ordres. ([COL : Catalogue of life](#)).



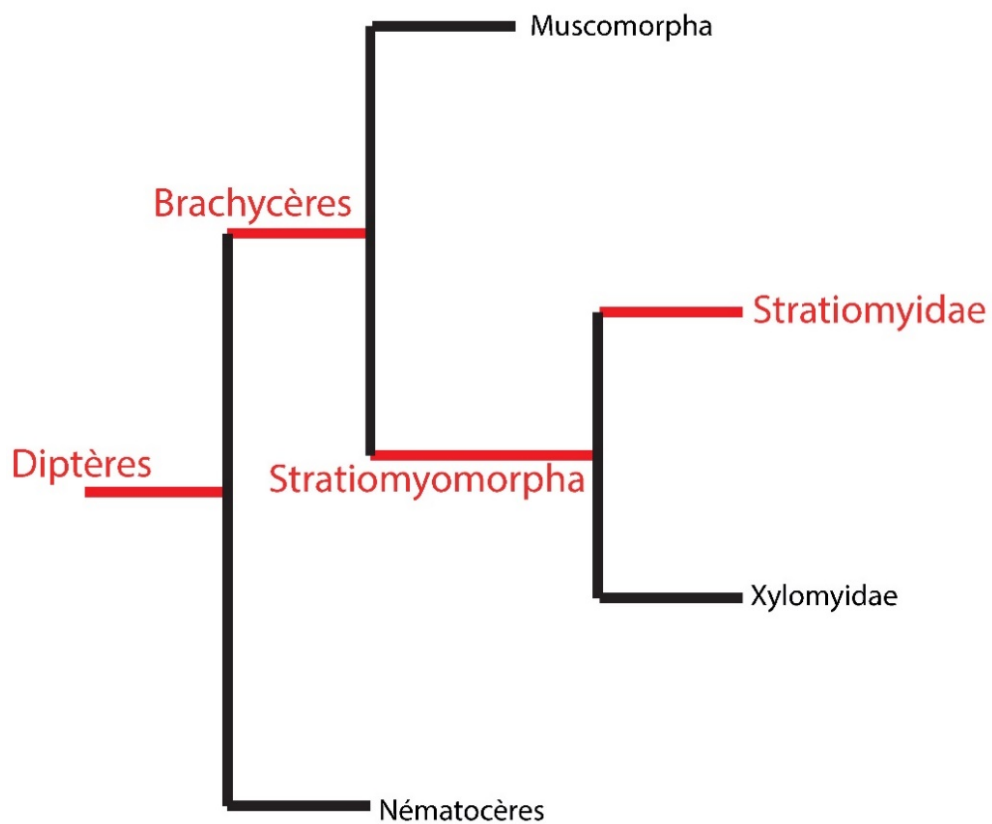


Figure 17 : Cladogramme des diptères au Stratiomyidae.

*Les brachycères possèdent des antennes courtes, ils ont soit une forme de Stratiomyidae, les Stratiomyomorpha, soit une forme muscoïde (Muscomorpha).*

Décrite dans les mémoires de l'académie des sciences de l'académie impériale de France en 1860 (Tome 21 - Partie 1 - page 1243) comme des insectes à deux ailes ; à trompe charnue, rétractile, terminée par deux lèvres ; à antennes plus longues que la tête, coudées à leur base allongée, formant une sorte de fuseau sans poil isolé ; corps long à abdomen ovale, obtus, un peu déprimé et élargi ; corselet à écusson armé de deux pointes en arrière ; ailes croisées dans le repos (figure 18).

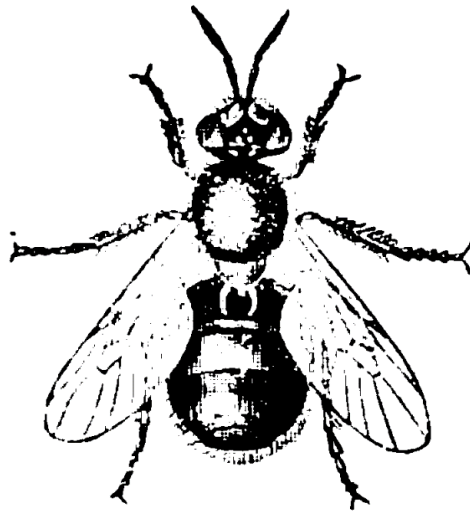


Figure 18 : Dessin d'observation d'un Stratiomyidae

(Mémoire de l'académie royale - p1284)

Environ la moitié des espèces de mouches soldats ont des larves aquatiques, et les adultes ont souvent tendance à mimer d'autres familles d'insectes (Woodley 2001).

Prenons par exemple les espèces du genre *Stratiomys* qui imitent relativement bien les guêpes ou abeilles (figure 19).



Figure 19 : *Stratiomys chamaeleon*

CC-BY SA 3.0

Mais avant d'aller plus loin, j'aimerais que l'on prenne le temps de se demander pourquoi d'ailleurs, cette famille s'appelle "soldat". Du reste, dans les autres langues les *Stratiomyidae* sont appelés « *soldier fly* » en anglais, « *moscas soldados* » en espagnol et « *Waffenfliegen* » (littéralement mouche d'arme, ou mouche armée) en allemand. Qu'est-ce que ces mouches, pour la plupart inoffensives pour l'homme, ont bien pu faire pour mériter une telle dénomination ?

A en croire le Professeur L.C. Miall dans "The natural history of aquatic insect", 1912, p 196, "From the back of the thorax two spines stand out towards the abdomen, and hence the name *Stratiomys*, which means the "armed Fly." Le nom de soldat viendrait donc de la présence de ces deux épines sur le scutellum de ces mouches, comme on peut le voir très bien sur la figure 20.



Figure 20 : *Odontomyia ornata*.

Université de Wageningen CC-BY 3.0. La présence des deux épines sur le scutellum (marquées d'une flèche rouge) donnent le nom à nos mouches soldats.

Toujours est-il qu'à donner un nom militaire à cette famille d'insectes, cela caractérise leurs noms vernaculaires. Voici quelques exemples, qui j'en suis sûr, égayeront votre prochain dîner mondain : *Stratiomys potamida* (le général rayé), *Odontomyia argentata* (Silver Colonel), *Odontomyia hydroleon* (Barred Green Colonel), *Odontomyia ornata* (Ornate Brigadier – figure 20), sans oublier les mouches soldat noires (*Hermetia illucens*), mouches soldat bleues (*Exaireta spinigera*), les mouches soldat jaunes (*Inopus flavus*) et mouches soldat de la canne à sucre (*Inopus rubriceps*) (Bull 1976)

Mais c'est aussi au sein de cette famille que l'on trouve l'animal dont le nom scientifique est le plus long à savoir : *Parastratiosphecomyia stratiosphecomyioides*.

Les *Stratiomyidae* sont un groupe relativement primitif, et certains insectes piégés dans l'ambre ont été datés du crétacée moyen (Liu *et al* 2016).

L'ancêtre commun de toute cette armée de *Stratiomyidae* aurait, selon les dernières estimations, vécu il y a quelques 143 millions d'années (Brammer & Von Dohlen 2007)

D'un point de vue écologique, les *Stratiomyidae* présentent une large variété de comportements et d'habitats. Les larves sont majoritairement détritivores, même si celles qui sont aquatiques peuvent se nourrir d'algues. Plus anecdotiquement, elles peuvent être herbivores ou même prédatrices, tel que *Zabrachia minutissima* qui est un prédateur des scolytes (Kenis *et al* 2004)

Les larves ayant un mode de vie terrestre sont très souvent retrouvées dans des matières en décomposition ou encore dans des excréments d'animaux.

*Inopus rubriceps* est un ravageur de culture des cannes à sucre et certaines vivent sous l'écorce des arbres (Gerard and Burton 1983).

Chez les *Stratiomyidae*, le nombre de mues successives peut atteindre 10 à 12 chez *Inopus rubriceps* (Gerard and Burton 1983) mais se situe en général autour de 6.

Comme tous les diptères, les *Stratiomyidae* sont des holométaboles, ils vont alors passer par une étape immobile qu'est la puppe. Une fois adultes, les mouches se nourrissent de nectar, ou peuvent ne pas se nourrir du tout, à l'instar d'*Hermetia illucens*.

Nous prendrons le temps au cours du chapitre suivant de détailler le cycle de vie d'un diptère, *Stratiomyidae* : *Hermetia illucens*.

## *Hermetia illucens*

---

Quarante-deux!  
H2G2

---

Elle est décrite, en 1758, par Linné dans son *Systema Naturae*, et rangée dans le règne animal parmi les diptères muscomorpha. Elle est décrite de la manière suivante :

“Antennes filamenteuses, avec des ailes bleues, massues, avec un corps noir, un segment d'abdomen translucide le long de leurs côtés. Elle vit en Amérique du Sud”

Il précise ensuite la description qu'en fait Charles de Geer (un biologiste suédois) : “Antennes longues, noires. Le deuxième segment de l'abdomen est glauque, transparent sur les côtés.” Pour information glauque en latin veut dire entre bleu et vert.



Figure 21 : *Hermetia illucens*

Crédit photo : Cycle Farms

En 1762, Linné, on lui pardonnera, décrit de nouveau la BSF et la nomme *Musca Leucopa*. Puis, c'est au tour de Latreille, lui-même qui avait décrit la famille des *Stratiomyidae* en 1802, de "ranger" la mouche soldat noire, dans le genre *Hermetia*, qu'il décrit en 1804.

La première vraie description de son cycle de vie se fera dans les années 1920, dans les pages du bulletin de la société entomologique d'argentine par Andrés Copello (Copello 1926) - Fait amusant, le bulletin avait été créé quelques mois auparavant en 1925.

Pendant de nombreuses années, *Hermetia illucens*, sera laissée de côté par la science, apparaissant par-ci par-là dans quelques textes, principalement des articles qui traitent de faune locale.

### Description des différents stades d'*Hermetia illucens* et Cycle de vie

Regardons donc, de plus près, ce que cette mouche a comme particularité.

#### Imago

Il s'agit de mouches de taille moyenne qui mesurent entre 15 et 20mm de long (May 1961), une couleur noire, avec des reflets métalliques bleu vert, et parfois légèrement marron au bout de l'abdomen. Deux ailes se replient l'une sur l'autre au-dessus du corps ainsi que des ailes en haltères (figure 22).

Comme tous les insectes, elle possède un corps en 3 parties : Tête, Thorax et Abdomen.

La tête porte des antennes, qui sont composées de trois segments et qui font environ deux fois la taille de la tête. Pour la vision, la mouche soldat possède trois ocelles et deux grands yeux composés (Oonincx et al 2016), légèrement plus écartés chez le mâle que chez la femelle. Chez l'adulte, les pièces buccales sont atrophiées, ce qui ne change rien au succès reproducteur mais il a été supposé que cela l'empêchait de se nourrir (Olivera et al 2016).

D'autres études au contraire, montrent que, bien que ces pièces buccales soient atrophiées, l'adulte reste capable de se nourrir, et possède un tractus digestif entièrement fonctionnel (Lupi et al 2019, Bruno et al 2019). Différents essais de nourriture ont été réalisés,

menant à la conclusion que l'adulte survivait plus longtemps avec de la nourriture sous forme liquide (ou juste de l'eau) mais que cela n'augmentait pas le nombre d'œufs pondus (Bertinetti *et al* 2019). Par ailleurs, l'adulte est capable, à l'instar d'autres espèces du genre *Musca*, de régurgiter de la salive, pour faciliter l'absorption d'aliments solides (Bruno *et al* 2019).

Un dimorphisme sexuel a d'ailleurs été trouvé au niveau des pièces buccales (Pezzi *et al* 2020).

Le thorax est la partie du corps, qui, comme chez tous les ptérygotes, porte les ailes, deux paires d'ailes. Cependant chez les diptères, la paire d'ailes postérieure est atrophiée et est dite en haltère. *Hermetia illucens* n'échappe pas à la règle. Son thorax est en trois parties, le premier segment, le prothorax, porte la première paire de pattes, le mésothorax, porte la deuxième paire de pattes ainsi que les ailes (celles responsables du vol à proprement parlé) et enfin le métathorax, troisième et dernier segment, porte la troisième paire de pattes ainsi que les ailes en haltère (figure 22).

Les pattes sont en trois parties, et ont un segment blanc. Les ailes antérieures sont membraneuses et possèdent des nervures, caractéristiques de la famille des *Stratiomyidae*.

Sur la face dorsale, entre le thorax et l'abdomen, il y a le scutellum, qui chez la plupart des *Stratiomyidae*, possède deux épines, qui donnèrent le nom aux mouches soldat. Les épines sont absentes du scutellum chez *Hermetia illucens* (Oliveira *et al* 2015).

L'abdomen est la troisième partie du corps des insectes. Chez *Hermetia illucens*, il est composé de 8 segments, le premier est celui qui porte une partie translucide (Oliveira *et al* 2016), ajoutant encore à son apparence élancée, lui donnant un aspect de "taille de guêpe". C'est de cette fenêtre que le nom latin "*illucens*" tire son origine. Le dernier segment qui porte les organes reproducteurs, est d'ailleurs la principale partie où s'exerce le dimorphisme sexuel (figure 23).



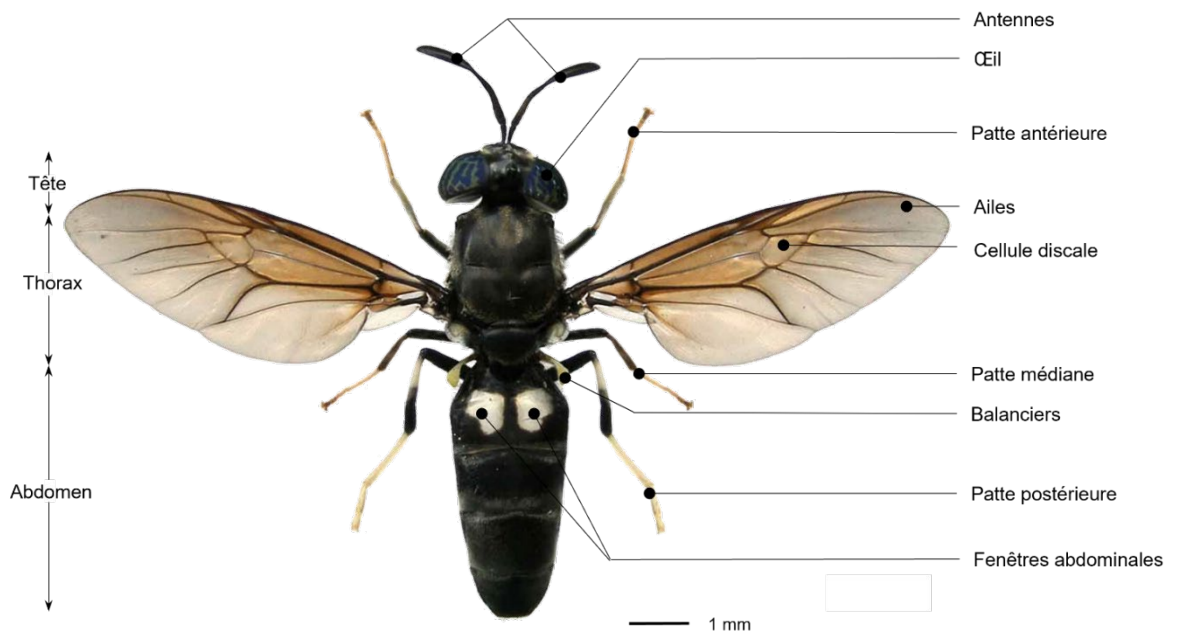


Figure 22 : Description des différentes parties d'une *Hermetia illucens* adulte

Avec l'aimable autorisation de Dominique Marité

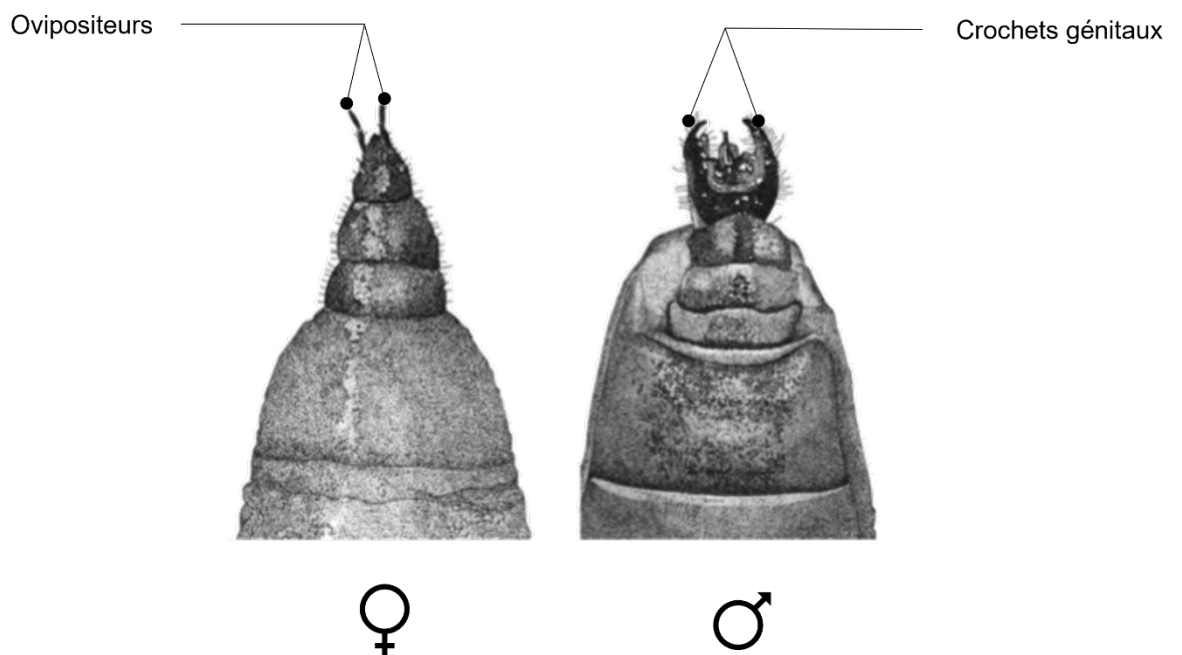


Figure 23 : Dimorphisme sexuel chez *Hermetia illucens*.

Adapté de Caruso et al 2014



Dans la nature, les adultes mâles sont principalement retrouvés en zone végétalisée, tandis que les femelles seront davantage au niveau des sites de pontes (près de matières en décomposition et installation pour volailles - Tomberlin 2001). Cela est principalement dû au comportement de reproduction que nous allons maintenant détailler.

### La reproduction

Au début de l'émergence, les imagos ne sont pas encore tout de suite prêts pour se reproduire, il faut environ deux jours avant que l'ensemble des pièces sexuelles soient matures (Tomberlin 2001).

Chez les BSF, la reproduction est un processus assez complexe, où les mâles vont avoir un comportement de lek afin de se reproduire avec les femelles. Ce comportement, se caractérise par une lutte entre les mâles pour un territoire donné, puis les femelles volent au-dessus des mâles posés au sol ou sur des feuilles. Un mâle va venir se "proposer" à la femelle en vol, si ce partenaire lui convient, alors la reproduction s'engage, pendant 20 minutes environ, temps pendant lequel le mâle va déposer ses spermatozoïdes dans les voies génitales de la femelle (Figure 24).

Les BSF sont polyandres, ce qui signifie que plusieurs mâles vont pouvoir féconder une même femelle (Hoffmann, 2021). La femelle possède trois spermathèques, mais nous ne savons pas encore s'il y a un choix qui est opéré par la femelle pour le placement des spermatozoïdes et le choix des œufs qui seront fécondés. Une fois la reproduction achevée, la femelle va chercher un site de ponte. Les odeurs de matières organiques en décomposition vont l'attirer et elle va pouvoir pondre. Par ailleurs, la lumière joue un rôle important dans le mécanisme de la reproduction avec des longueurs d'onde optimales dans les élevages (Awal Rabiul *et al.*, 2022).



Figure 24 : Reproduction chez *Hermetia illucens*.

Crédit photo : Cycle Farms

### Les œufs

Les femelles BSF pondent entre 200 et 600 œufs selon les différentes études. L'une d'entre elles (Jones et Tomberlin 2020) montre que les femelles plus grosses pondent plus d'œufs. Les œufs sont de petites tailles environ 1 mm, de forme légèrement oblongue (figure 25), un poids de 25 µg environ et de couleur blanc crème. Après 4 à 14 jours selon la température et le taux d'humidité (May 1961) les larves de premier stade vont pouvoir émerger.



Figure 25 : Œufs d'*Hermetia illucens*

Crédit photo : Cycle Farms

### Les larves

Les larves de BSF possèdent 12 segments (figure 26) en incluant la capsule céphalique et vont passer par 6 stades larvaires (figure 27 A-

F), ayant à chaque fois une mue (Kim *et al* 2010 ; Barros *et al* 2018). Selon les conditions environnementales (température, humidité, accessibilité à la nourriture), les différents stades larvaires s'enchaînent plus ou moins vite. Il faut environ 14 jours pour passer du premier instar au sixième, et la taille va passer de 1.8mm à 20mm (Liu *et al* 2019). Les larves commencent à se nourrir immédiatement après l'éclosion avec une consommation de plus en plus importante, surtout après le troisième stade larvaire (Liu *et al* 2019). Le poids des larves en 14 jours augmente d'un facteur 10000.

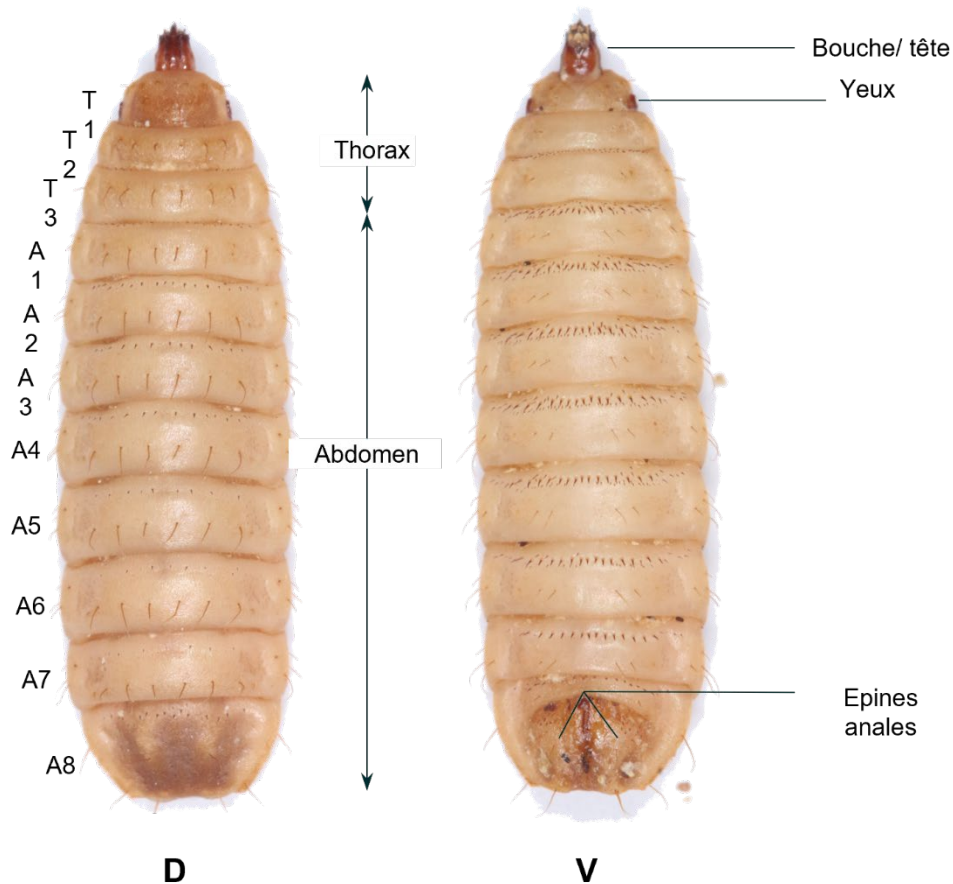


Figure 26 : Description des différentes parties d'une larve d'*Hermetia illucens*

Crédit Cycle Farms



Figure 27 : Evolution des larves d'*Hermetia illucens*

Visualisation des 6 instars (A à F) et de la pupa (G) - Crédit Cycle Farms

Lors du dernier stade larvaire a lieu la pupaison (figure 27 – G), il s’agit d’une étape où les larves cessent de se nourrir, vont changer de milieu, pour rechercher un endroit sec, et vont “durcir” leur exosquelette, qui va de fait, s’assombrir.

### Les pupes

La pupaison requiert environ 2 semaines (Hall and Gerhardt 2002). Comme chez la plupart des holométaboles, la pupa est immobile, et c’est la prochaine mue, qualifiée de mue imaginale, qui permettra l’émergence de l’adulte. De façon générale, les mâles émergent avant les femelles, ce qui peut permettre d’éviter la consanguinité, les mâles s’éloignant alors des sites d’émergences. Les femelles, qui émergent légèrement plus tard, sont en général plus grosses que les mâles.

### Cycle de vie et facteurs influant

Le cycle de vie d'*Hermetia illucens* peut donc être résumé en visualisant le passage des différents stades (figure 28).

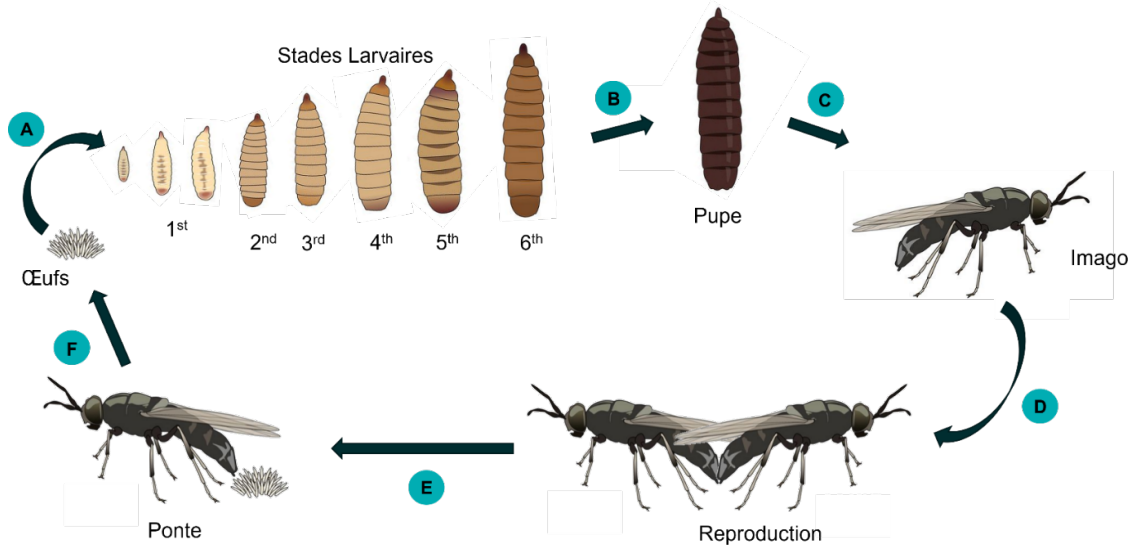


Figure 28 : Résumé du cycle de vie d'*Hermetia illucens*

Crédit : Cycle Farms

Certaines substances ou certaines conditions vont venir altérer ce cycle de vie, et cela que ce soit de façon négative :

- La cyromazine et le Pyriproxyfen, sont deux insecticides, qui, à un certain dosage, vont tuer les larves. (Tomberlin *et al* 2002).
- Le pH aura aussi un impact dans les premiers jours de développement des larves. Cependant, quel que soit le pH initial, dès le 4ème jour de développement, le pH du milieu se stabilise autour de 9 (Meneguz *et al* 2018).

Ou positive :

- L'augmentation de l'aération a pu tripler le poids sec de larves cultivées sur des coproduits d'amande (Palma *et al* 2018).
- L'ajout de probiotiques, telles que les bactéries du genre *Rhodococcus*, peut également augmenter le poids des larves et la bioconversion des déchets (Franks *et al* 2021). *Bacillus velezensis* ajouté au substrat a permis une augmentation du taux de protéine et de la survie des larves de BSF (Pei *et al.*, 2022).

Cependant la mouche soldat va pouvoir résister à certains pathogènes, voire en réduire les proportions. Ainsi, dans un milieu dans lequel se développent des larves de BSF, les proportions d'*Escherichia coli* et de *Salmonella enterica*, vont être réduites (Liu et al 2008 ; Erickson et al 2003). En revanche, ce n'est pas le cas de tous les pathogènes tels que des *Enterococcus* ou encore du bactériophage  $\phi$ X174 (Lalander et al 2013)

Par ailleurs les larves de BSF, sont capables sans modifications de leur cycle de vie, de bio accumuler certains métaux lourds tel que le cadmium et l'arsenic, de résister à toute une variété de mycotoxines, de pesticides ainsi qu'à certains produits pharmaceutiques (Lievens et al 2021 ; Camenzuli et al 2018 ; Diener et al 2015).

### Répartition de la mouche soldat noire

Un faisceau de données pointe une origine Sud-Américaine de l'espèce (nord de l'Amérique du Sud et Amérique centrale). Les apports récents de la génétique des populations de l'espèce qui seront développés dans le chapitre 02 de cette thèse (Hauser and Woodley, 2015 ; Kaya et al 2021) ont confirmé cette origine.

Retenons néanmoins qu'une hypothèse possible est que, profitant des voyages humains, elle s'est propagée sur l'ensemble du globe, s'acclimatant à quasiment toutes les conditions climatiques (Rozkošný 1983). Elle est retrouvée aujourd'hui sur tous les continents à l'exception de l'antarctique. Cependant, plusieurs observations contredisent le scénario de « l'auto-stop ». Par exemple, on a retrouvé dans le sarcophage d'Isabelle de Naples (1470 - 1524) une larve de mouche soldat noire (Benelli et al 2014). Or il a été proposé que son arrivée en Europe, à Malte, n'ait eu lieu qu'en 1926 (Lindner 1936). Alternativement, elle pourrait s'être immiscée lors des premiers voyages transatlantiques en bateau, dès 1493 avec Christophe Colomb, comme le suggère Benelli et al (2014). Ces données anecdotiques montrent que le timing et le scénario de dispersion sont encore très mal connus.

Plus récemment de nouvelles localités ont été détectées sur une grande partie de l'Europe occidentale, portant le nombre de pays

Européens où elle est présente à 25, ainsi que dans la quasi-totalité des pays du bassin méditerranéen (Demetriou *et al.*, 2022).

Mais alors, comment passe-t-on d'un insecte dont on connaît le cycle de vie, la répartition, etc., à une utilisation industrielle ? La réponse se trouve, en partie, dans un article de 2002 (Sheppard *et al* 2002) où les auteurs ont décrit un système d'élevage des mouches soldat. La clé, ont-ils découvert, était de trouver le mélange précis de température, d'humidité et, surtout, d'éclairage pour stimuler la reproduction des mouches.

Et tout ça pour quoi ?

Mais avant de répondre à cette question, observons à nouveau à quel point nous aimons créer toujours plus de boîtes, à l'image de la classification complète d'*Hermetia illucens*.

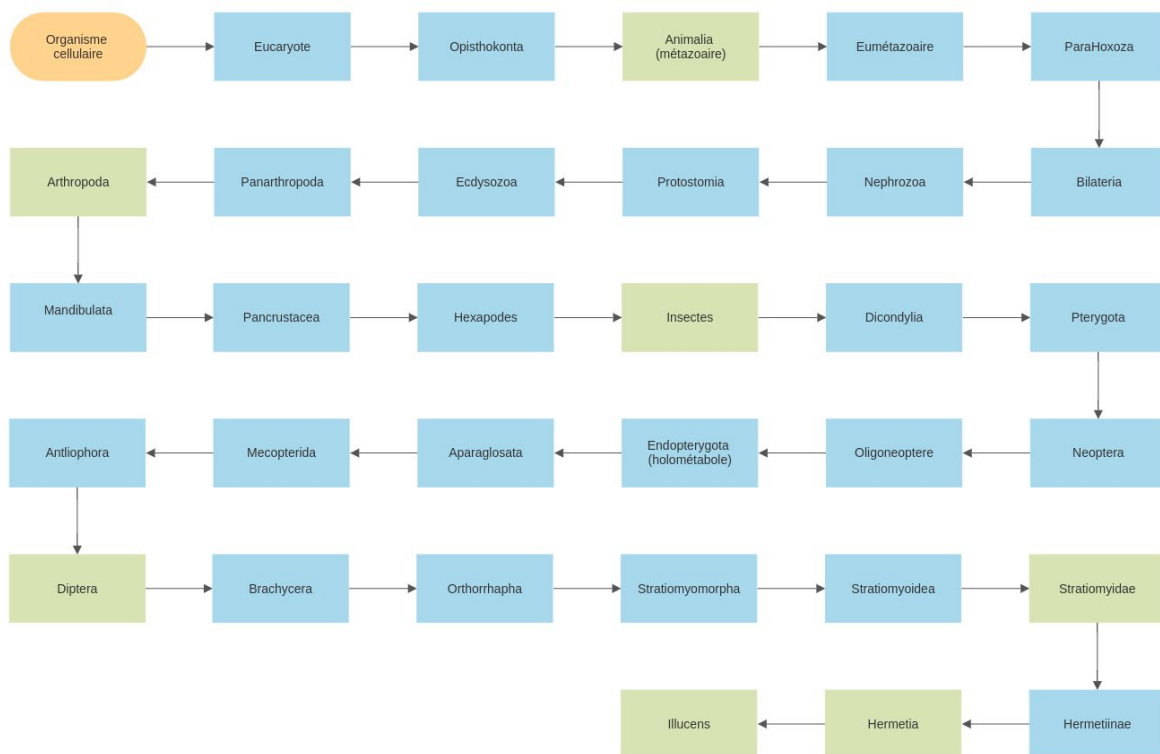


Figure 29 : Ensemble des rangs taxonomiques de la mouche soldat noire.

Les 7 rangs principaux sont en vert (dans l'ordre : Règne, Embranchement, Classe, Ordre, Famille, Genre, Espèce).  
Les rangs intermédiaires sont en bleu.



## Partie 2 : *Hermetia illucens* et *Homo sapiens*

---

*Dans cette nouvelle partie, nous allons nous arrêter sur l'insecte star de cette thèse, Hermetia illucens, et comment il interagit avec l'espèce de celui qui est en train d'écrire ces mots, ou de les lire, Homo sapiens. Nous nous intéresserons dans un premier temps à des aspects plus ou moins sympathiques de la vie d'un insecte nécrophage et détritivore, puis nous verrons ce que nous pouvons extraire de tout cela - littéralement et financièrement –*

*Mutualisme : Relation durable entre deux espèces ou deux populations, avantageuse pour les deux.*

*Larousse*

---



*L'étude.*

*Peinture générée par Craiyon*

*Craiyon, est un modèle d'IA capable de dessiner des images à partir de n'importe quel texte!*



## Entomologie médico-légale

---

*[Parlant à un cadavre dont des insectes sortent du ventre]*

*“Madame, je crois que vous avez besoin d’un exterminateur.”*

*Alexx Woods - Les experts Miami -*

---

Après avoir passé des dizaines d’années à apparaître dans la littérature scientifique, majoritairement pour signifier sa présence dans certaines parties du monde; au travers, notamment, de bulletins de sociétés entomologiques diverses et variées, au Colorado et en Utah par exemple (James 1936), la reine de cette thèse, va commencer à faire parler d’elle dans un domaine assez éloigné de celui sur lequel on l’attendait : La datation de l’intervalle post mortem.

En 1982, deux scientifiques rapportent la découverte d’une espèce d’*Hermetia* dans un corps en décomposition (Lee and Cheong 1982). En 1994, un scientifique du FBI publie le premier papier traitant de l’utilisation de la mouche soldat noire, comme insecte permettant de dater l’intervalle post-mortem (Lord et al 1994). Évidemment la BSF n’est pas le premier insecte utilisé en entomologie médico-légal, ni même le premier diptère (Catts et Goff 1992). Sa présence sur une grande partie du globe permet, en revanche, d’avoir des informations qui sont utilisables par un grand nombre d’enquêteurs.

La mouche soldat noire va, en effet, venir pondre ses œufs dans les 20 à 30 jours qui suivent un décès (Lord et al 1994), et en fonction des conditions de température et d’humidité, on pourra, au moment de la découverte du corps, donner l’âge des larves et ainsi estimer l’intervalle post-mortem.

Des exemples d’utilisation de la présence de BSF dans des corps ont été donnés dans la littérature scientifique. En 2001 en Italie, pour la première fois en Europe (Truchetto et al 2001). En 2008 au Brésil (Pujol-Luz et al 2008) et 2011 en Espagne (Martinez-Sanchez et al 2011).

Dans les cas présentés en Europe, les auteurs indiquent à chaque fois que la présence d'*Hermetia illucens* est non seulement un indicateur utilisable, mais que par ailleurs son interaction avec les autres espèces d'insectes (par compétition) doit être pris en compte pour le décompte de l'intervalle post-mortem. Ces informations avaient d'ailleurs déjà été utilisées dans le cadre de lutte biologique contre les mouches domestiques (Furman *et al* 1959 ; Bradley et Sheppard 1984).

Et alors, me direz-vous, de quel genre de cadavre s'amourache *Hermetia illucens* ? Et bien il semblerait que les BSF aient une préférence pour les cadavres en état de décomposition avancée et secs, je vous épargne donc les images trop explicites (figure 30).

## Datation de l'intervalle post - mortem

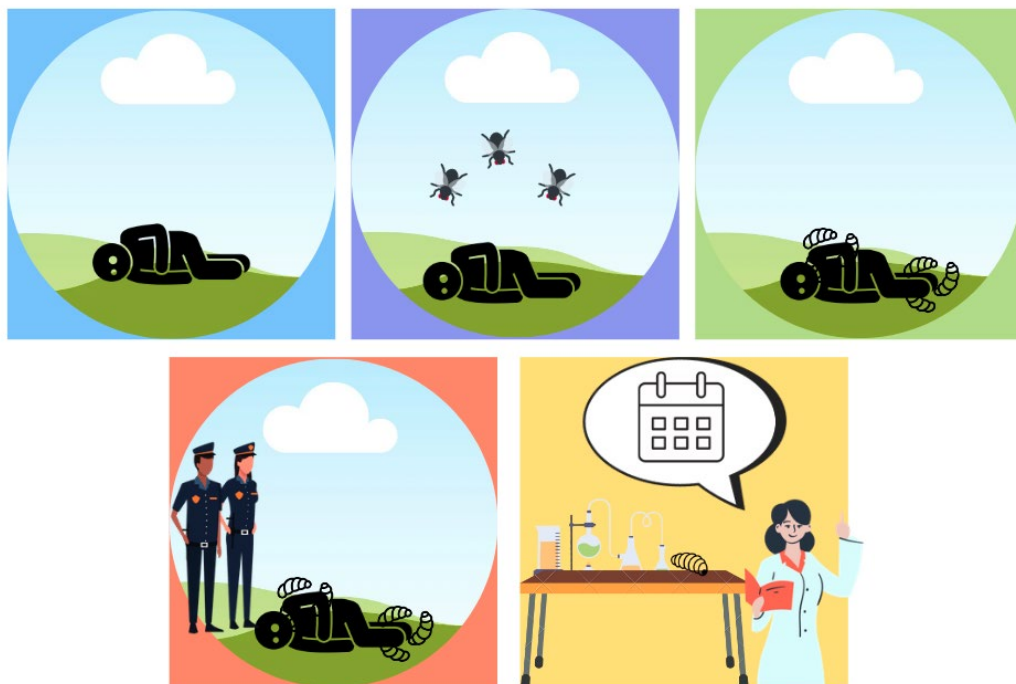


Figure 30 : Datation de l'intervalle post-mortem – reconstruction

Un corps est dans un environnement sec et en décomposition depuis 20 à 30 jours (1) c'est à ce moment-là que des *Hermetia illucens* adultes vont venir pondre leurs œufs (2). Les larves se développent (3). Quand les autorités vont trouver le corps (4) les larves seront analysées et en fonction de la température de l'environnement et de l'âge de ces dernières on pourra donner la date de la mort (5).

## Quid des déchets ?

---

“Eveeee”

Wall-E - Wall-E Disney Pixar 2008

---

Alors si la datation de l'intervalle post-mortem n'est finalement qu'une utilisation anecdotique de la BSF, il est un domaine où son utilisation est prise très au sérieux, la réduction et la transformation des déchets.

*Hermetia illucens* possède des capacités tout à fait remarquables lorsqu'il s'agit de réduire les biodéchets. Et dans un monde où l'on prévoit être autour de 9,7 milliards d'ici 2050 (United Nations, 2021), je vous laisse imaginer la quantité de déchets. En ne parlant ici que des déchets municipaux solides et à ce rythme de croissance, cela pourrait passer des 2,01 milliards de tonnes de déchets estimés en 2016 à quelques 3,4 milliards de tonnes de déchets municipaux solides en 2050. (Kaza et al 2018).

En effet, et à titre indicatif, on estime que chaque être humain produit entre 0,11 et 4,54 kg de déchets par jour (figure 31), et cela augmente dans les pays développés (Kaza et al 2018).

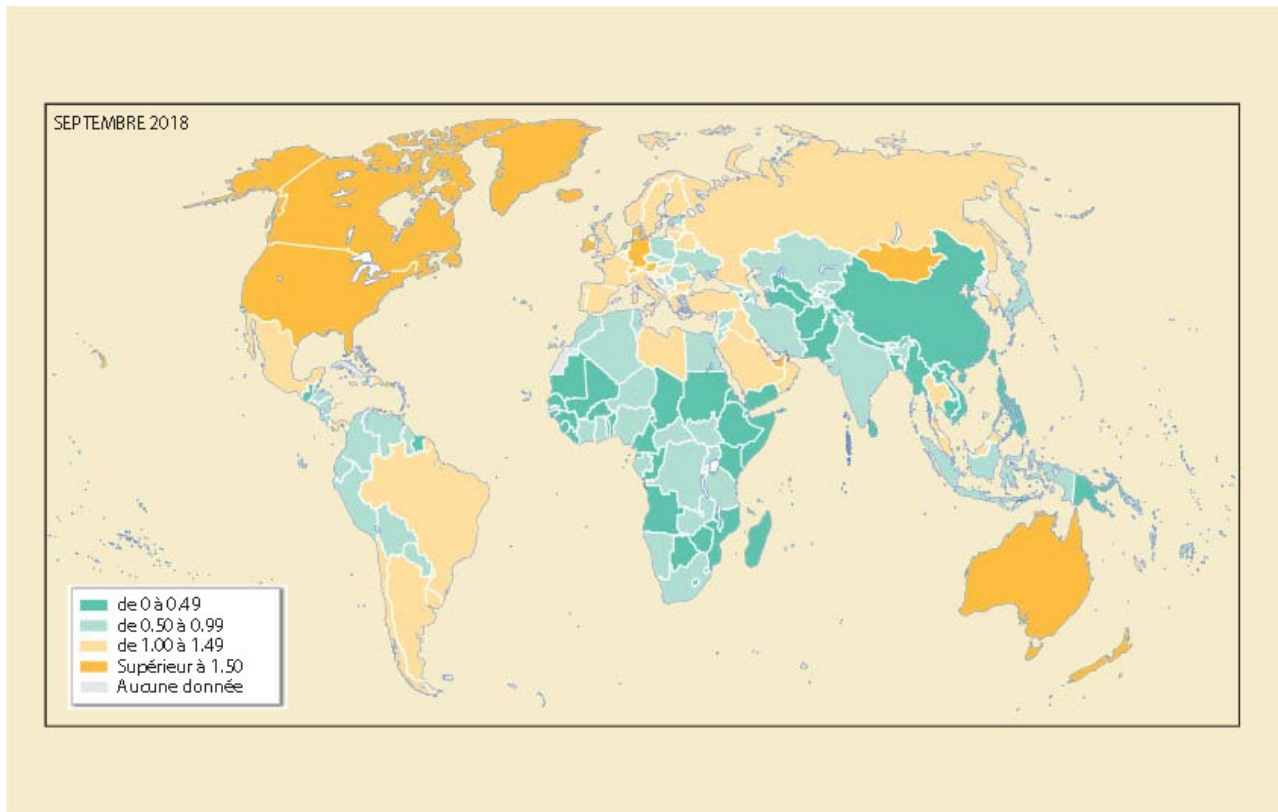


Figure 31 : quantité de déchet par habitant et par jour

La quantité de déchet produite par habitant et par jour est donnée en kilogramme. Elle varie de 0,11 kg par jour et par habitant dans les régions sub-saharienne à 4,54 kg par jour et par habitant en Amérique du Nord. La France est à 1,38 kg par jour et par habitant.  
(Adapté de Kaza et al 2018)

Cependant, il est également important de prendre en compte les régions où sont générés les déchets, car bien qu'un habitant d'Amérique du Nord ou d'Europe génère plus de déchet au quotidien qu'un habitant de l'Asie de l'Est et du Pacifique, le nombre d'habitants est bien plus important dans ces régions (figure 32).

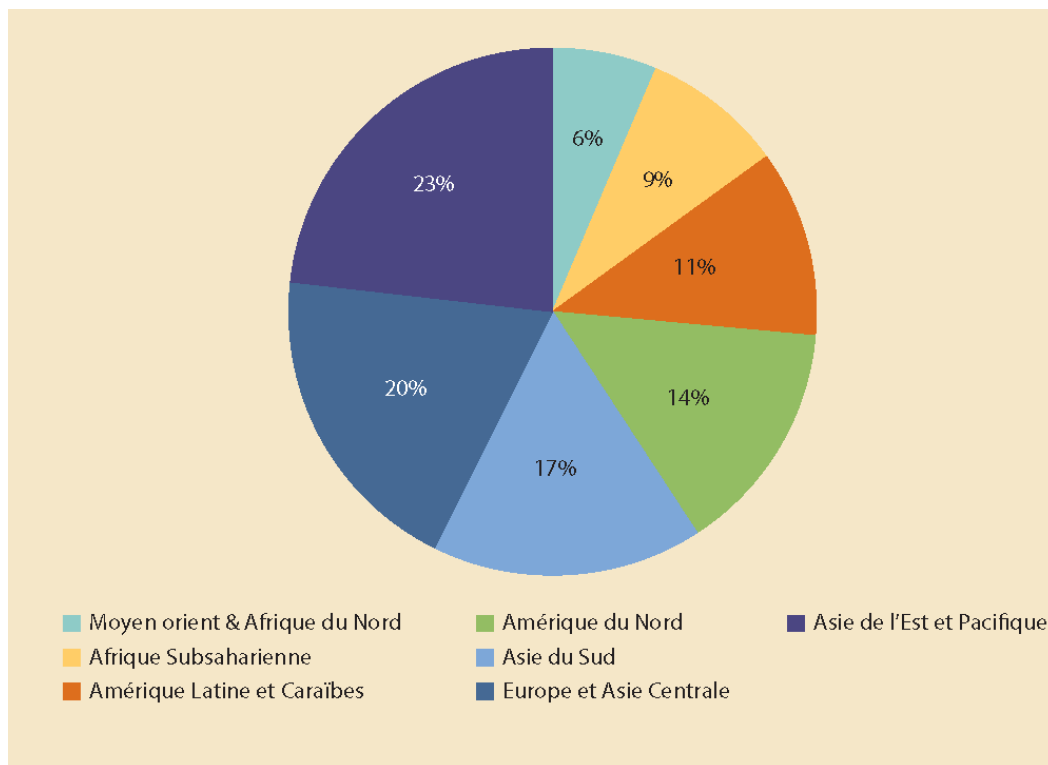


Figure 32 : Déchet généré par région du monde

Les pays d'Asie de l'Est et du Pacifique associé à l'Europe et à l'Asie centrale produisent 43 % des déchets du monde en termes de volume. Les régions du Moyen-Orient et de l'Afrique du Nord et de l'Afrique subsaharienne sont celles qui produisent le moins de déchets, représentant ensemble 15 % des déchets mondiaux. La région Asie de l'Est et Pacifique génère le plus de déchets en termes absolus, soit environ 468 millions de tonnes en 2016, et la région Moyen-Orient et Afrique du Nord en génère le moins, soit 129 millions de tonnes (adapté de Kaza et al 2018).

Au niveau mondial, près de la moitié de ces déchets sont organiques (Figure 33). La composition en déchet varie d'un pays à l'autre et est corrélée au PIB par habitant.

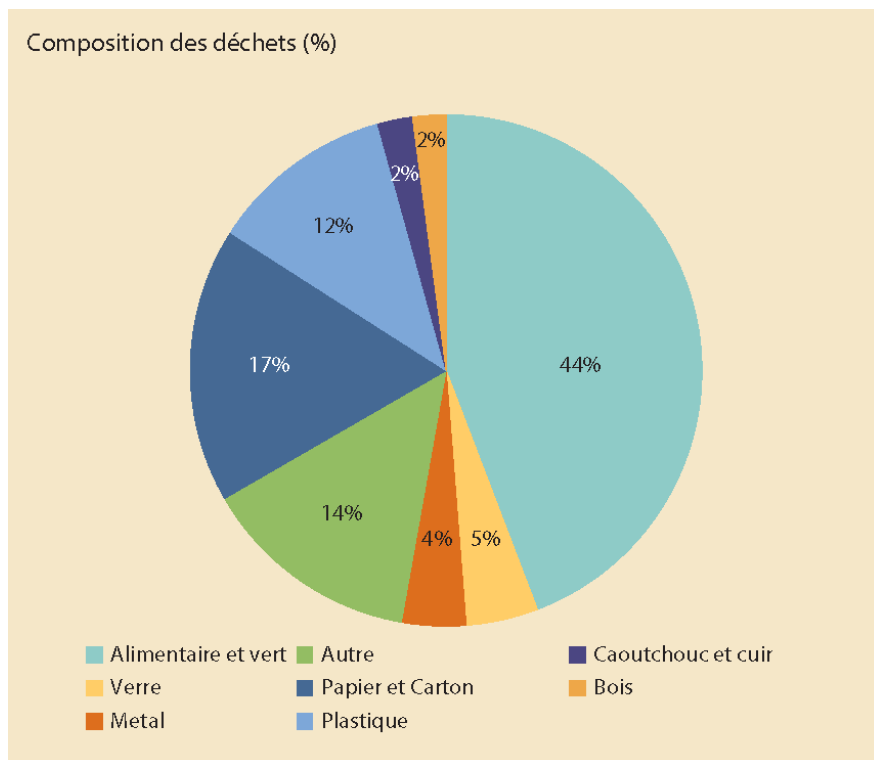


Figure 33 : Composition mondiale des déchets

Il y a donc, une corrélation directe entre le niveau de vie et la quantité de déchets produits (Kaza et al 2018). De manière générale, la quantité de déchet par habitant augmente avec l'augmentation du PIB (produit intérieur brut) par habitant.

En 2018, l'union européenne a produit 2,6 milliards de tonnes de déchets (déchets provenant à la fois des particuliers et des industries), soit, si on le rapporte au nombre d'habitants : 5,1 tonnes par habitant (SDES, 2020).

Par ailleurs, en France, chaque année 10 millions de tonnes de produits alimentaires sont perdues ou gaspillées, ce qui représente près d'un cinquième de la production (ADEME, 2019) et 100 millions de tonnes de déchets alimentaires pour l'Europe (Winkel et al 2016).

La gestion des déchets devient donc une étape clé pour les décennies à venir. D'autant plus qu'une mauvaise gestion des déchets, par souci d'économie par exemple, entraîne des problèmes sanitaires (produits toxiques, production de méthane, développement de la vermine ...) ainsi qu'un surcoût par rapport à ce que cela aurait coûté,

de bien les gérer directement (Hoornweg and Bhada-Tata, 2012). Actuellement seul 5,5% des déchets sont compostés et un tiers sont placés dans des décharges ouvertes (Kaza et al 2018).

Imaginez-donc que nous ayons sous la main un insecte qui se régale de matières en décomposition, et qui, étant donné son histoire évolutive, se trouve adapté à vivre dans ces déchets (comprenez bien ici, que ce ne sont pas des déchets qu'aux yeux de l'être humain, dans les élevages de BSF ils sont considérés comme matière première).

Et en effet, nous l'avons.

Ainsi en fonction des aliments qui seront donnés aux larves, celles-ci pourront consommer 37,3 % d'un aliment pour volaille, et jusqu'à 68% de déchets organiques (Diener et al 2011). Concrètement un kilogramme d'œufs de mouches soldat donnera en 10 jours 10 tonnes de larves vivantes consommant 40 à 50 tonnes de déchets alimentaires.

Un autre aspect intéressant de la mouche soldat noire est également présenté par Diener et al (2011) : un mélange un pour un de déchets organiques et de boues fécales (excrément déshydraté), a conduit les larves de BSF à une réduction de 66,6% de ces déchets. Ce genre de matière première, hostile pour la plupart des animaux ne l'est pas pour *Hermetia illucens*. C'est en effet grâce à sa grande résistance face aux pathogènes que cela est possible (Voir Partie 1 - *Hermetia illucens* - Cycle de vie).

Et c'est d'ailleurs grâce à cette grande capacité, que l'on a trouvé aux larves de BSF, de résister et d'assainir son environnement, des bactéries, que l'on a commencé à utiliser *Hermetia illucens* avec cet objectif. Que ce soit le fumier de volailles ou le lisier de porcs, l'action de la mouche soldat noire va alors assainir et de plus, les nutriments sont réduits de manière significative entre 37 et 93 % selon les éléments avec l'azote réduit lui à 71% (Newton et al 2005). L'ensemble des macro et microéléments prélevés du lisier de porcs sont accumulés dans les larves.

L'expérience en laboratoire a également montré que les mauvaises odeurs du lisier étaient capturées par les larves d'*Hermetia illucens* (Newton et al 2005).

*Hermetia illucens* a été utilisée avec succès sur des tourteaux de palmeraie, elle vient aussi s'inviter dans les plantations de café pour dégrader les coques de café, ces dernières qui étaient jusqu'à présent, laissées de côté (Noerhatini et al 2022).

Par ailleurs, nous l'avions déjà vu avant, mais *Hermetia illucens* permet aussi de faire de la lutte biologique contre les autres espèces de mouches (comme *Musca domestica* – figure 34 - et *Fannia canicularis*) et ce, sans pour autant envahir les milieux anthropisés (Bradley et Sheppard 1984).



Figure 34 : *Musca domestica* - Dessin observation (Beink)

*Hermetia illucens* renter en compétition avec *Musca domestica* permettant une lutte biologique. En effet contrairement aux mouches domestiques les mouches soldats noires ne sont pas gênante pour l'homme.

Cependant, des actions restent à mener. Les insectes, *Hermetia illucens* en tête, sont en effet l'une des solutions pour diminuer la quantité de déchets produits, mais ce n'est pas la solution miracle. Aussi bien au niveau habitude de vie, recherche fondamentale (sur les insectes et les potentiels risques) et législation, une marge de progrès est possible (Lalander et Vinneras 2022).



Rappelons une évidence : le meilleur déchet est celui qui n'est pas produit (figure 35) mais quand bien même, réduirions-nous drastiquement nos déchets, il y en aura toujours.

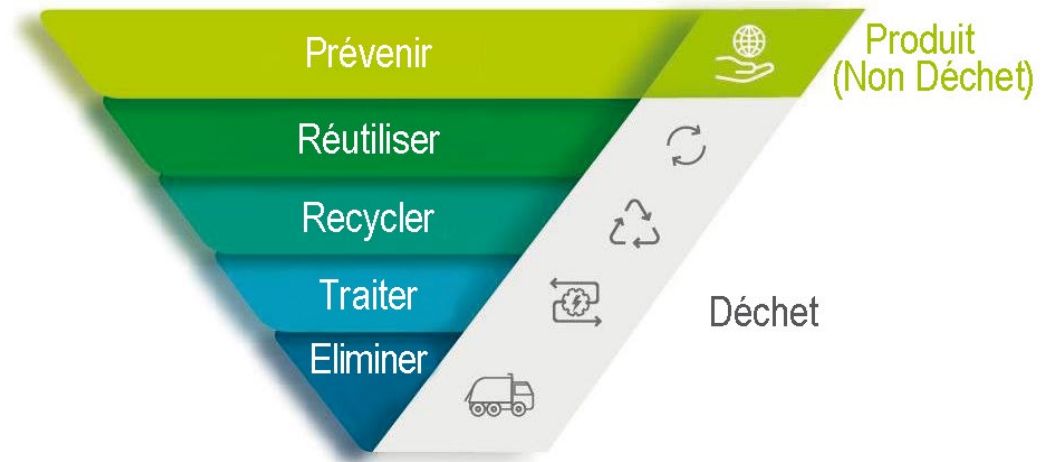


Figure 35 : Pyramide du traitement des déchets

Adaptée de Lalander et Vinneras 2022

Les insectes sont prometteurs et la mouche soldat noire en est un bon exemple (Amrul *et al* 2022). Mais si ce n'était que la diminution des déchets, cela rendrait déjà *Hermetia illucens* fort intéressante, mais rappelons-nous, que les pièces buccales de l'adulte étant atrophiées (mais pas non fonctionnelles - voir partie 1 *Hermetia illucens* - Cycle de vie), l'adulte va tirer avantage de réserves acquises au stade larvaire, et pourra effectuer sa mue imaginale. Elle pourra se reproduire et pondre même sans manger (Lupi *et al* 2019). Et nous allons, nous, *Homo sapiens* pouvoir tirer profit de ces réserves.

## Microbiologie et digestion

Évidemment, cette capacité à digérer tout ce qui l'entoure (ou presque) ne se fait pas sans un système digestif extrêmement efficace.

Le microbiote d'un animal est l'ensemble des micro-organismes qui le compose. De manière plus spécifique, nous allons nous intéresser ici au microbiote intestinal.

L'intestin des mouches soldat noires (comme celui de tous les insectes) peut être divisé en trois parties :

L'intestin antérieur, l'intestin moyen, région la plus longue où a lieu la quasi-totalité de la digestion et de l'absorption, et enfin l'intestin postérieur où l'eau, les sels et certaines molécules sont absorbés avant l'élimination des matières fécales.

Le pH varie tout le long de cet intestin, et plus particulièrement dans l'intestin moyen qui peut lui-même être divisé en trois parties (figure 36). Dans ce dernier le pH est autour de 7 dans la partie antérieure, entre 2 et 3 dans la partie moyenne et 8 dans la partie finale (Bruno *et al* 2019 - Gold *et al* 2018).

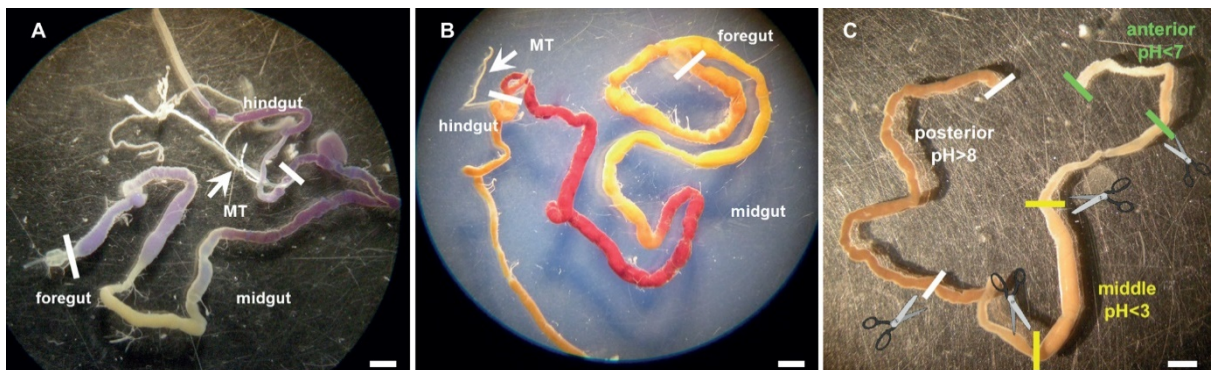


Figure 36 : Parties et pH d'un intestin d'*Hermetia illucens*.

A & B : anatomie de l'intestin d'*Hermetia illucens* coloré au bleu de bromophénol (A) et au rouge de phénol (B). Les barres représentent 2mm, et MT sont les tubes de Malpighi (organe excréteur des panarthropodes). On y voit les trois parties de l'intestin : l'intestin antérieur (foregut) qui commence à la bouche. L'intestin moyen (midgut) partie la plus longue et enfin l'intestin postérieur (hindgut) qui se termine par l'anus. C : on y voit les variations de pH le long de l'intestin moyen (midgut) avec une première section avec un pH neutre, suivi d'une partie très acide et enfin d'une troisième partie alcaline (adapté de Bruno *et al* 2019).

Le régime alimentaire a un impact sur la croissance des larves, il a également un impact sur l'endroit de l'intestin où l'on retrouve les bactéries. Il ne semble pas avoir d'influence sur le nombre total de bactéries dans l'intestin mais plutôt sur les communautés bactériennes qui peuplent l'intestin (Bruno *et al* 2019). De manière générale, la richesse bactérienne va en diminuant le long de l'intestin, à l'inverse, la densité en bactéries augmente lorsque l'on progresse le long de l'intestin, pour être maximale, quel que soit le régime alimentaire dans l'intestin postérieur.

Un microbiote cœur a été décrit chez les BSF (Klammsteiner *et al* 2020), celui-ci est composé de *Dysgonomonas*, d'*Enterococcus*, et d'*Actinomyces* (figure 37).

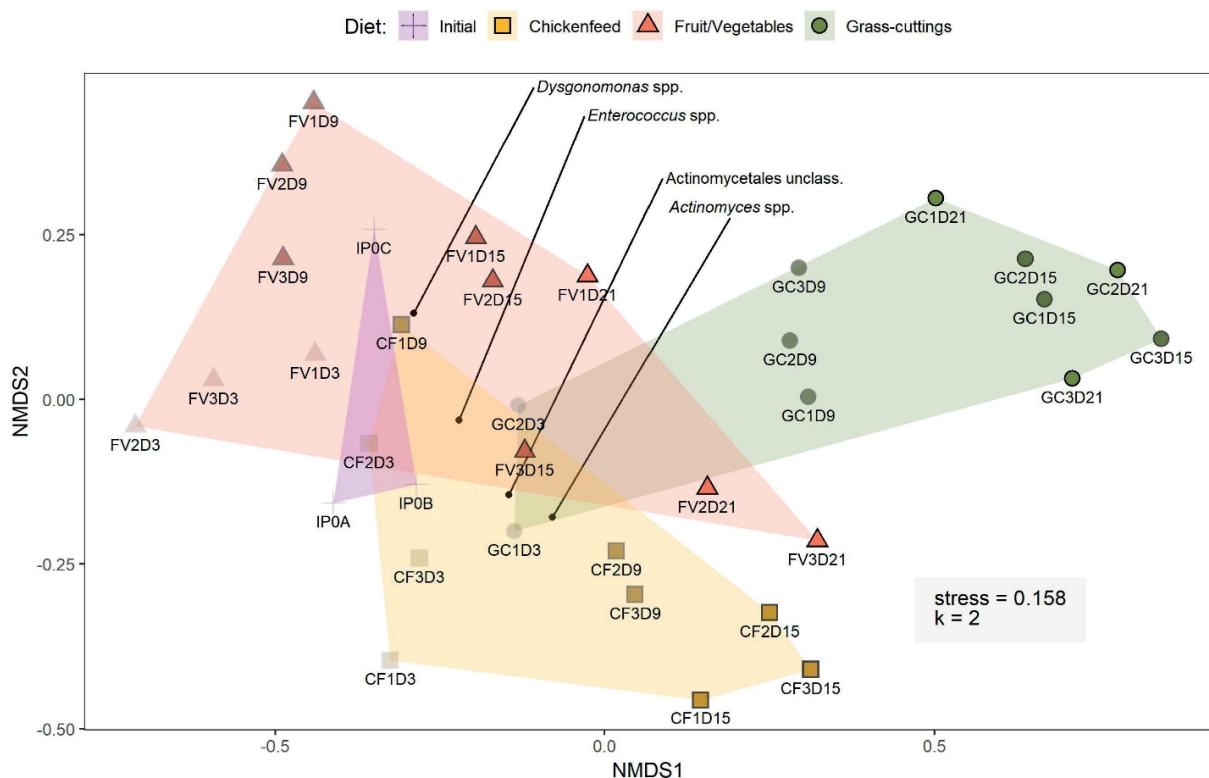


Figure 37 : Analyse Nonmetric Multidimensional Scaling (NMDS) du microbiote de larves d'*Hermetia illucens* nourries avec différents substrats

L'analyse initiale est celle des bactéries présentes avant l'ajout d'un quelconque substrat, puis mesurées avec différents substrats. Les points noirs indiquent les OTU (unité taxonomique opérationnel) présents dans au moins 80% des échantillons et qui constituent le microbiote cœur des BSF (Klammsteiner *et al* 2020).

Le microbiote des larves de mouches soldat noires va jouer un rôle important dans plusieurs domaines :

- En premier lieu, associés aux enzymes digestives (principalement amylase, lipase et protéase), il agit sur la digestion des aliments, donc sur les aspects nutritionnels (Kim *et al* 2011). Les larves de BSF vont être capables de se nourrir sur un grand nombre de substrats différents, la classant dans les insectes polyphages.

- Ensuite le microbiote va agir, de concert avec d'autres molécules, tel que des peptides antimicrobiens, pour la protection contre les pathogènes. Mais également sur la détoxification et la dégradation des antibiotiques (Zhang *et al* 2020). Ainsi que la réduction des odeurs (Beskin *et al* 2018) et le choix du lieu de ponte (Jordan *et Tomberlin* 2021).

Les bactéries intestinales sont également impliquées dans le métabolisme énergétique et moléculaire, la biosynthèse des polysaccharides, et également dans des aspects moins attendus tel que le transport d'ions inorganiques, le cytosquelette et le transport membranaire (Zhineng *et al* 2021). Par ailleurs, d'une génération à une autre, il existe une faible variabilité inter-cycle dans les communautés bactériennes, au cours de cycles d'élevages successifs lorsque le protocole d'élevage est standardisé (Gorrens *et al* 2022)

Pour résumer, nous pouvons considérer, comme Gold *et al.* (2018), les larves de mouches soldat noires comme des bioréacteurs pour les déchets organiques (figure 38). C'est la combinaison des microorganismes et des capacités intrinsèques des larves, qui vont faire de la BSF une championne de la dégradation.

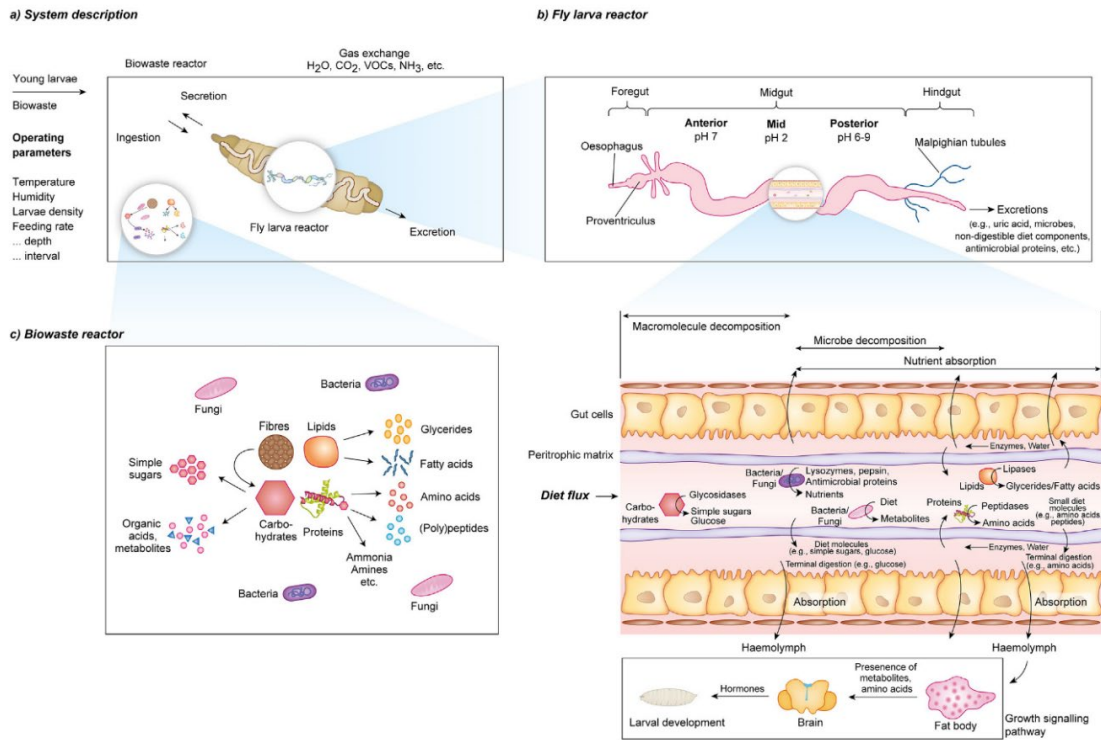


Figure 38 : Résumé schématique du système digestif des larves de BSF

A : Les aliments ingérés par les larves passent par l'intestin qui peut être considéré comme un réacteur de biodéchets. B : l'intestin est divisé en trois parties avec des variations de pH importantes dans la partie médiane. C : Dégradation à l'aide d'enzymes et de bactéries des nutriments présents dans les biodéchets. La majorité de la digestion se fait au niveau de l'intestin moyen, les nutriments sont dégradés dans la première partie de l'intestin qui sont absorbés par la suite permettant la croissance des larves (d'après Gold et al 2018).

De manière similaire, le microbiote du substrat, c'est-à-dire les communautés bactériennes présentes dans l'aliment pour BSF, va être impacté par la densité en larves et la nature de ce substrat (Schreven et al., 2022).

Nous voilà donc un peu plus avancés maintenant. Nous savons le comment, maintenant passons au pourquoi. Pourquoi digérer ? Pour se développer, faire ses mues, faire des réserves évidemment.

Alors ces réserves seront sous deux formes principales. Des protéines et des lipides. Nous, *Homo sapiens*, avons trouvé comment les utiliser, comment mettre à profit ces deux familles de molécules, et d'autres familles encore.

## Utilisation comme source de protéine

Les insectes sont depuis longtemps considérés comme une source alternative de protéines (Calvert *et al* 1969). *Hermetia illucens* n'a pas tardé à attirer l'attention. Dès 1973 en complément pour le poulet (Hale 1973) puis en 1977 pour le porc (Newton *et al* 1977) et en 1981 pour les poissons (Bondari et Sheppard 1981).

En effet, les larves de BSF vont pouvoir contenir, en fonction du substrat sur lequel elles ont été nourries, 32 à 58% de protéines, en masse. Une fois débarrassée de la partie lipidique, cela peut monter jusqu'à 55 à 65%. Ce genre de valeur peut également être atteinte avec des insectes comme le ver de farine. À titre de comparaison, le bœuf en contient 19 à 26% (Van Huis *et al* 2013).

L'espace nécessaire pour produire des protéines, ainsi que le potentiel de réchauffement climatique, sont bien moindre avec des insectes qu'avec des sources de protéines animales plus traditionnelles (lait ; porc ; poulet ; bœuf). Cependant, la quantité d'énergie nécessaire est semblable aux autres élevages (Oonincx et de Boer 2012). Ces aspects sont valables également chez *Hermetia illucens*. De même, là où les protéines végétales sont parmi les plus durables, les protéines d'insectes ne sont pas encore compétitives, mais pourraient le devenir à moyen ou long terme. Notamment sur le plan environnemental, cela est d'ailleurs déjà le cas en matière d'utilisation d'eau et de terre (Smetana *et al* 2019).

Ces protéines peuvent être utilisées pour nourrir différents animaux, notamment des poissons. Elles présentent une excellente digestibilité pour le poisson et les autres animaux. Les protéines d'insectes de manière générale sont qualifiées de protéines complètes car elles contiennent tous les acides aminés (Leni *et al* 2020) contrairement à la majorité des protéines végétales.

La composition du substrat, sur lequel les mouches auront été élevées, aura forcément un grand impact sur le taux et la qualité des protéines de la larve. En effet, un élevage de mouches sur des déchets fruitiers, apportera environ 300g de protéine par kilogramme de matière sèche, tandis que sur des coproduits brassicoles, ce sera

plus de l'ordre de 530g de protéine par kilogramme de matière sèche (Meneguz *et al* 2018).

Les essais pour faire de la farine d'insectes à partir de larves de BSF déshydratées ont été nombreux. Plusieurs études, plus ou moins concluantes, ont montré que cette protéine pouvait servir pour nourrir des différents animaux :

Animaux	Impact	Etude
Porcs	Aucun impact (ni positif ni négatif)	Biasiato <i>et al</i> 2019
Chèvres	Potentiel amélioration avant et après sevrage ainsi que sur la croissance. Source de probiotique via l'apport de bactérie lactique	Astuti <i>et Wiryawan</i> 2021
Volailles	Compensation d'un potentiel déficit en acides aminées des BSF par rapport au soja. Potentiel jaunissement de la peau et de la chair des poulets.	Heuel <i>et al</i> 2022
Chiens	Amélioration de la digestibilité, des capacités anti-inflammatoires et antioxydantes	Lei <i>et al</i> 2019
Tilapias	Effet positif sur les transporteurs de protéines et de graisses. Pas d'impact sur la croissance. Potentiel amélioration de l'immunité.	Kishawy <i>et al</i> 2022
Truites	Diminution du rapport hépatosomatique (indice sur les réserves énergétiques). Diminution de l'albumine et augmentation de la globuline. Aucun impact sur la croissance.	Mikołajczak <i>et al</i> 2022
Soles	Effet positif sur la croissance à faible dose de remplacement. Au-delà de 50% de remplacement, apparition de dommages intestinaux.	Li <i>et al</i> 2022
Saumons	Apport mineur, ou non significatif.	Belghit <i>et al</i> 2019

Par ailleurs, bien que certaines de ces études n'aient apporté aucune preuve d'une amélioration de la croissance des animaux, par le remplacement de l'aliment "habituel" (figure 39), par de la BSF, les impacts écologiques et économiques ne sont pas toujours pris en compte. Actuellement la plupart des poissons d'élevage sont nourris avec des aliments à base de farine de poissons (carcasses de poissons réduites en poudre) et huile de poisson. Le remplacement



de tout ou partie de cet aliment par de l'aliment à base d'insectes, peut s'inscrire dans une démarche écologique et potentiellement économique, selon les coûts des matières premières (Mikołajczak *et al* 2022).

Smetana *et al* (2016) ont comparé les impacts environnementaux de la production de ces protéines d'insectes avec des aliments plus classiques. Ils ont conclu que les BSF élevées sur des produits issus de l'industrie alimentaire (drêches de distillerie) permettaient d'obtenir un aliment ayant le même impact environnemental que les aliments classiques pour poulet. Cependant, si cet aliment était produit avec des déchets municipaux, l'impact environnemental était largement plus faible.

Parameters	Treatments <sup>1</sup>				SEM	P-value
	CON	BSFL5	BSFL10	BSFL20		
<b>Growth performance</b>						
Initial BW (g)	4.81	4.67	4.75	5.00	0.0693	0.399
Final BW (g)	10.69	10.31	10.76	10.65	0.1299	0.684
BW gain (g)	5.88	5.64	6.02	5.64	0.1134	0.645
Specific growth rate (%/day)	1.38	1.37	1.42	1.31	0.0245	0.453
Survival rate (%)	100	100	100	100	0.0000	1.000
<b>Feed utilization</b>						
Daily intake rate (%/day)	1.29	1.30	1.32	1.30	0.0067	0.606
Feed intake (g, as fed)	5.79	5.64	5.90	5.87	0.0669	0.568
Feed conversion ratio	0.99	1.00	0.98	1.04	0.0140	0.456
Protein efficiency ratio	2.15	2.14	2.14	2.01	0.0301	0.343
<b>Environmental sustainability</b>						
Fish meal use (g/kg of fish gain)	345.33 <sup>a</sup>	325.00 <sup>a</sup>	295.00 <sup>b</sup>	260.83 <sup>c</sup>	10.2714	< 0.001
Fish oil use (g/kg of fish gain)	138.13 <sup>a</sup>	123.00 <sup>b</sup>	103.25 <sup>c</sup>	73.03 <sup>d</sup>	7.4534	< 0.001
FIFO	1.76 <sup>a</sup>	1.63 <sup>a</sup>	1.45 <sup>b</sup>	1.21 <sup>c</sup>	0.0641	< 0.001

Figure 39 : Analyse de l'inclusion de protéines issu de mouche soldat noires dans la croissance des truites européennes – *Salmo trutta*.

*L'inclusion de 5 10 ou 20% de protéines issus d'Hermetia illucens n'a aucun impact significatif sur les performances de croissance ou les quantités de nourriture utilisées pour la croissance des truites. Cependant l'inclusion de ces protéines a un impact environnemental positif au regard de la quantité de farine de poisson utilisée pour nourrir les truites (d'après Mikołajczak et al 2022).*

Pour ce qui est des qualités organoleptiques, une étude menée en nourrissant des canards avec des larves vivantes, a conclu que les œufs avaient une bonne saveur et une belle couleur. (Gunawan 2022)



Enfin des études ont montré que la BSF possédait une cinquantaine de peptides antimicrobiens (AMP), soit la plus grande famille d'AMP jamais identifiée chez un insecte (Zhan *et al* 2020).

Un peptide antimicrobien est un enchaînement d'acides aminés (au même titre qu'une protéine - mais de plus petite taille) qui possède une action sur les pathogènes de l'hôte. Ils font partie de la défense immunitaire innée.

Chez la BSF, un extrait lyophilisé de larves, avait des activités antimicrobiennes, notamment contre les staphylocoques dorés, résistant à la méticilline (Park *et al* 2014). Plus récemment, un peptide antimicrobien a été caractérisé, il aurait des activités perturbatrices de l'architecture de la membrane cellulaire des bactéries (Di Somma *et al.*, 2021).

Un hydrolysat de protéines s'est avéré avoir des propriétés antioxydantes, qui ont pu être caractérisées via des techniques de chromatographie liquide haute performance. Certains se sont avérés présenter des risques d'allergie, cependant, la plupart sont non allergènes et non toxiques et peuvent donc être utilisés dans la transformation d'aliments, de cosmétique et en médecine (Lu *et al* 2022).

## Utilisation comme source de lipides

*Le gras c'est la vie  
Karadoc - Kaamelott - Livre 2 épisode 76*

Si, comme on vient de le voir, les protéines constituent une part importante des réserves des larves de mouches soldat noires, l'autre stockage est constitué de lipides. Et ces lipides ne sont pas en reste question apport.

En effet, ils peuvent être d'une part, utilisée comme aliment, rentrant dans la composition de certains produits à destination des animaux comme les poissons (Li *et al* 2022). Ils peuvent être également utilisés pour produire des biocarburants. 1000 larves croissant sur un kilogramme de fumier de bovins / porcs / poulets donneront 36 / 58 / 91 g de biocarburant (Li *et al* 2011). Des essais plus poussés ont également été fait en Corée (figure 40), où un rendement de production supérieur à 85% a été constaté. (Park *et al* 2020).

Feedstocks	Oil (lipid) contents (% oil by wt. in dry biomass)	Oil yield (L oil ha <sup>-1</sup> year <sup>-1</sup> )	Biodiesel productivity (kg biodiesel ha <sup>-1</sup> year <sup>-1</sup> )
Soybean	18	636	562
Rapeseed	41	974	862
Palm	36	5,366	4,747
Jatropha	28	741	656
Microalgae	50	97,800	86,515
BSFL	30 – 40	162,000	145,800a

Figure 40 : Comparaison des quantités de biocarburant produites en fonction des matières premières utilisées

*La quantité de biocarburant en utilisant des larves de BSF (BSFL) est plus importante qu'avec l'utilisation de soja de colza de palmier de jatropha ou de microalgue. L'étude a considéré 25000 larves par m<sup>2</sup> (Park et al 2020).*

Il est également possible de faire rentrer *Hermetia illucens* en cours de procédé, dans un objectif de production de biocarburant. En partant de 200 grammes de tiges de maïs, majoritairement composées de résidus lignocellulosiques (non utilisable par la mouche soldat noire en l'état), on peut effectuer une première dégradation avec des vers de farine (*Tenebrio molitor*). Puis, par la suite, utiliser des larves de mouches soldat noires (figure 41). Ce genre de procédé a conduit à une réduction des déchets de 51,8% (en matière sèche), la production de 8,5g d'insectes, 1,95 g d'acide gras lui-même converti en 1,76g de biodiesel. Soit un taux de conversion du gras d'insecte brut en biocarburant de 90% (Wang et al 2017).

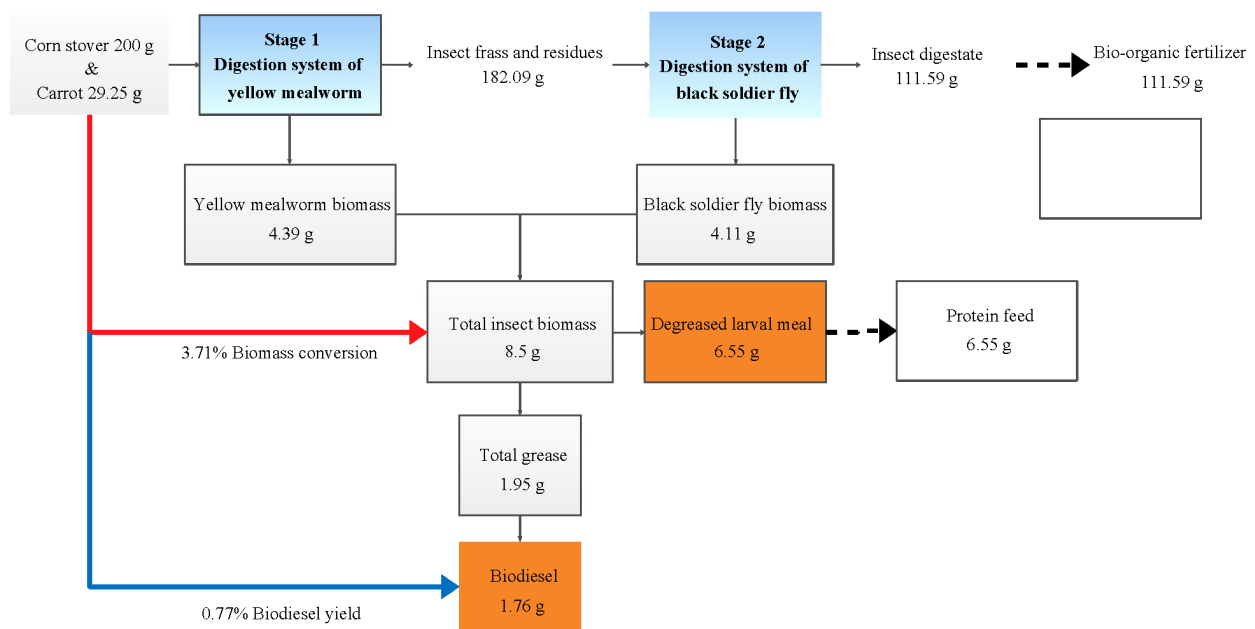


Figure 41 : Résumé schématique de la production de biocarburant par l'action combiné de *Tenebrio molitor* et *Hermetia illucens* à partir de fanes de maïs.

Adapté de Wang et al 2017.

Si l'on regarde d'un peu plus près, la composition en acide gras des larves de BSF, on se rend compte qu'elle va dépendre du substrat utilisé. Meneguz et al (2018) ont élevé des BSF sur des produits aussi variés que des fruits, légumes, coproduits de brasserie et viticoles. Ils ont obtenu majoritairement des acides gras polyinsaturés, tandis qu'Ewald et al (2019) ont montré qu'élever des BSF, sur des substrats tel que, du pain, des déchets alimentaires, ou encore du poisson,

donnaient majoritairement des acides gras saturés et des acides gras mono insaturés.

Enfin ces acides gras peuvent être utilisés dans les cosmétiques, des essais de crème pour les mains ont notamment été réalisés et se sont avérés concluants (Verheyen *et al* 2018). De même, des applications médicales autour des lipides des BSF verront très probablement le jour dans les prochaines années. L'utilisation d'acide laurique dans le traitement de l'acné (Nakatsuji *et al* 2009 ; Almeida *et al.*, 2022) en est un exemple. Des recherches ont récemment porté leurs fruits en utilisant les acides gras produits par la BSF dans la lutte contre *Klebsiella pneumoniae* multirésistant aux antibiotiques (Mohamed *et al.*, 2022).

## Autres utilisations et aspects réglementaires

*Dura lex, sed lex*

Jusqu'à présent, nous pouvons résumer l'utilisation de la mouche soldat noire à trois grands domaines :

- La production d'aliments, via les protéines et les lipides.
- La réduction de biodéchets (figure 42).
- La production de biodiesel via les lipides.

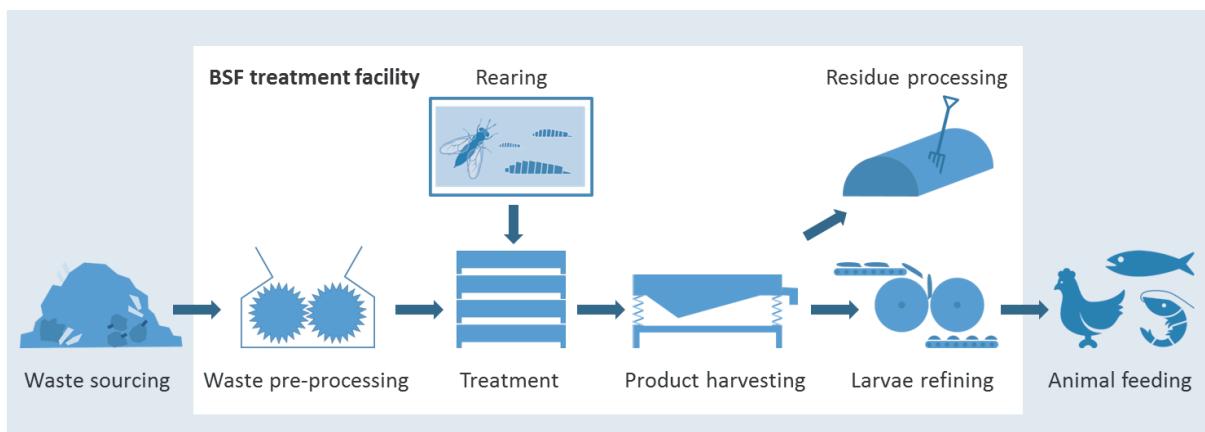


Figure 42 : Résumé schématique du processus de transformation de biodéchets par l'utilisation de mouches soldat noires.

Ce schéma présente l'utilisation d'*Hermetia illucens* à visée de production d'aliment pour des élevages (Zurbrügg et al 2018)

Mais d'autres aspects de la mouche soldat peuvent encore être exploités. Il y a tout d'abord la chitine, ce polymère osidique, qui constitue l'exosquelette des insectes. Ce dernier a des utilités diverses : il possède des propriétés antimicrobiennes, peut être utilisé dans les cosmétiques ou en tant que cicatrisant sur les brûlures. C'est également un agent dépolluant des eaux.

L'avantage est, que l'extraction de la chitine des BSF ne rentre pas en compétition avec les autres extractions (protéines et lipides) mais peut se faire en complément (Guarnieri et al., 2022).

Le rendement en chitine est de 5% (Lee et al 2022) ce qui est bien plus faible que chez les crustacées, mais les quantités disponibles pourront pallier ce rendement (Guarnieri et al., 2022). La quantité de chitine est également impactée par le régime alimentaire des larves (Meneguz et al 2018).

Quant aux aspects de nutrition humaine directe, le criquet migrateur (*Locusta migratoria*) et le ver de farine (*Tenebrio molitor*) sont déjà autorisés sur le marché européen pour la consommation humaine. Ce n'est pas encore le cas pour *Hermetia illucens* qui reste donc en attente de validation. En effet, pour qu'une espèce soit approuvée par les autorités, il faut que sa composition ne mette pas en danger les consommateurs. De nombreuses recherches ont donc été faites dans ce domaine. Elles tendent à montrer que les larves de mouches soldats noires n'accumulent pas de produits toxiques (organiques ou inorganiques) au-delà de la législation, exception faite de certains métaux (comme le plomb et le cadmium notamment). D'autres études continuent de paraître régulièrement, ce qui fera peut-être un jour pencher la balance définitivement pour obtenir l'autorisation ou non (Lievens et al 2021 - Shah and Çetingül, 2022).

Par ailleurs, les résidus, laissés par la dégradation de la matière organique, appelés « frass », peuvent également être utilisés comme amendement au même titre qu'un engrais (Schmitt et de Vries 2020). Ces résidus ont un impact positif sur le taux de chlorophylle dans les plants de riz, et permet ainsi de réduire la quantité de fertilisant chimique (Reswita et al., 2022).

Nous sommes ici dans une économie circulaire complète. Néanmoins, certaines études rappellent que, pour être réellement considérées comme un gain environnemental, les larves de BSF devraient être élevées directement sur des déchets organiques à faible teneur en protéine. Ces protéines produites par les larves devraient être utilisées directement pour l'alimentation humaine (et non des animaux de rente). Il ne faut donc pas croire que tout serait résolu avec la mouche soldat en 2022 (Seyedalmoosavi et al 2022). Cette étude conclut volontiers sur le fait

que la BSF serait davantage une espèce complémentaire, qu'une réelle prouesse pour l'économie circulaire.

En Europe, la BSF est autorisée pour l'alimentation animale mais sous certaines conditions (Alagappan *et al* 2021). Il faut rappeler en effet, que même si les insectes ne sont pas destinés à l'alimentation humaine directement, ils ne peuvent pas être élevés dans des substrats contenant des sous-produits animaux, eux-mêmes étant considérés comme des animaux d'élevages.

Maintenant que nous avons vu tous ces aspects, peut être commencez-vous à comprendre l'intérêt de cette mouche ! En tous cas, il n'y a pas que les industriels qui se soient saisis du sujet, les scientifiques l'ont également pris à bras le corps. D'où l'augmentation significative du nombre de publications autour de la mouche soldat noire. Même si ces chiffres doivent être modulés dans un contexte d'augmentation générale du nombre de publications scientifiques, tous sujets confondus (figure 43).

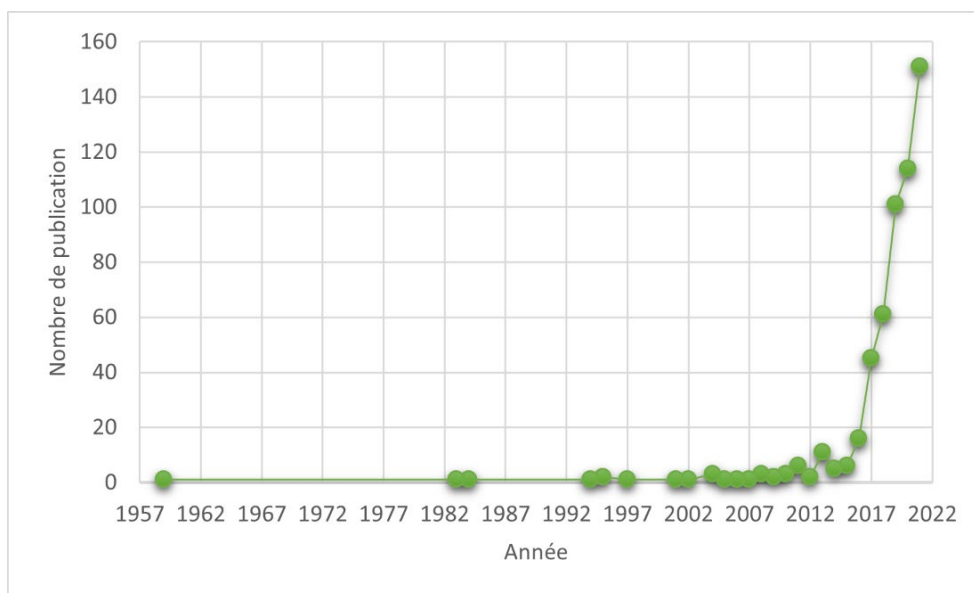


Figure 43 : Augmentation du nombre de publications traitant d'*Hermetia illucens* entre 1955 et 2021

Données issues de PubMed NCBI.

Dans des sujets aussi variés qu'il est possible de voir (figure 44).





## Aspect financier

---

*Il n'y a pas de déchets,  
que des choses que l'on ne sait pas encore vendre.*  
JG

---

S'il y a bien un aspect qu'il faut désormais prendre en compte lorsque l'on parle de l'élevage de la mouche soldat noire, c'est celui du modèle économique qui sous-tend le monde de l'insecte. En effet, bien que la BSF ne soit pas le seul insecte utilisé par des entreprises pour faire de la production de protéines et de lipides sur des co-produits, elle reste cependant en tête ([All about feed, 2017](#)).

Parlons business donc, peu de données sont encore présentes sur les différents insectes utilisés par les industries mais selon « meticulous research », pour ce qui est de tous les insectes : les marchés de l'insecte pourraient dépasser le milliard de dollars en 2023 ([Guiné et al 2020](#) – figure 45) et atteindre plus de 9 milliards d'euros d'ici 2030. Avec un CAGR (taux de croissance annuel) autour de 28,3% entre les années 2022 et 2030 (à titre de comparaison, le marché du pet food est autour de 5 milliards d'euros) et dépasser les 3 millions de tonnes d'ici 2030 ([meticulous research, 2022](#)).

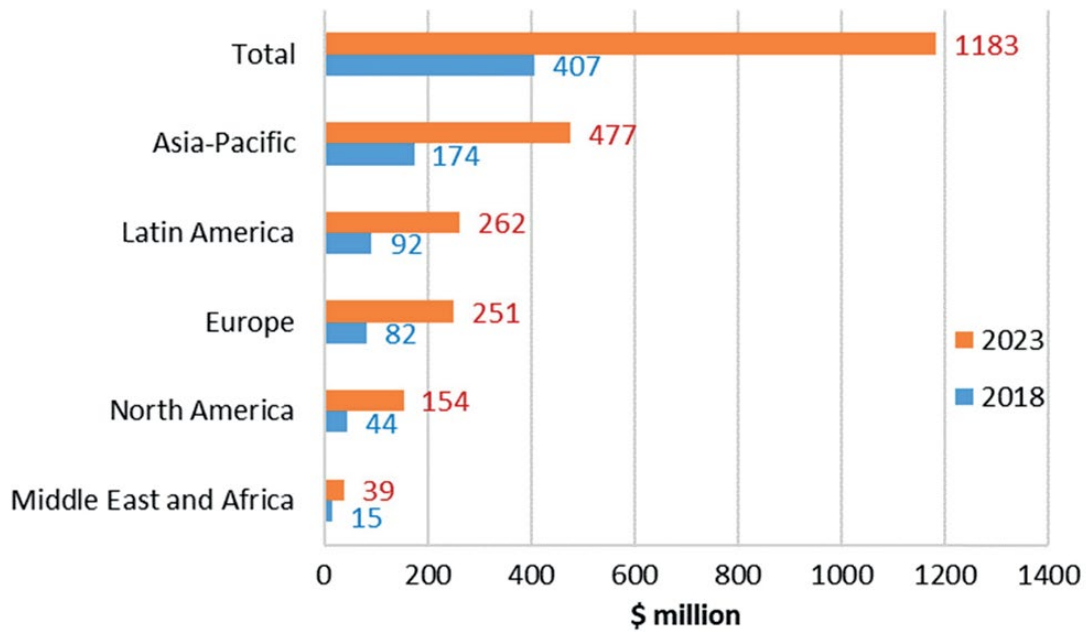


Figure 45 : Croissance attendu du marché des insectes comestibles

Evolution entre 2018 et 2023 dans les différentes régions du monde (Guiné et al 2020).

Par ailleurs, l'élevage d'insectes va permettre la création d'emplois, et bien qu'il soit difficile d'obtenir des chiffres fiables, nous pouvons regarder l'évolution du nombre d'entreprise entre 2005 et 2020 (Cohen et Duchemin 2020 – figure 46).

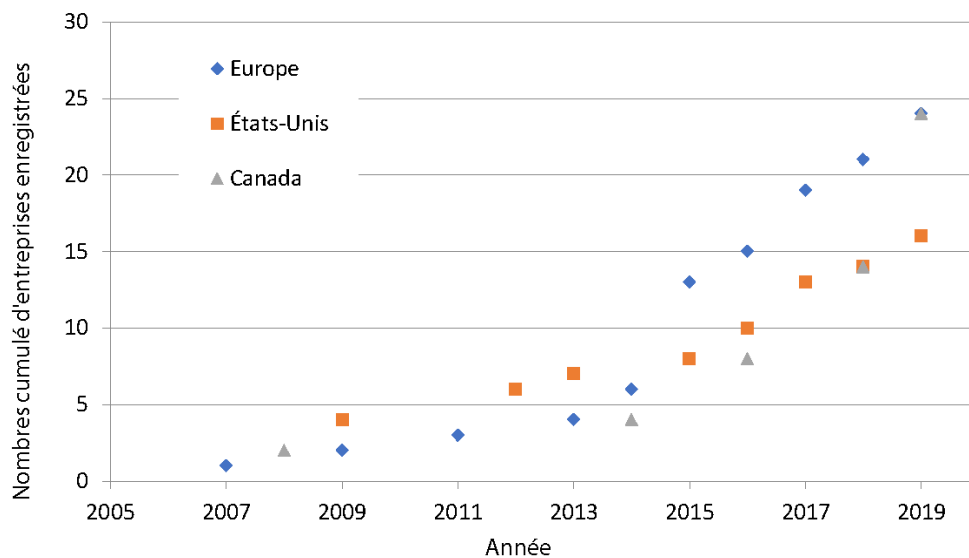


Figure 46 : Evolution du nombre d'entreprises d'élevages d'insectes

Evolution sur trois régions du monde entre 2005 et 2019 (Cohen et Duchemin 2020)

Et ces élevages, en comparaisons avec des élevages classiques utiliseront moins de nourriture, d'espace et d'eau (figure 47).

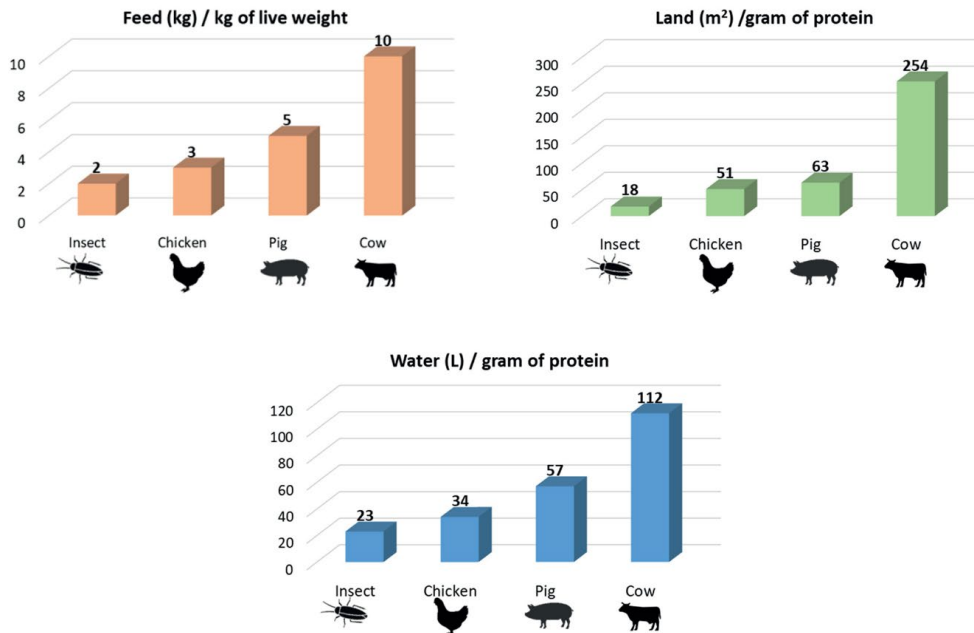


Figure 47 : Quantités comparées de nourriture d'espace et d'eau

Comparaison de différents élevages, les insectes nécessitent systématiquement moins d'intrants que les élevages comparés (Guiné et al 2020).

Si l'on se base sur les adhérents à l'IPIFF (International Platform of Insects for Food and Feed) de 2019 qui comptaient une soixantaine de membres parmi les principales entreprises de l'insecte, près des deux tiers travaillent avec la mouche soldat noire (IPIFF 2019).

En 2017, un article publié dans les échos, faisait état pour l'ensemble de la filière insecte, de levées de fonds cumulées autour de 100 millions d'euros (Les Echos, 2017).

Et deux ans après, en 2019, selon Crunchbase voilà les levées de fonds en dollars des 5 plus grosses entreprises de l'insecte :

175 millions pour Ynsect (*Tenebrio molitor* - France - créée en 2011)

122 millions pour AgriProtein (BSF - Afrique du Sud - créée en 2008)

120 millions pour EnviroFlight (BSF - Etats-Unis, créée en 2009)

55 millions InnovaFeed (BSF - France, créée en 2016)

45 millions Protix (BSF - Pays-Bas, créée en 2009)

Donc 4 des 5 plus grosses entreprises de l'insecte font de l'élevage de BSF et 2 d'entre elles sont françaises. En novembre 2020, Innovafeed a refait une levée de fonds de 140 millions de dollars.

En 2021, le total des levées de fonds était à plus de 400 millions de dollars.

Cependant, tout n'est pas encore gagné. En effet en terme économique, l'élevage d'insecte peine encore à convaincre. Une étude menée avec des farines d'insectes composées de *Tenebrio molitor*, montre une augmentation des coûts de l'alimentation en raison des prix de la farine sur le marché (ces coûts étant directement corrélés aux coûts de production). Par conséquent, bien que les avantages environnementaux soient évidents, l'industrie aquacole et la production d'insectes ne se rejoignent pas encore sur ces aspects purement économiques (Arru et al., 2019).

A l'heure actuelle, la production de farine d'insectes élevés sur des biodéchets n'est donc pas encore rentable et ce pour au moins deux raisons. La première est que le coût des autres sources protéiques (Soja notamment) reste en deçà de celui de la production à partir de BSF. Cela est principalement dû aux coûts de transport qui sont ridiculement faibles par rapport à l'impact écologique qu'ils ont. La deuxième raison est, qu'il n'y a pour l'instant aucun avantage

financier à favoriser le recyclage des biodéchets par des larves de BSF par rapport à une destruction pure et simple.

Nous voici donc avec un insecte qui peut être élevé facilement. En effet, en termes d'espace, les élevages de BSF peuvent être empilés et il n'y a pas besoin de supplémenter en eau. Mais tous les éleveurs, d'insectes ou non, sont en recherche permanente d'amélioration de leurs performances. La BSF ne fait pas exception. Une possibilité serait d'obtenir des mouches qui pondent beaucoup d'œufs, qui stockent davantage sous forme de protéines ou de lipides selon ce que l'on veut exploiter par la suite. Peut-on gagner en masse et en temps de développement autrement qu'en changeant la composition du substrat ? Peut-on trouver des *Hermetia illucens* plus résistantes à des températures plus basses ?

Et naviguant autour de toutes ces questions, quelle est la part de la génétique dans toutes ces améliorations que les industriels peuvent vouloir ?

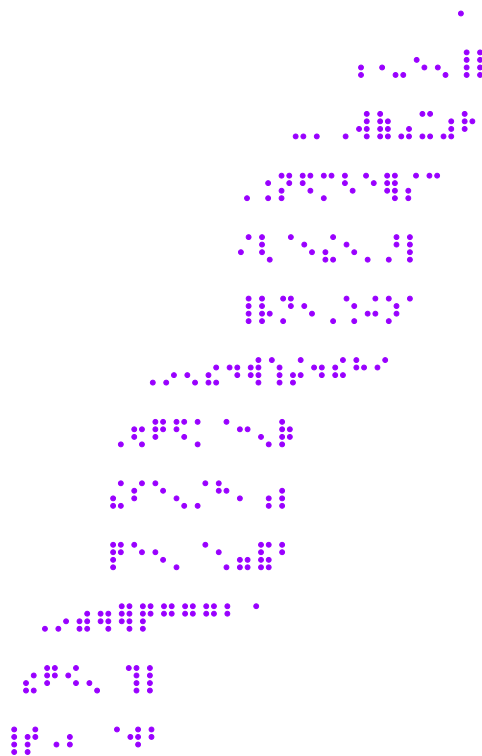
Alors sans plus tarder plongeons dans les aspects génétiques, la diversité et les potentiels d'amélioration qui s'offrent à nous.

## Partie 3 : Génétique

---

*Nous y voici finalement, troisième pilier de cette introduction de thèse, bienvenue dans ma cuisine ... moléculaire. Nous avons bien parlé d'*Hermetia illucens* jusqu'à présent et je vous rassure, cela va continuer. Mais nous allons continuer notre voyage au cœur des tissus, des cellules, des noyaux et des organites de notre petit diptère.*

---



Double hélice d'ADN – ASCII art

## Structure des génomes et séquençage.

---

ATGGCCAGATGCATTGCTTGA

---

### Structure des génomes

Un génome est l'ensemble des informations génétiques contenues dans chacune des cellules d'un organisme. Chez les eucaryotes, tel qu'*Hermetia illucens*, cela comprend d'une part, les chromosomes (autosomes et gonosomes) situés dans le noyau des cellules mais également l'ADN mitochondrial contenu lui dans les mitochondries. Dans ce mot : « génome » on retrouve la racine « gène » aussi vais-je le définir. En termes moléculaire, un gène se définit comme la séquence complète d'acide nucléique nécessaire à la synthèse d'un polypeptide ou d'un ARN fonctionnel.

#### Génome mitochondrial

Les mitochondries sont des organites, dérivant d'une bactérie symbiotique ancestrale, présents chez la plupart des eucaryotes (Karnkowska *et al.*, 2016), elles ont leur propre génome. Le génome mitochondrial d'*Hermetia illucens* (figure 47) est circulaire, comme tous les génomes mitochondriaux. Il fait un peu plus de 15 kb (kilobases) il contient 13 gènes codant pour des protéines, 22 gènes codant pour des ARN de transfert (tRNA) 2 codant pour des ARN ribosomiaux (rRNA) et une région non codante de contrôle (Qi *et al* 2017). Les mitochondries ont une transmission maternelle et sont présente en nombreuses copies dans les cellules des organismes.

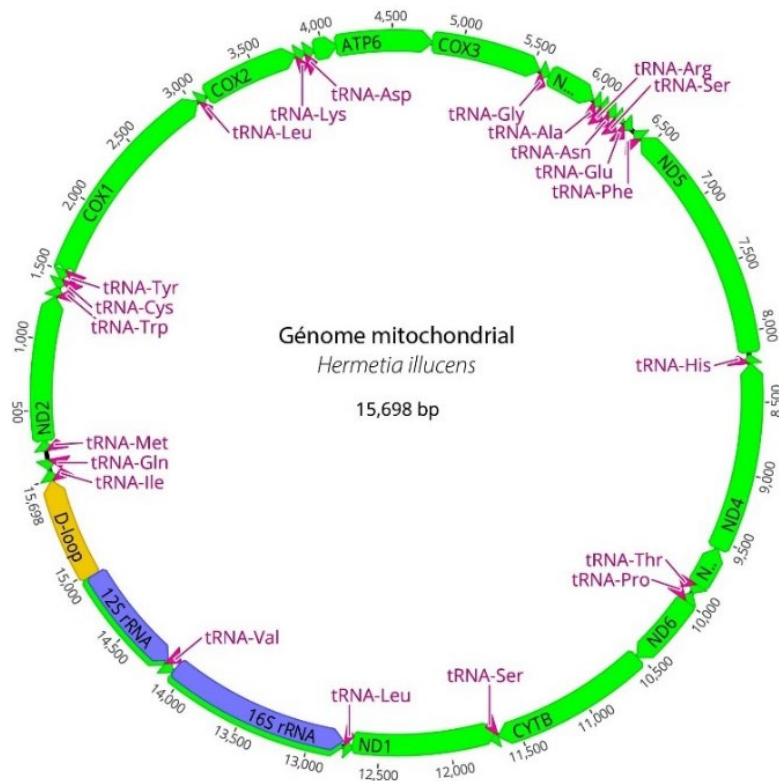


Figure 48 : Génome mitochondrial d'*Hermetia illucens*

Les gènes codant pour des protéines sont en vert, les tRNA en violet, les rRNA en bleu et les régions non-codantes sont en jaune (d'après Qi et al 2017, visualisation Geneious)

### Génome nucléaire

Le génome nucléaire, situé dans le noyau chez les eucaryotes, contient l'information génétique encodée sous la forme d'une double hélice d'ADN. Cette même hélice, par l'intermédiaire de protéines (histones et non histones), est compactée à différents niveaux de densité jusqu'au chromosome (Lodish et al., 2005).

Dans une cellule somatique, la ploïdie est le nombre de copies de chaque autosome. Le génome d'*Hermetia illucens*, est diploïde (Boyes 1973). Au sein d'un génome, nous allons retrouver les gènes codant pour des protéines (gènes uniques ou gènes dupliqués). Chez les eucaryotes, ces gènes sont composés d'exons et d'introns. Nous allons également retrouver les gènes codant pour des ARN fonctionnels. Le reste des génomes est composé d'ADN répété et d'ADN intercalaire (Lodish et al., 2005).



Les génomes sont de tailles très variées (elles sont données en mégapaire de bases Mpb). Le nombre de gènes contenu dans un génome n'est pas corrélé à sa taille. En effet, à titre d'exemple, chez les eucaryotes, *Drosophila melanogaster* à un génome de 118 Mpb avec environ 16500 gènes (Adams *et al.*, 2000) et le nématode *Caenorhabditis elegans* à un génome de 100 Mpb avec plus de 22500 gènes (Waterston *et al.*, 1993). Au départ les premières études révélaient que le génome d'*Hermetia illucens* était composé de 7 paires d'autosomes et d'une paire de gonosomes de type XY (Boyes 1973 – figure 49). Cependant des études plus récentes ont fait apparaître qu'il ne comptait que 6 autosomes (figure 50) et une paire de chromosomes sexuels. Le génome de la BSF fait 1010 Mpb, et code pour 16478 gènes (Generalovic *et al.*, 2021), ce qui le place parmi les grands génomes pour les diptères (figure 51).

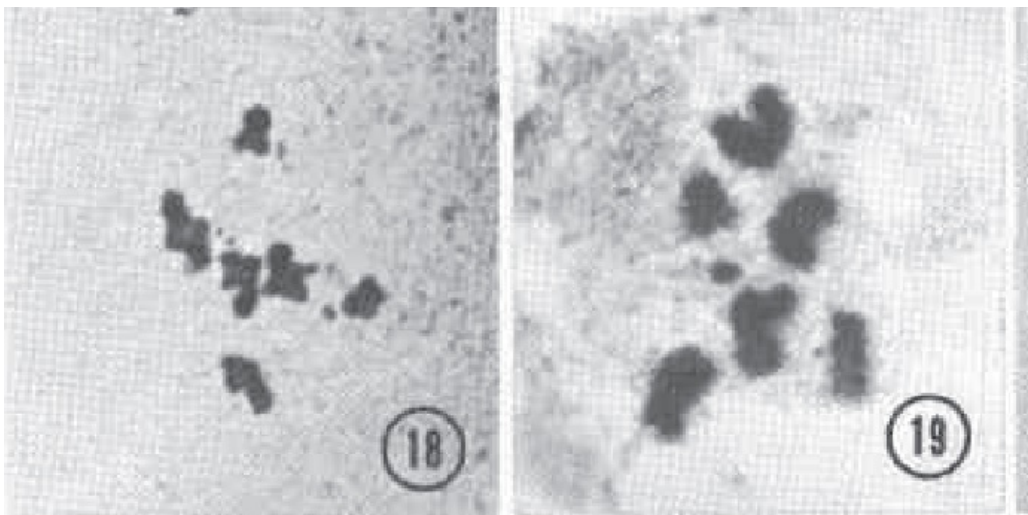


Figure 49 : Chromosomes d'*Hermetia illucens*

Caryotype (x1740) des chromosomes d'*Hermetia illucens* en métaphase I. Les chromosomes ont été extraits de cellules spermatiques (Boyes 1973)

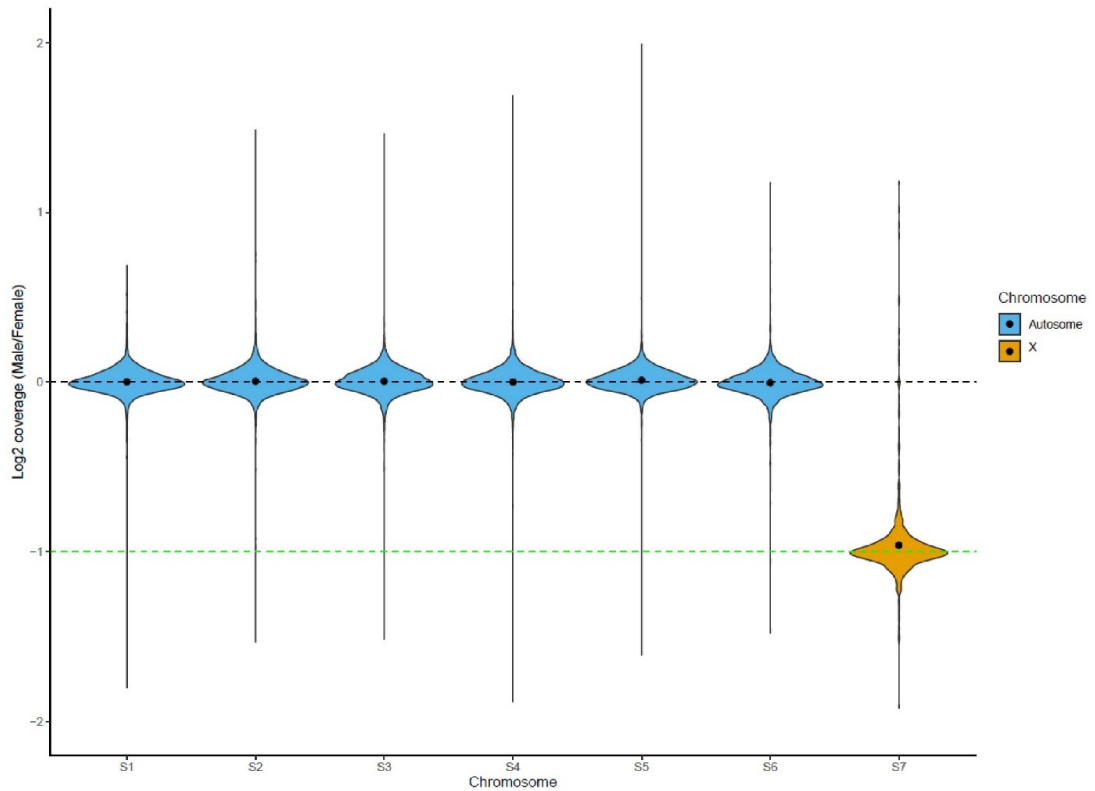


Figure 50 : Couverture Log2 (mâle/femelle) des données de reséquençage du génome entier

Les six autosomes et le gonosome X d'*Hermetia illucens*. Le chromosome sept révèle la moitié de la couverture attendue d'un autosome ressemblant à celui d'un chromosome sexuel X dans l'assemblage *H. illucens* (Generalovic et al 2021).

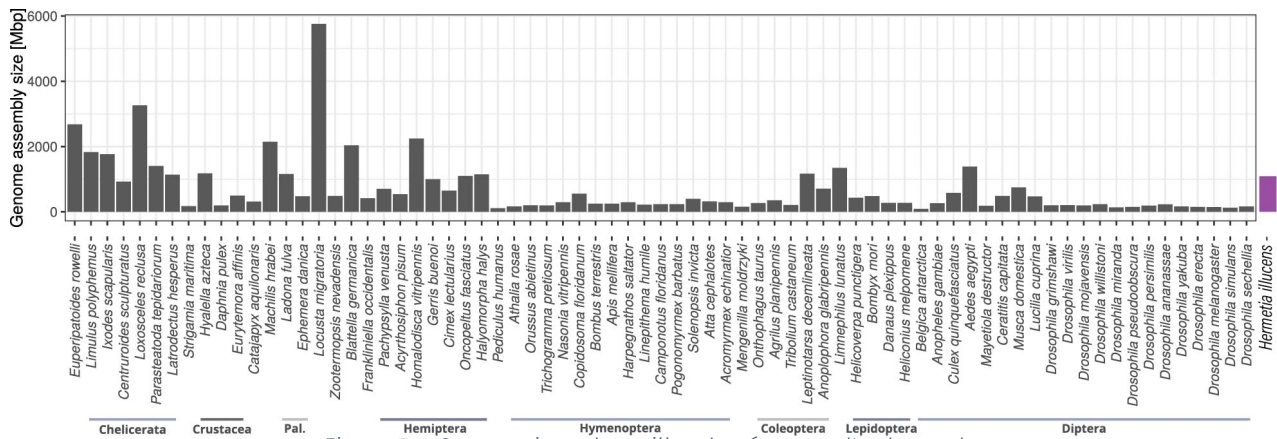


Figure 51 : Comparaison des tailles de génomes d'arthropodes

Les tailles sont données en Mpb, *Hermetia illucens* a été rajoutée en violet. La mouche soldat noire a un génome de taille importante par rapport aux autres diptères brachycères. Adapté de Petersen et al 2019.

## ADN répété

Nous l'avons dit un peu plus tôt, les éléments répétés sont une partie importante des génomes eucaryotes et ils jouent un rôle clé dans la génération de variations structurales et plus généralement dans l'évolution des génomes.

Il y a tout d'abord l'ADN répété de séquence simple, il s'agit de courtes séquences (de 1 à 500 nucléotides) répétées en tandem (de mille à un million de fois). On les rencontre en particulier près des centromères et des télomères (Lodish *et al.*, 2005).

En fonction de la taille de la séquence répétée (1 à 500 pb) et du nombre de répétition on distingue les satellites (longs et très répétés), les minisatellites (plus courts car moins répétés, tel que les télomères par exemple) et les microsatellites qui sont de court motif. La longueur d'une région contenant ces courtes séquences répétées en tandem est très variable d'un individu à l'autre, au sein d'une même espèce. Les différences de longueur de certaines de ces régions sont à la base de techniques de génétique des populations (Lodish *et al.*, 2005).

Par ailleurs il y a l'ADN répété dispersé, il s'agit de séquences répétées, en général plus longues que les précédentes. Elles ne sont pas organisées en tandem mais sont dispersées dans le génome. Ces séquences portent également le nom d'élément transposable (TE). Elles se sont en effet insérées dans les génomes au cours de leur évolution. Certaines ont encore la capacité de pouvoir s'y déplacer. Ces éléments ont été découverts à la fin des années 40 par la biologiste américaine Barbara McClintock (elle obtiendra pour cela le prix Nobel ... en 1983). Par ailleurs, le nombre d'éléments transposables peut être très différent même entre 2 espèces proches. En effet entre *Drosophila melanogaster* et *Drosophila simulans* il y a 3 fois plus de TE chez la première que chez la seconde (Vieira and Biémont, 2004).

Il existe deux classes d'éléments transposables. Ceux de classe 1 qui passe par un intermédiaire ARN pour leurs répliquations et qui se retrouvent de manière générale multipliés dans le génome (les rétrotransposons avec et sans LTR). Ceux de classe 2 qui ne passent

pas par un intermédiaire ARN et dont le mécanisme général est plutôt de type « couper-coller » (Wicker *et al.*, 2007 -figure 52).

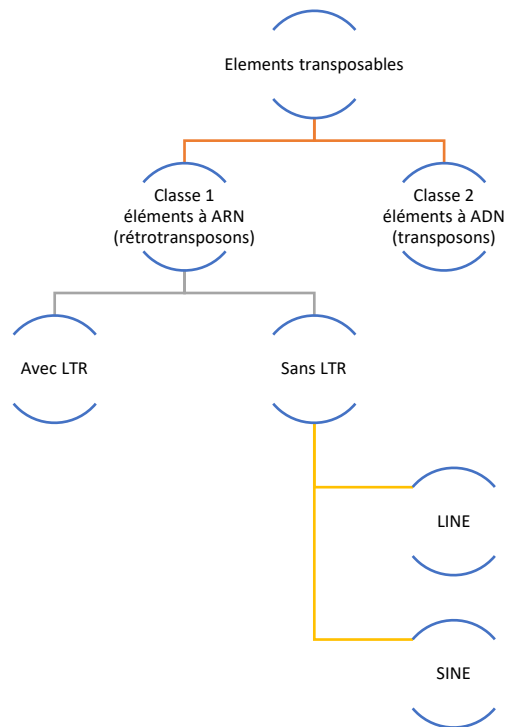


Figure 52 : Classification simplifiée des éléments transposables

### Les transposons d'ADN (classe 2)

Aussi appelé éléments à ADN, les transposons d'ADN sont des séquences d'ADN linéaires qui ne peuvent être isolés à l'état libre. Ils sont capables de promouvoir leur translocation d'une molécule d'ADN à une autre ou d'un endroit d'une molécule d'ADN à un autre (Lodish *et al.*, 2005).

### Les rétrotransposons à LTR (classe 1)

Ces éléments sont caractérisés par la présence de longues répétitions terminales (LTR) qui encadrent la région centrale codante. Ils se déplacent en faisant intervenir une transcriptase inverse qui

est codé directement par la partie centrale (Lodish *et al.*, 2005). Ils sont abondants dans les génomes des drosophiles (Peterson *et al.*, 2019 – figure 53).

### Les rétrotransposons sans LTR (classe 1)

Ils constituent une grande partie des éléments répétés des génomes, ils sont composés de LINE (Long Interspersed Nuclear Element) qui eux ne codent que pour une protéine qui fait tout le travail (séparation répliation intégration) de l'élément. Et enfin des SINE (Short Interspersed Nuclear Element) qui eux vont utiliser les protéines des autres éléments de classe 1 (LINE par exemple) pour réaliser leur transposition.

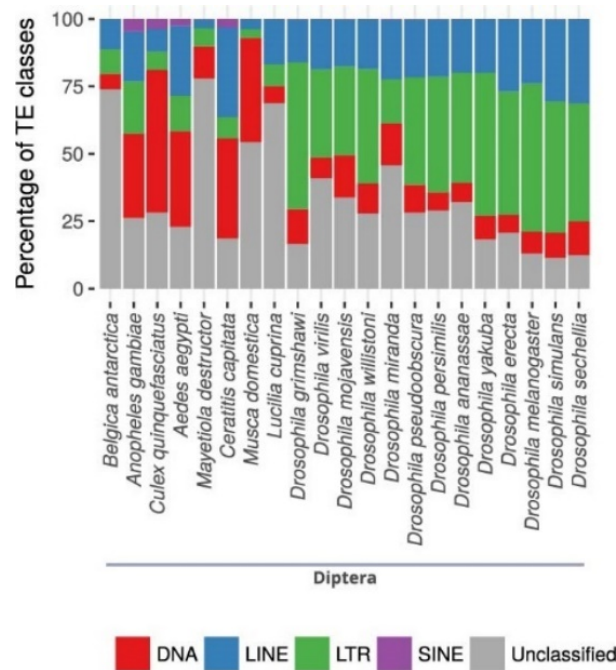


Figure 53 : Proportion des différentes familles de TE chez les diptères. Adapté de Petersen *et al.* 2019.

## Rôle des éléments transposables dans l'évolution des génomes

Il est généralement admis que la taille des génomes est corrélée à la quantité d'éléments transposables (Kidwell, 2002). De plus, lors de leur transposition, les éléments transposables peuvent avoir des actions délétères, par scission d'un gène par exemple. De même que leurs départs peuvent entraîner des cassures d'ADN (Bourque *et al.*, 2018). Les TE peuvent également être responsables de changement dans la structure même de l'ADN par la formation de méthylation (McDonald *et al.*, 2005). Ils sont également suspectés d'avoir un rôle important dans la spéciation (Jurka *et al.*, 2007 ; Ricci *et al.*, 2018).

Par ailleurs, les éléments transposables peuvent également jouer un rôle agissant directement sur le phénotype de l'hôte. L'exemple le plus célèbre étant probablement le changement de couleur de la phalène du bouleau (*Biston betularia*) dont la couleur originelle, pâle, et la couleur de la nouvelle forme, noire, est due à la présence d'un large élément transposable dans un intron du cortex de gène (Hof *et al.*, 2016).

## Le séquençage

Après que la structure de l'ADN fut révélée au monde en 1953, la découverte de l'ADN polymérase en 1956 fut une autre découverte majeure. Elle permis en effet en 1977 d'avoir les premiers résultats de séquençage et en 1986 d'avoir la première publication sur la PCR (Réaction de Polymérisation en Chaîne).

*- Pour information, ce qui est résumé ici en 4 lignes, a nécessité des décennies de recherches, l'implication de centaines d'hommes et de femmes, des milliers d'heures de travail et a été récompensé par la distribution de 5 prix Nobel -*

Alors si évidemment, je ne vais pas rentrer dans les plus fins détails de l'histoire, bien qu'elle soit passionnante, je voudrais quand même vous parler des techniques de séquençage qui ont été employées au

cours de ma thèse, et plus particulièrement le MinION (figure 54) développé par Oxford Nanopore Technologie ([Branton et al 2008](#)).



*Figure 54 : MinION développé par Oxford Nanopore Technologie*

*Le MinION est un séquenceur portable, capable de séquencer de longs fragments d'ADN à une vitesse proche des 500 nucléotides par secondes.*

Le principe de base de la technologie MinION est la mesure d'une différence de potentiel à l'intérieur d'un pore protéique (de l'ordre du nanomètre) provoqué par le passage d'une molécule d'ADN (figure 55).



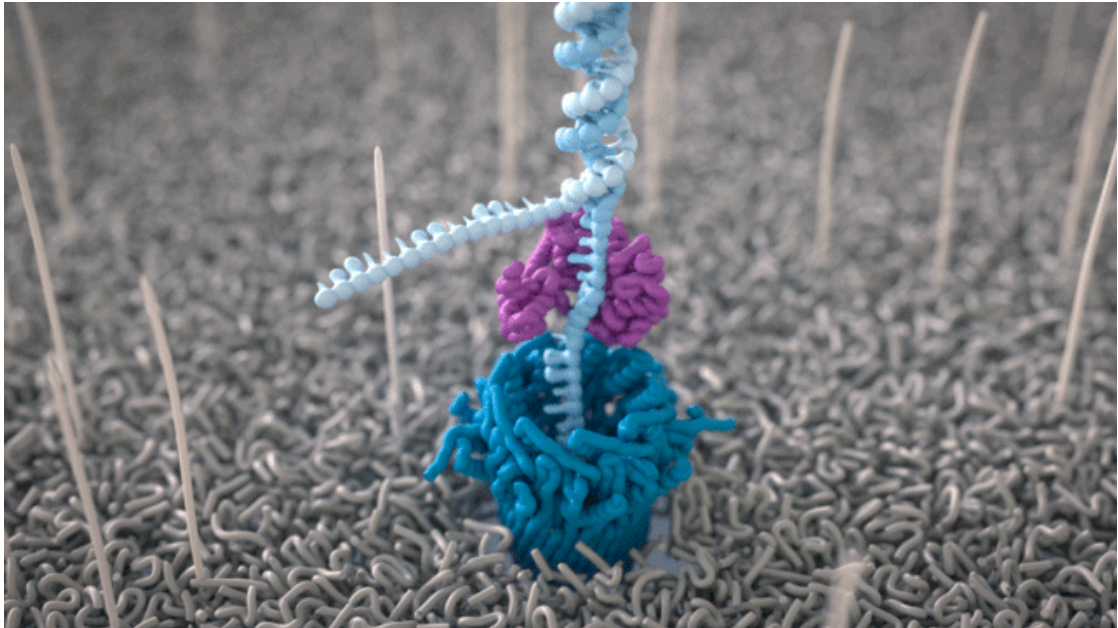


Figure 55 : Détail d'une molécule d'ADN séquencé dans l'un des pores moléculaires du MinION

*Le pore, en bleu, est enchâssé dans une membrane, en gris, l'un des deux brins d'ADN, en bleu clair, est guidé via une protéine, en violet, au travers du pore.*

Ces pores, (au nombre de 2048 dans les cellules actuelles) sont intégrés sur une membrane électro-résistante et sont chacun reliés à une électrode. La molécule d'ADN est guidée vers le pore au moyen d'une protéine attachée préalablement à la molécule d'ADN à séquencer et seul l'un des deux brins est guidé à l'intérieur.

Le passage du brin d'ADN au travers du pore, va provoquer une perturbation du champ électrique le traversant. L'électrode va alors mesurer le signal qui sera dû à la présence de bases à l'intérieur du pore. Le brin d'ADN continue d'avancer et l'électrode continue de mesurer à une vitesse de 500 bases par seconde. Le signal électrique étant spécifique de la présence de certaines bases dans le pore, il est transcrit directement en séquence d'ADN.

Le premier avantage de cette technologie, est la vitesse de séquençage et la taille des fragments mesurés. C'est-à-dire que l'on peut désormais obtenir un génome complet en quelques heures. Par ailleurs, les technologies Nanopore sont capables de séquencer des fragments de plusieurs milliers de bases. Un autre avantage apporté par MinION est sur le coût de séquençage, en effet le coût moyen est



désormais bien plus faible. Enfin il s'agit d'une technologie portable qui a pour vocation d'être employée directement sur le terrain, permettant ainsi un séquençage rapide *in situ*.

Cependant, cette technologie n'est pas exempte de défaut. Le principal problème étant le taux d'erreur que peut être amené à faire le séquenceur, principalement au niveau des homopolymères. Le logiciel de base Calling, fourni par le constructeur, est capable de faire diminuer ce taux d'erreur, mais il semble qu'en 2021, il restait autour de 6% (Delahaye et Nicolas 2021). Cela dit, il est possible de combiner les technologies de séquençage pour corriger les longs fragments séquencés par le MinION qui comportent des erreurs avec des fragments, plus courts, obtenus avec une technologie qui présente moins d'erreurs. Ce séquençage hybride offre actuellement le meilleur des deux mondes.

Une fois les séquences obtenues, il convient d'assembler les différents fragments d'ADN pour tenter de reformer le plus fidèlement possible la molécule initiale (chromosome, ou ADN mitochondrial par exemple). Le principe de l'assemblage est en soit assez simple : il repose sur l'homologie de deux ou plusieurs fragments sur un espace donné. Il existe deux méthodes qui sont principalement utilisées dans les logiciels d'assemblage moderne : la méthode OLC (pour Overlap Layout Consensus – figure 56) et une méthode basée sur des graphes de de Bruijn (Li et al 2011). Parfois, comme dans le cas du logiciel MaSuRCA les deux méthodes sont combinées.

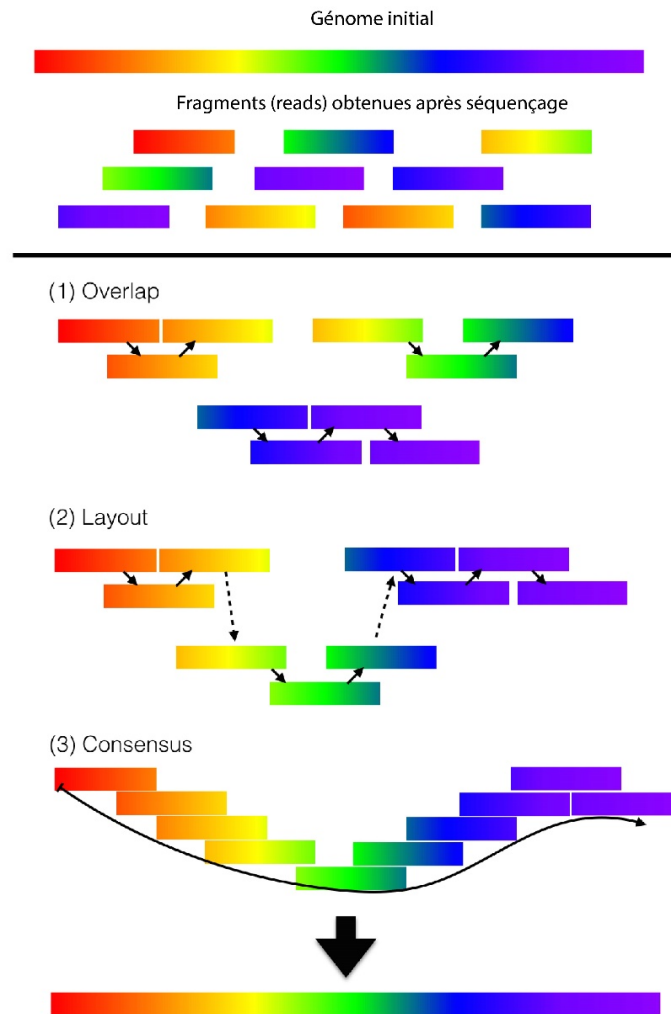


Figure 56 : Exemple d'assemblage OLC

*A partir d'un génome extrait des cellules d'un tissu, ou d'un organisme (donc présent en plusieurs copies), par séquençage on obtient des fragments (reads) d'ADN. Le logiciel réalisant l'assemblage va alors repérer les régions se chevauchant (Overlap), et celles-ci vont être assemblées, simplifiées et corrigées par l'étape suivante (Layout), enfin la séquence la plus probable est conservée pour chaque contig (Consensus).*

Quelquefois, il n'y a pas assez d'information génétique pour pouvoir assembler de manière fiable les fragments jusqu'à la reconstruction du chromosome. On parle alors de contig. On peut par la suite rassembler ces contigs entre eux en acceptant une plus grande incertitude. On parle alors de scaffold (figure 57).

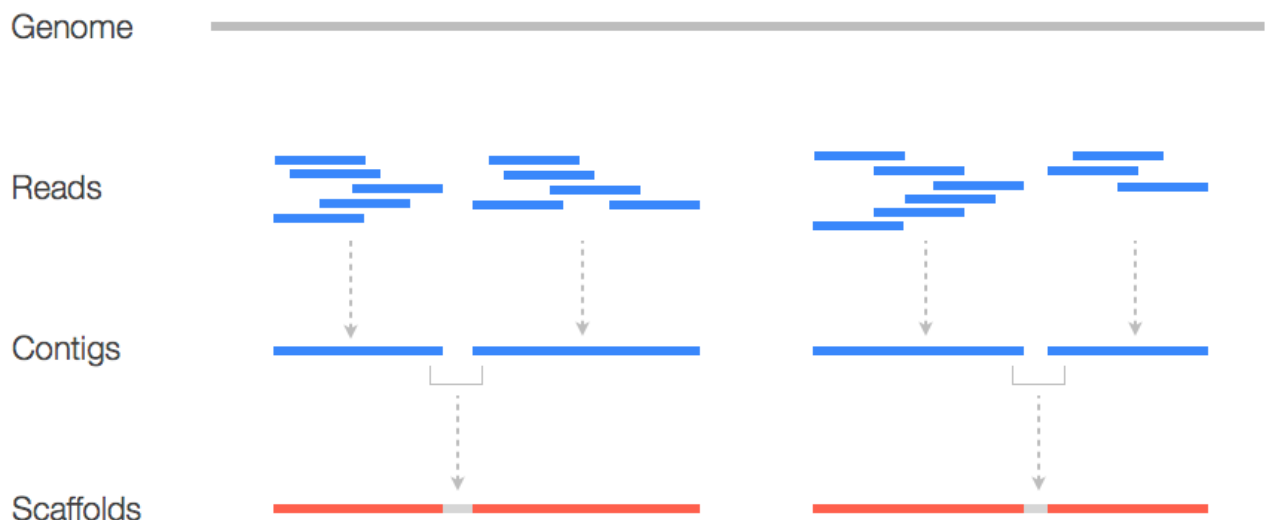


Figure 57 : Résumé d'un assemblage avec une étape de scaffolding

Les contigs obtenus peuvent être de nouveau assemblés, via une étape de scaffolding. Il s'agit de rassembler des contigs, dont la séquence est robuste, entre eux et dont la distance entre eux est connue mais pas la séquence.

## Génétique des populations

---

*Les plus grandes idées sont les plus simples  
William Golding - Sa majesté des mouches*

---

Si nous avons au cours de la première partie défini ensemble les notions d'individu et d'espèces, nous n'avons pas parlé encore de populations. J'ai volontairement laissé ce sujet en suspens pour la fin.

Pourtant si l'on reprend la définition d'espèce, nous y parlons déjà de population.

Alors je vous propose de définir maintenant de quoi nous parlons, et surtout de quoi nous parlerons.

Pas besoin de faire un grand suspense ici, une population c'est tout simplement un ensemble d'individus appartenant à la même espèce, inter-fertiles et vivants sur un territoire de taille suffisamment restreinte pour qu'ils puissent tous s'entrecroiser. Les échanges entre populations sont supposés rares.

La génétique des populations permet de mesurer la diversité génétique au sein et entre les populations. En comparant la distribution de la diversité génétique, la génétique des populations permet également d'expliquer l'origine des populations ainsi que leurs routes d'invasion.

### La diversité génétique et son étude

On entend par diversité génétique, les variations des gènes entre les individus ou entre les populations. Par ailleurs, une diminution de la diversité génétique au sein d'une population, ou d'une espèce, peut avoir des effets dramatiques, allant d'une diminution des possibilités d'adaptation à un environnement changeant (Frankham, 2003), jusqu'à une extinction de l'espèce ou de la population (Frankham, 2005). Afin d'étudier cette diversité génétique, on regarde le polymorphisme génétique, c'est-à-dire les variations entre les

séquences d'ADN des différents individus (pour un gène, un groupe de gènes, un locus ou encore une combinaison de locus). Bien que la plupart des mutations entraînant un polymorphisme génétique n'aient aucune influence, on parle alors de mutations silencieuses (Kimura, 1983), ou soient délétères, certaines d'entre elles vont apporter de la diversité au sein des populations ainsi qu'un potentiel d'adaptation.

Pour étudier les populations, et retracer des histoires évolutives et invasives, l'accès au génome (ou à certaines parties) est une étape indispensable. Par la suite on peut s'intéresser à plusieurs marqueurs moléculaires dont deux principaux : les microsatellites et les SNP.

Les microsatellites, ces courtes séquences d'ADN (1 à 4 nucléotides) sont répétées en tandem dans les génomes. La longueur totale, varie d'un individu à l'autre mais sont conservés pour des individus phylogénétiquement proches. En conséquence, il est possible d'utiliser ces séquences afin de déterminer les relations entre des individus d'une même espèce. On peut ainsi déterminer les structures des populations, les flux migratoires, ainsi que les possibles routes d'invasion. Les microsatellites ont l'avantage d'être neutres au regard de la sélection naturelle, c'est-à-dire qu'ils n'y sont pas soumis. En revanche, il n'est pas possible de déterminer via cette méthode des temps de divergence entre les populations. Pour cela, d'autres méthodes existent et elles se basent sur l'apparition de mutation, SNP (pour Single Nucleotide Polymorphism).

Quand les populations sont suffisamment éloignées, le nombre de différences mutationnelles entre les séquences, indique les temps relatifs de divergence. Si l'on suppose un taux de mutation, on peut obtenir un temps absolu de divergence. Cette méthode a été mise au point dans les années 60 et a depuis été affinée (Zuckerandl and Pauling 1965).

Chez *Hermetia illucens*, il a été prouvé que l'élevage en masse provoquait une chute drastique de la diversité génétique ainsi qu'une divergence rapide entre populations sauvages et populations captives (Hoffmann et al., 2021).

Enfin, et avant même de parler d'amélioration, le choix d'une souche peut considérablement impacter les performances zootechniques obtenues (figure 58). L'étude des interactions entre une génétique

particulière et le substrat (analyse GxE) est primordiale dans le cadre d'élevages de masse, que ce soit pour choisir la meilleure génétique parmi plusieurs, ou la plus adaptée en fonction du substrat (Sandrock et al., 2022).

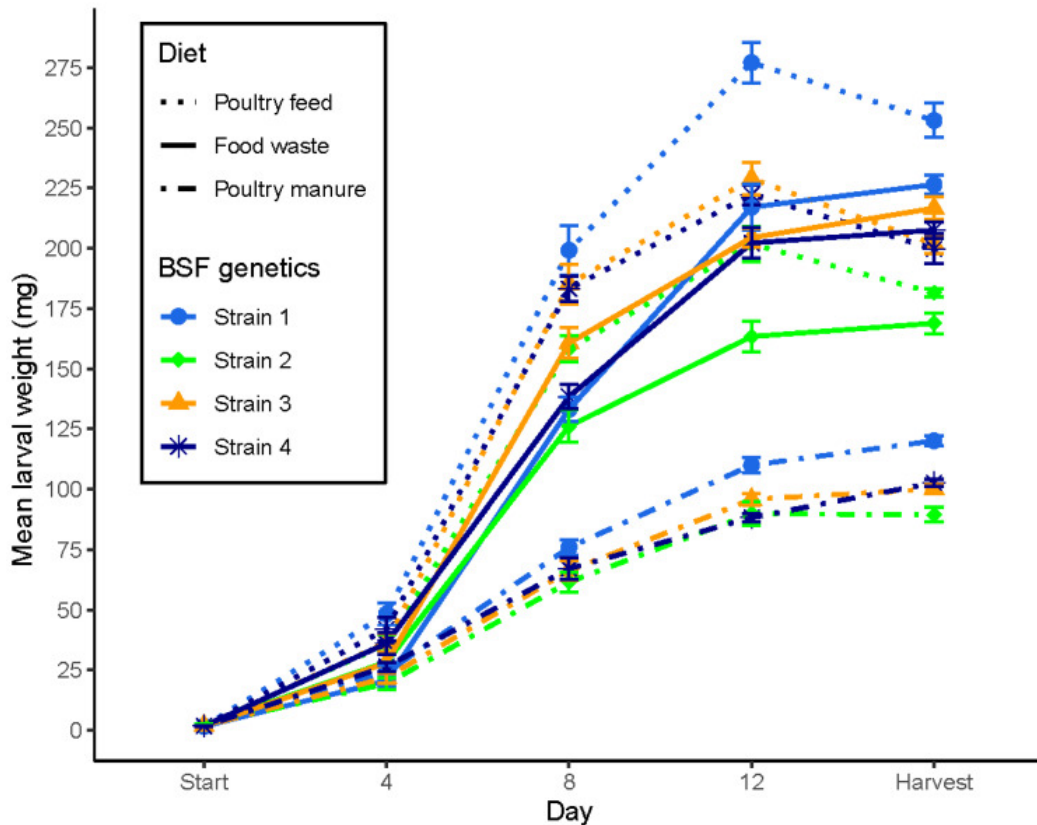


Figure 58 : Effets comparés de souches génétiquement différentes sur différents substrats

La comparaison de quatre souches génétiquement différentes sur trois types de substrats, montre dans cette étude que peu importe le substrat considéré, une souche est toujours supérieure aux autres en termes de gain de masse (Sandrock et al., 2022).

### *Hermetia illucens*, comme espèce d'étude en génétique des populations

Le CO1 pour cytochrome oxydase 1, est un gène mitochondrial (voir figure 48) qui est facilement séquençable et qui diffère d'une espèce à une autre tout en restant suffisamment proche pour définir de

quelle espèce il s'agit. Utilisé en routine dans l'identification d'espèces, il peut également s'avérer utile pour distinguer des populations. C'est donc sans grande surprise que celui-ci a été utilisé dès 2017 en Corée du Sud, pour déterminer la diversité génétique des BSF du territoire (Park *et al* 2017) ; 10 haplotypes ont été retrouvés en échantillonnant 245 individus sur l'ensemble du territoire (figure 59).

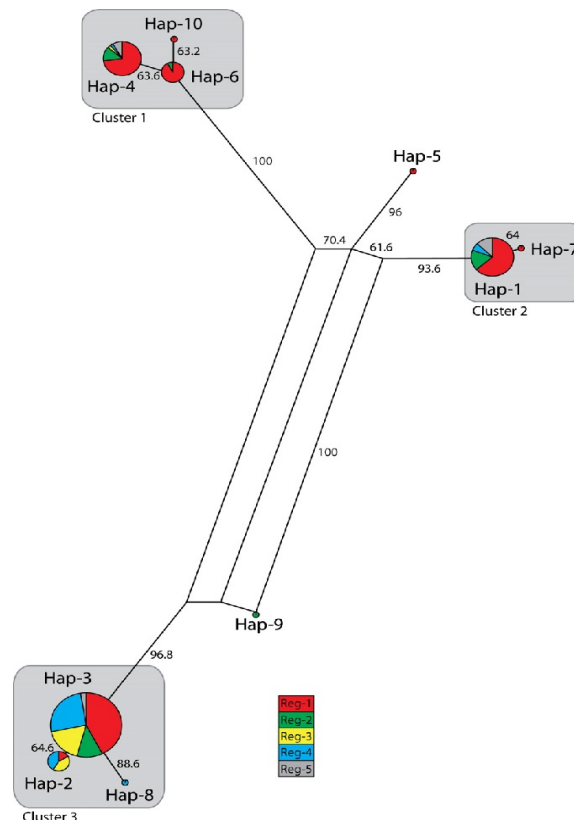


Figure 59 : Réseau d'haplotype CO1 d'*Hermetia illucens* en Corée du Sud

*En utilisant les séquences de 245 individus provenant de 5 régions de Corée du Sud, les auteurs ont découvert 10 haplotypes via un algorithme « median joining ». 3 clusters ont été délimités avec des haplotypes très proches les uns des autres (Park et al 2017)*

Les premières conclusions de cet article, s'étonnaient de la distance génétique entre les haplotypes trouvés. Les auteurs ont émis l'hypothèse, qu'une dispersion immédiate à la suite de l'introduction, pouvait expliquer les résultats qu'ils avaient. Ils complètent en pointant du doigt que cette propagation serait probablement due à l'activité humaine vu les faibles capacités de dispersion de la BSF.

Une autre étude, 3 ans plus tard (Stahls *et al* 2020), a quant-à-elle, collecté des BSF dans plusieurs régions du monde (7 régions

climatiques différentes) et également dans plusieurs entreprises d'élevage. La reconstruction d'un réseau d'haplotypes a permis la découverte de 56 haplotypes avec près de 600 individus séquencés (figure 60). Ces séquences de 658 nucléotides de long, n'avaient que 12% de sites variables (c'est-à-dire informatifs pour la reconstruction d'haplotypes). Les auteurs ont également montré que bien que les haplotypes les plus éloignés aient une divergence de 4,9% les individus étaient toujours inter-fertiles.

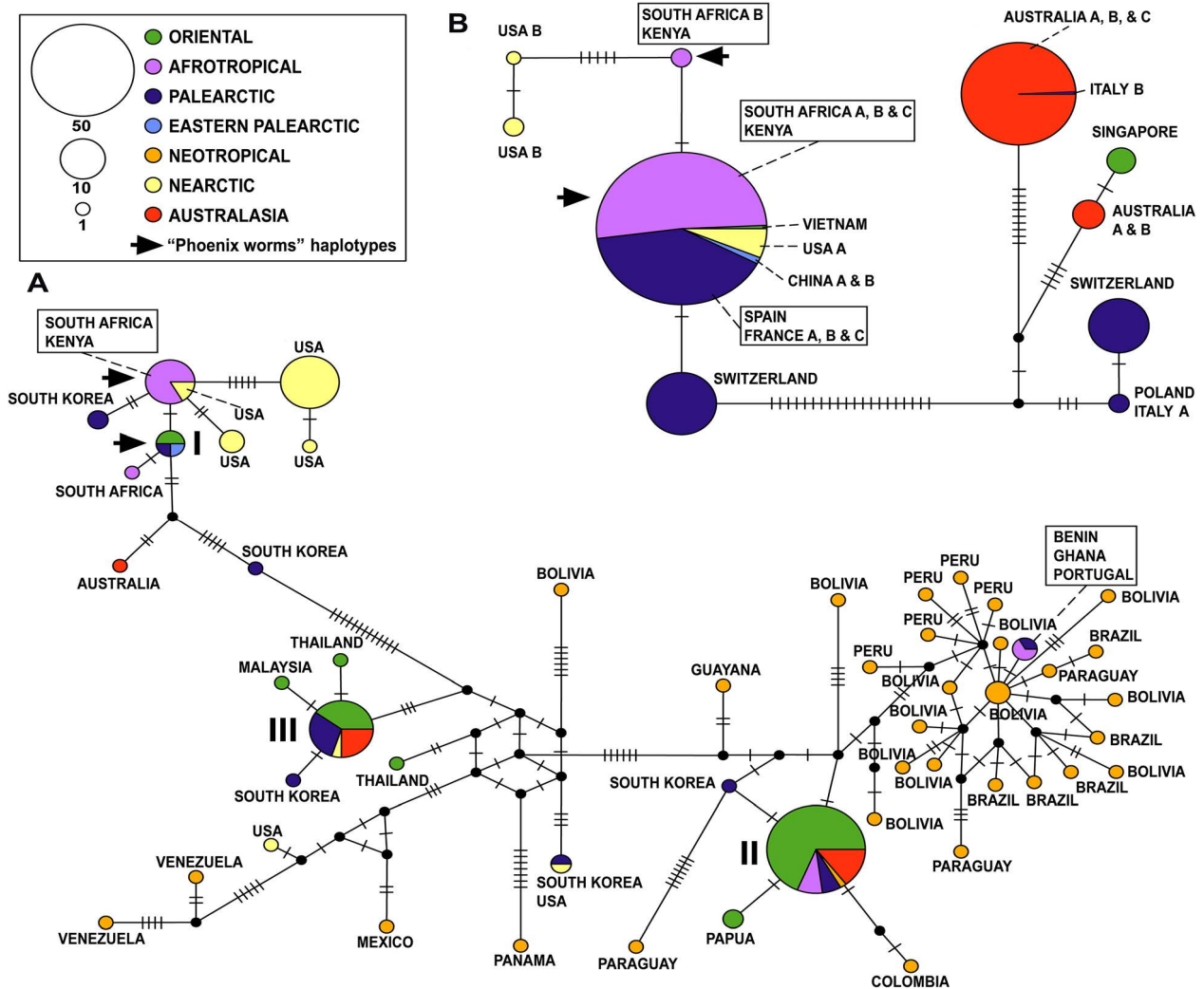


Figure 60 : Réseaux d'haplotypes CO1 après un échantillonnage mondial sur 7 zones biogéographiques différentes.

A : Réseau d'haplotypes d'*Hermetia illucens* capturées dans la nature. L'algorithme de reconstruction utilisé est le « median joining » les cercles noirs représentent les individus putatifs non échantillonnés. Le nombre de traits entre deux cercles représente le nombre de mutations. B : Réseau d'haplotypes d'*Hermetia illucens* provenant d'élevages (Stahls et al 2020).



En plus de présenter ces résultats sur la diversité génétique, l'équipe de Stahls (2020) a aussi fait remarquer une diversité phénotypique, tant au niveau de la taille des individus, que de la coloration de l'abdomen allant du noir au marron. Par ailleurs, ils précisent que si la divergence qu'ils ont observée sur le CO1 s'étend à tout le génome mitochondrial, les haplotypes les plus distants se seraient séparés il y a 0,74 à 1,35 million d'années. Ils remettent en question, par ce résultat, soit l'origine géographique de la BSF, soit le rôle de l'être humain dans sa dispersion.

Une autre étude, l'année d'après ([Kaya et al 2021](#)) viendra cependant apporter une autre hypothèse, expliquant les écarts génétiques entre les différentes populations. En utilisant 15 marqueurs microsatellites sur un échantillonnage de 2862 individus collectés dans 57 pays (figure 61).

Les auteurs ont tout d'abord montré la présence de points chauds génétiques dans le centre de l'Amérique du Sud, tendant à prouver qu'il s'agit du berceau de l'espèce, ainsi que des expansions successives vers l'Amérique du Nord. Pour les autres continents où l'espèce a été retrouvée (tous, pôles exceptés), ils montrent soit des effets de fondations issus d'introductions indépendantes, soit l'introduction d'un mélange d'individus génétiquement distincts.

Ils montrent ainsi, qu'une population tête de pont, a colonisé la Polynésie, puis la région australienne. Les descendants de cette dernière région, mélangés à d'autres profils génétiques auraient alors envahi l'Afrique puis l'Europe.

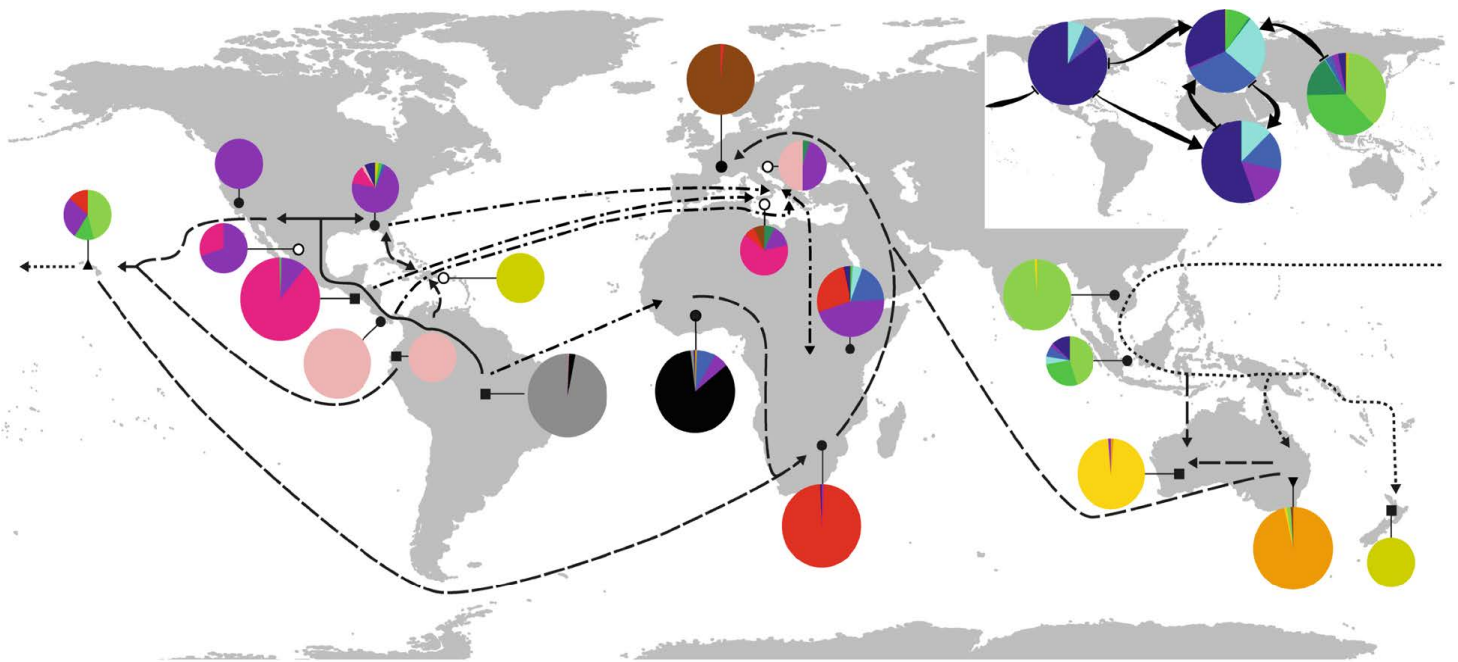


Figure 61 : Dynamique de la distribution mondiale déduite des groupes génétiques d'*Hermetia illucens* déterminés via des marqueurs microsatellites.

Carte principale : la taille des camemberts est proportionnelle au nombre d'individus. 16 clusters différents ont été trouvés, représentés par 16 couleurs différentes. L'expansion depuis l'aire d'origine est représentée par des flèches pleines. Les flèches pointillées représentent les colonisations en séries. Les flèches en tirets représentent des mélanges d'introduction démographiquement indépendantes et les flèches avec des points et des traits représentent des événements fondateurs issus de sources uniques. Carte secondaire : représentation des liens entre les populations captives des différents continents (Kaya et al 2021).

Ce qui est également intéressant dans cette même étude, est la proximité génétique entre les individus captifs à l'échelle mondiale (figure 61 - carte secondaire). Enfin, ils concluent qu'outre le rôle que la BSF pourrait avoir en tant qu'intermédiaire écologique (pour la réduction des déchets et l'apport de nourriture pour les élevages traditionnels), elle semble être une candidate intéressante pour étudier la divergence rapide et l'adaptation locale dans un cadre écologique des insectes envahissants.

## Sélection et amélioration

L'étude de la diversité génétique n'est pas une fin en soi dans le cadre de la BSF. Comme nous l'avons vu, elle présente un intérêt industriel non négligeable. Les entreprises se lancent donc dans des processus de sélection afin d'améliorer les populations qu'ils possèdent. Les buts vont varier selon les objectifs de chacun mais les techniques employées seront globalement les mêmes, à savoir sélectionner dans une ou plusieurs populations les individus ayant des caractéristiques zootechniques d'intérêt.

Plusieurs entreprises se sont donc déjà lancées dans cette course à l'amélioration. Les premiers à avoir proposé une amélioration significative, l'on fait en réalisant des mutants par la méthode CRISPR-Cas9 (Zhan *et al* 2020). Ils ont tout d'abord réalisé des mutants KO pour le gène PTTH (prothoracicotropique hormone), qui produit une hormone responsable de la mue et de la métamorphose des insectes. Le résultat obtenu était des larves significativement plus grosses et plus lourdes (figure 62) le dernier stade larvaire passant également de 4-5 jours à plus de 85 jours chez les mutants.



Figure 62 : Comparaison des mutants PTTH et sauvages

Les *Hermetia illucens* sauvages (indiquées WT), sont significativement plus petites et plus légères que les mutants pour l'hormone prothoracicotropique (notées  $\Delta$ HiPtth). La méthode d'édition de gènes CRISPR-Cas9 a été utilisée pour produire ces mutants (Zhan *et al* 2020).

Cette même étude (Zhan *et al* 2020), a également fait des mutants sans ailes, toujours avec la méthode CRISPR-Cas9. Ceci pourrait permettre d'éviter les échappements dans les entreprises. Ils ont

ciblé le gène *Vg* (Vestigial). Les auteurs présentent donc des mouches adultes sans ailes (figure 63) indiquant qu'il n'y a pas d'autres anomalies de développement ou morphologique. Cependant, sans préciser si ces individus étaient toujours capable de se reproduire.

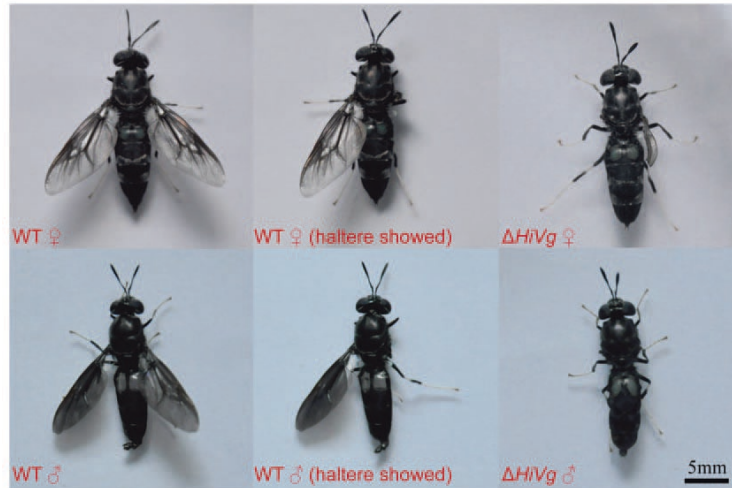


Figure 63 : Comparaison des mutants sans ailes et sauvage

*Les BSF sauvages (WT) ont encore leurs ailes (celles pour le vol et celles en haltères) tandis que les mutants *Vg* ( $\Delta HiVg$ ) n'ont plus d'ailes et ne présentent pas d'autres caractéristiques morphologiques.*

Par ailleurs, les premières études visant à améliorer les souches commerciales, sans manipulations génétiques, font leur apparition (Cai et al 2022). Une comparaison entre une souche domestiquée et une population consanguine issue de cette souche (ponte d'une seule femelle) n'a pas montré de différence morphologique ni de différence sur le temps de développement. En revanche, le nombre d'œufs était significativement plus important dans la souche consanguine que dans la souche domestique. Les auteurs ont donc émis l'hypothèse que la consanguinité pourrait améliorer la production d'œufs (figure 64).

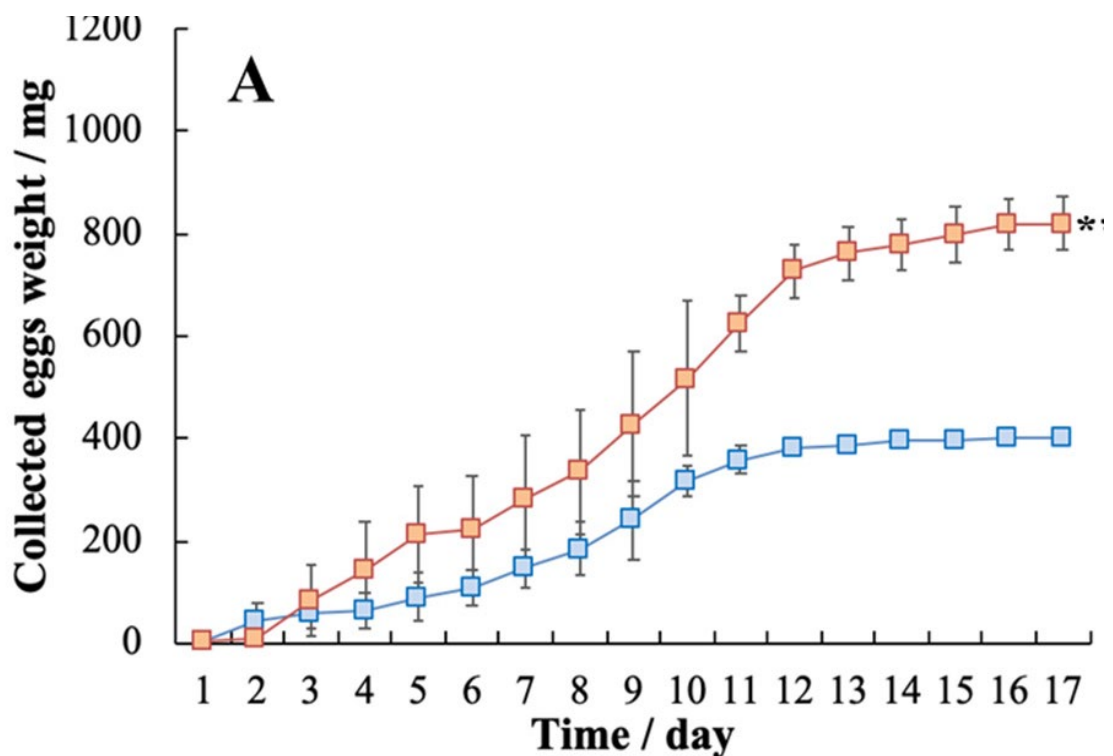


Figure 64 : Evolution du nombre d'œufs collectés entre une souche d'élevage (bleu) et une souche consanguine issue de la souche d'élevage (rouge).

La souche consanguine pond en moyenne plus d'œufs que la souche d'élevage (d'après Cai et al 2022)

Il faut cependant prendre du recul sur cette publication, car les variations observées peuvent être dues à la variabilité génétique déjà présente au sein de la population de base sans lien avec la consanguinité.

Un programme de sélection massale qui a commencé en 2019, au sein de l'entreprise Protix et qui s'est étalé sur 16 générations a montré un gain de masse moyen des individus de 39% (figure 65). Notamment un gain de 32% pour les protéines et 21% pour les lipides (Facchini et al., 2022).

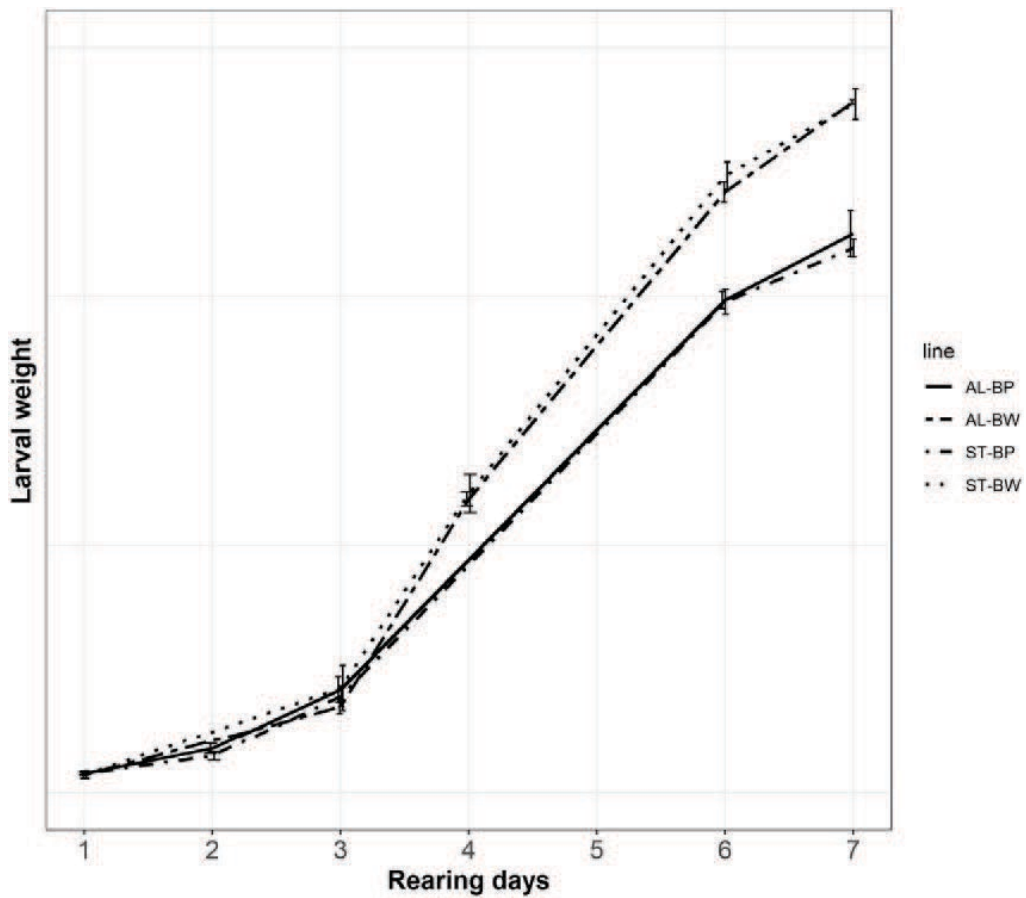


Figure 65 : Comparaison du poids des larves sous sélection massale

Deux lignées sont comparées : la lignée de base (BP) et la lignée sélectionnée via le processus de sélection massale (BW). Pour chacune, deux conditions de nourriture sont appliquées soit standard (ST) soit ab libitum (AL). Quelques soient les conditions de nourriture la lignée sélectionnée pour le poids est significativement supérieure à la population de base à partir du 4<sup>ème</sup> jour d'élevage (Facchini et al., 2022).

Ces résultats, bien que prometteurs sont aussi à prendre avec un certain recul, les auteurs n'ayant pas partagé dans leur publication les poids exacts des larves (des deux lignées) ainsi que la pression de sélection apposé sur la lignée sélectionnée.

# Problématiques

---

*Au bout d'un chemin il y a toujours au moins une question.*  
JG

---

Nous voici à la fin de notre introduction, et vous voilà prêt pour aborder les chapitres qui vont suivre avec les connaissances nécessaires pour comprendre mes avancées. Vous avez probablement appris quelques éléments (et moi aussi en la rédigeant). Rappelons tout de même, les principaux points de discussion et questions qui ont été à la base de cette thèse et qui constitueront le corpus des trois prochains chapitres de ce manuscrit :

- Bien que ce ne soit plus le cas désormais, au démarrage de cette thèse en janvier 2019, le génome et le catalogue de gènes d'*Hermetia illucens* n'était pas connu.
- Très peu de données existent sur la génétique et la génomique de la BSF, or ce point est crucial pour espérer améliorer les performances de productions via des procédés de sélection par exemple.
- L'histoire naturelle de l'espèce, en particulier le scénario de dispersion n'est pas encore fixé et plusieurs scénarii sont encore envisageables.
- Peut-on retrouver dans les différentes populations des liens avec les zones biogéographiques d'origine ? De même, peut-on dater les séparations entre les différents haplotypes déjà identifiés ?
- Quelle est la proportion d'éléments transposables dans le génome d'*Hermetia illucens* et celle-ci est-elle corrélée à différentes populations ?

# Chapitre premier : Un génome pour les gouverner tous.

## Résumé

Mettre en élevage une nouvelle espèce implique de la domestication. Vouloir adapter l'espèce à des environnements et des conditions requiert de la sélection. Effectuer ces travaux aujourd'hui, peut-être grandement facilité par l'apport de la génétique, des connaissances des gènes et de la diversité des différentes populations présentes et à notre disposition.

Aussi pour ce qui est de la mouche soldat noire, très peu d'informations quant à sa génétique étaient disponibles au démarrage de ma thèse en janvier 2019. Je me suis lancé dans une manipulation passionnante qu'est le séquençage, l'assemblage et l'annotation d'un génome qui pourrait servir de référence. Je me suis senti tel un pionnier dans ce domaine. Malheureusement la compétition est parfois rude et avant que je n'aie l'occasion de terminer ce travail, une autre équipe publiait leurs résultats avec un génome de référence. Aussi ce premier chapitre est un travail inachevé, une pointe de déception, mais surtout une grande quantité de connaissance que j'ai accumulée.



## Introduction

*Hermetia illucens*, la mouche soldat noire, est devenue au cours de la décennie précédente un animal d'élevage (Gold et al., 2018), un moyen de diminuer les déchets organiques (Amrul et al., 2022) tout en produisant des protéines, des lipides et d'autres molécules dont nous avons parlé dans l'introduction du présent manuscrit (Wang and Shelomi, 2017).

Cependant, comme avec bon nombre d'animaux d'élevage, les besoins de l'être humain, ne se plient pas toujours à la réalité biologique des êtres vivants. Aussi, les hommes et les femmes entrent-ils dans un processus de sélection sur les animaux et végétaux qu'ils élèvent. Les processus de sélection sont nombreux et peuvent se baser sur plusieurs paramètres tel que, par exemple, la sélection massale, qui consiste à ne garder d'une génération à l'autre, que les individus les plus gros. De même, la domestication est également un processus de sélection (Driscoll et al., 2009). En effet, à propos de la mouche soldat noire, on ne conserve que les individus qui ont pu se reproduire et pondre dans les systèmes d'élevage que nous mettons en œuvre pour les élever.

Une fois ces premières étapes de domestication réalisées, nous composons avec la diversité génétique restante au sein de nos élevages. Selon la taille des populations et fonction de la dérive génétique, la diversité génétique des élevages va diminuer. Cette diminution a été montrée chez de nombreuses espèces végétales (El-Kassaby and Ritland, 1996 ; Ait Aabd et al., 2019) et animales (Notter, 1999 ; Barker, 2011). Pour le dire autrement, avec la domestication, une population restreinte et une sélection massale, par exemple, la consanguinité des élevages augmente (De Donato et al., 2005). La consanguinité est la production de descendants par des individus proche génétiquement Cette dernière va provoquer une diminution de l'hétérozygotie, réduire la diversité allélique (Calleri et al., 2006) qui peut impacter la résistance aux maladies (Paterson et al., 1998).

La consanguinité peut être un problème et ce, même si les conditions se maintiennent de façon absolument identique d'une génération à l'autre (Turner, 1960). Ce qui n'arrive jamais car cela serait faire fi des

variations de température, d'humidité, ou encore de l'arrivée d'un pathogène dans les élevages. Aussi, il est indispensable, pour les éleveurs de BSF, de, non seulement maîtriser la biologie générale de la mouche soldat noire, mais également d'être capable de faire de la sélection sur l'espèce, ou encore, de croiser différentes populations pour améliorer leurs élevages. Ces améliorations passent par plusieurs points, notamment connaître la diversité génétique de l'espèce, aussi bien dans les élevages que sauvages. Cela passe également par une connaissance génétique plus fondamentale qu'est le génome de l'espèce élevée.

En effet, connaître le génome d'une espèce, avoir la carte génétique des gènes permet de relier un génotype à un phénotype comme cela a déjà été fait chez *Dendroctonus ponderosae* (Cullingham et al., 2020).

En 2019, aucun génome complet de BSF n'était publié à ce moment-là. Seul était disponible le génome mitochondrial (Qi et al., 2017) ainsi qu'une esquisse de génome sous la forme de reads Illumina non assemblés (plusieurs millions de fragments de 100 paires de bases) (Vicoso and Bachtrog, 2015).

Aussi, me suis-je attelé au séquençage et à l'assemblage du génome, le but étant *in fine* d'annoter ce génome. Pour obtenir un génome d'une espèce, il convient d'en extraire l'ADN dans un premier temps. Il faut avoir l'ADN le plus pur possible et selon le type de séquençage, le plus long possible. En effet, lorsque l'on va extraire l'ADN des cellules, celui-ci va se casser en fragments de tailles variables. Plus les méthodes pour l'extraire et le purifier seront brutales plus les fragments seront de petites tailles.

Ensuite, il faut effectuer le séquençage à proprement parler. Diverses techniques existent, reposant soit sur la complémentarité des bases, en incorporant des nucléotides marqués avec une molécule fluorescente, soit la différence de potentiel induite par la présence d'une séquence donnée de nucléotides, dans un pore moléculaire. Nous avons choisi cette double approche pour le séquençage du génome d'*Hermetia illucens*. Un séquençage illumina de deux individus, puis un séquençage de longs fragments en utilisant le séquenceur MinION de la société Oxford Nanopore Technologies.

Cependant un génome fut publié avant que mon travail ne soit terminé. Je vais donc vous présenter les résultats que j'ai obtenus précédés des méthodes que j'ai mises en place pour obtenir ce génome.

## Matériels et méthodes

### Matériel biologique

Les mouches soldat noires utilisées pour le séquençage de l'ADN proviennent de la société Cycle Farms. Elles ont été collectées au stade adulte et larvaire, et elles ont été stockées dans de l'alcool absolu et conservées à -20°C avant l'extraction d'ADN.

### Extraction et purification d'ADN

Nous avons effectué des extractions d'ADN des adultes et des larves. Le tube digestif des larves était préalablement retiré afin d'éviter les contaminations par le microbiote. Nous avons favorisé l'extraction d'ADN des adultes pour le séquençage, afin de différencier les mâles des femelles.

La première étape de l'extraction est une lyse cellulaire, pour cela les individus entiers (ou larves sans tube digestif) étaient broyés dans de l'azote liquide dans un mortier à l'aide d'un pilon. Le contenu était placé dans des tubes à hémolyse de 5 ml à 50°C, durant une nuit en utilisant les réactifs du kit d'extraction d'ADN Macherey-Nagel AXG 100. La suite du protocole s'est faite en suivant les recommandations de ce kit. Il s'agit d'une technique de chromatographie d'échange d'ions par gravité. Elle permet d'avoir un traitement léger sur l'ADN et d'éviter de le fragmenter. Nous avons également ajouté une seconde étape de purification à la fin, en ajoutant de l'éthanol 70% fraîchement préparé ainsi qu'une centrifugation (15 min, 15000 rpm à 4°C). L'ADN extrait est alors suspendu dans 100 µl d'un tampon TRIS-EDTA (pH 8).

## Contrôle de la pureté et dosage

Le contrôle de la pureté est réalisé à l'aide de la mesure des absorbances à 260 nm et 280 nm sur un Nanodrop (2000c ThermoFisher). Le dosage de la quantité d'ADN a été fait par une mesure de la fluorimétrie en utilisant un kit DNA Broad Range pour le Qubit (3.0 fluorometer Invitrogen) et l'intégrité de l'ADN a été évaluée par électrophorèse en gel d'agarose.

## Préparation de la librairie de séquençage MinION et illumina

L'ADN génomique extrait est composé d'un ensemble de fragments de différentes tailles. Afin de nous assurer que nous ne conservons que les fragments de plus grandes tailles pour le séquençage MinION, une étape utilisant le kit SRE (Short Reads Eliminator - Circulomics) est réalisée. La librairie a ensuite été préparée en utilisant le protocole d'Oxford Nanopore Technologie SQK-LSK 109.

D'autres essais ont également été réalisés afin d'essayer de ne garder que les plus grands fragments, c'est le cas des g-TUBE de la société Covaris pour récupérer des fragments de 6 à 20 kb.

Pour les librairies illumina 2 µg d'ADN génomique ont été envoyés à la plateforme Imagif : 1 µg provenant d'un mâle et 1 µg d'une femelle.

## Séquençage illumina et MinION

Le séquençage a été réalisé de deux manières différentes. Tout d'abord le microlitre d'ADN génomique a été séquençé par la plateforme Imagif, nous donnant pour les deux individus 20 équivalents génomes (soit 40X au total).

Ensuite plusieurs individus ont été séquençés via plusieurs flow cell MinION pour un total de 48X sur 4 individus. Nous avons utilisé le logiciel Guppy (V 6.1.3+) afin de réaliser le base Calling (transformation des informations brutes issues du séquenceur MinION en données fastq). La qualité et la taille des fragments issus

du séquençage MinION ont été analysés avec le logiciel NanoPlot (V1.22.0) issu du NanoPack (De Coster et al., 2018).

## Estimation de la taille du génome

L'estimation de la taille du génome a été réalisée en utilisant GenomeScope (Vurture et al., 2017) pour estimer la taille, l'hétérozygotie ainsi que la quantité d'éléments répétés au sein du génome.

## Assemblage du génome.

L'assemblage du génome a été fait en utilisant les séquences provenant des banques illumina et MinION. L'assembleur utilisé était Masurca (Zimin et al., 2013) (V 3.3.31) et Cabog. Les ressources bio-informatiques utilisées étaient celles du laboratoire EGCE. La vérification de la complétude du génome a été réalisée en utilisant le logiciel Busco (Simão et al., 2015).

## Résultats et discussion

Les extractions d'ADN nous ont procuré une grande quantité d'ADN génomique (Tableau 1). Sans différences significatives entre les mâles et les femelles, ni entre adultes et larves.

Adultes	Larves
1,1 ( $\pm$ 0,2) $\mu$ g	0,8 ( $\pm$ 0,3) $\mu$ g
28 réplcats	10 réplcats

Tableau 1 : Quantité d'ADN ( $\mu$ g) extrait de BSF

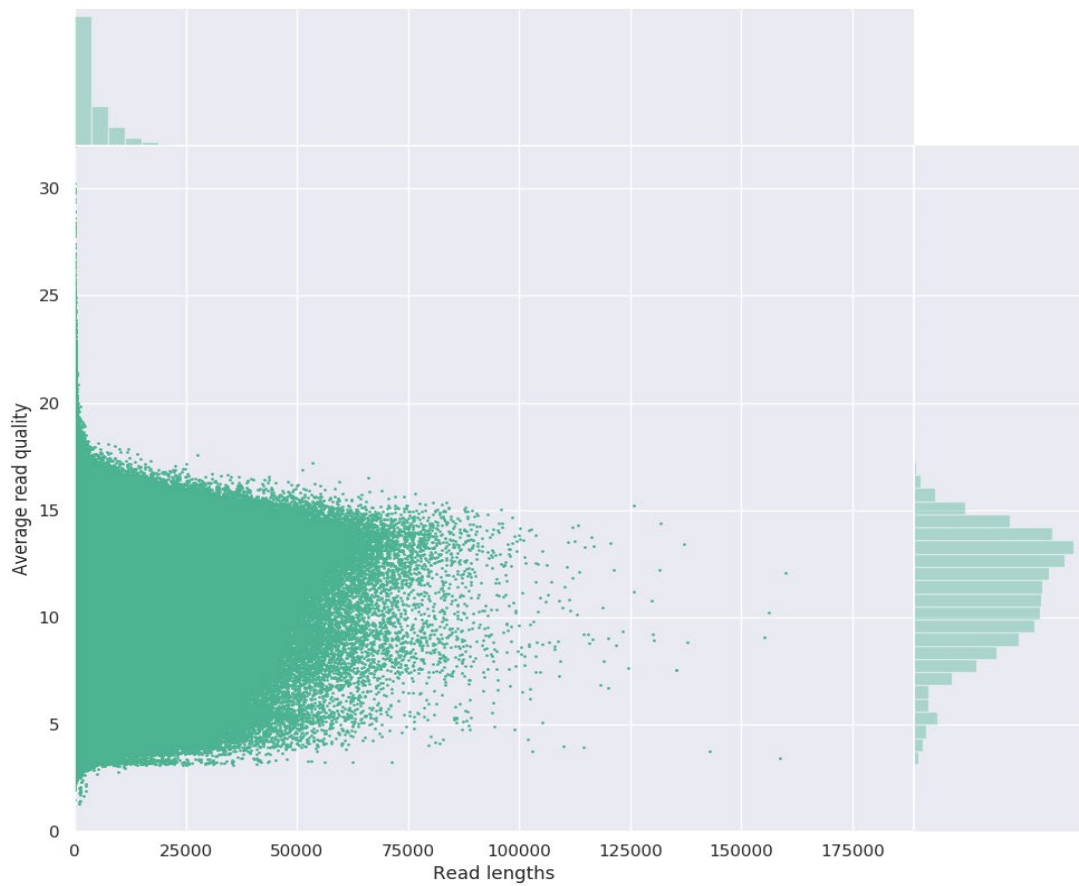
5 individus ont été séquencés au total, deux femelles et trois mâles. Des conditions différentes ont été appliquées en amont des séquençage MinION (SRE, g-TUBE ou aucun – tableau 2), parmi ces 5 individus, un mâle et une femelle ont été séquencés avec technique illumina.

Run	Traitement	Gigabase
Hilu01 (F)	g-TUBE	1,8
Hilu02 (M)	g-TUBE	7
JAFD94 b (F)	Aucun	11
	g-TUBE	0,9
	SRE	3,8
JAMD206 c (M)	SRE	13,7
JAMD206 d (M)	SRE	9,8
<b>Total</b>		<b>48</b>

*Tableau 2 : Résumé des individus séquencés au MinION ; femelle (F) ou mâle (M), des traitements appliqués ainsi que de la quantité de de bases séquencée. (g-tube de chez Covaris, ou kit Short Reads Eliminator de chez Circulomics)*

Les données de séquençage MinION, après le base-Calling via le logiciel Guppy ont été analysées avec NanoStats.

La taille moyenne des reads est de 4,5 kb avec un score de qualité phred de 11,2 (figure 1 et 2). Le plus long reads a une taille de près de 160 kb avec un score de qualité phred de 12,1.



*Figure 1 : Tracé point par point de la taille des fragments avec leur qualité en score phred pour l'ensemble des données MinION*

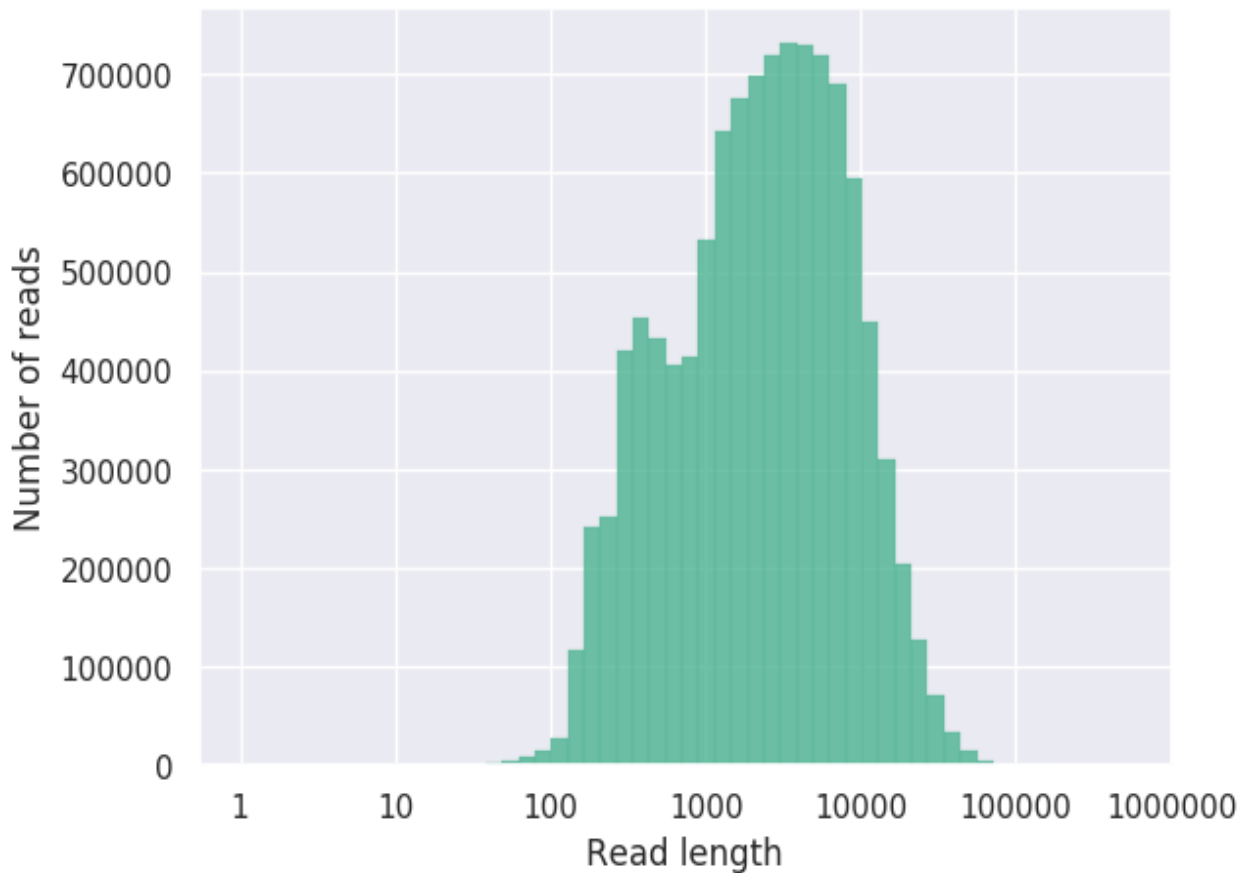


Figure 2 : Histogramme de la longueur des reads après transformation logarithmique pour l'ensemble des données MinION.

Les deux microgrammes d'ADN (un microgramme pour une femelle et un microgramme pour un mâle) ont été séquencés par la plateforme Imagif. Ils nous ont donné chacun environ 20X de taux de couverture avec un score de qualité phred de 30 en moyenne.

Nous avons estimé la taille du génome à l'aide de ces données Illumina et du logiciel GenomeScope en utilisant des Kmer de taille 21. D'après nos estimations le génome de la mouche soldat noire faisait environ 0,8 Gigabase (figure 3), cependant les assemblages qui auront lieu par la suite, trouveront une taille autour de 1,01 Gigabase (Generalovic et al ; 2021). Notre analyse estime qu'il y a environ 80%



de 21-mer unique ce qui suggère que ce génome a un contenu répété autour de 20%, or les analyses postérieures suggèrent un pourcentage d'éléments répétés autour de 60%.

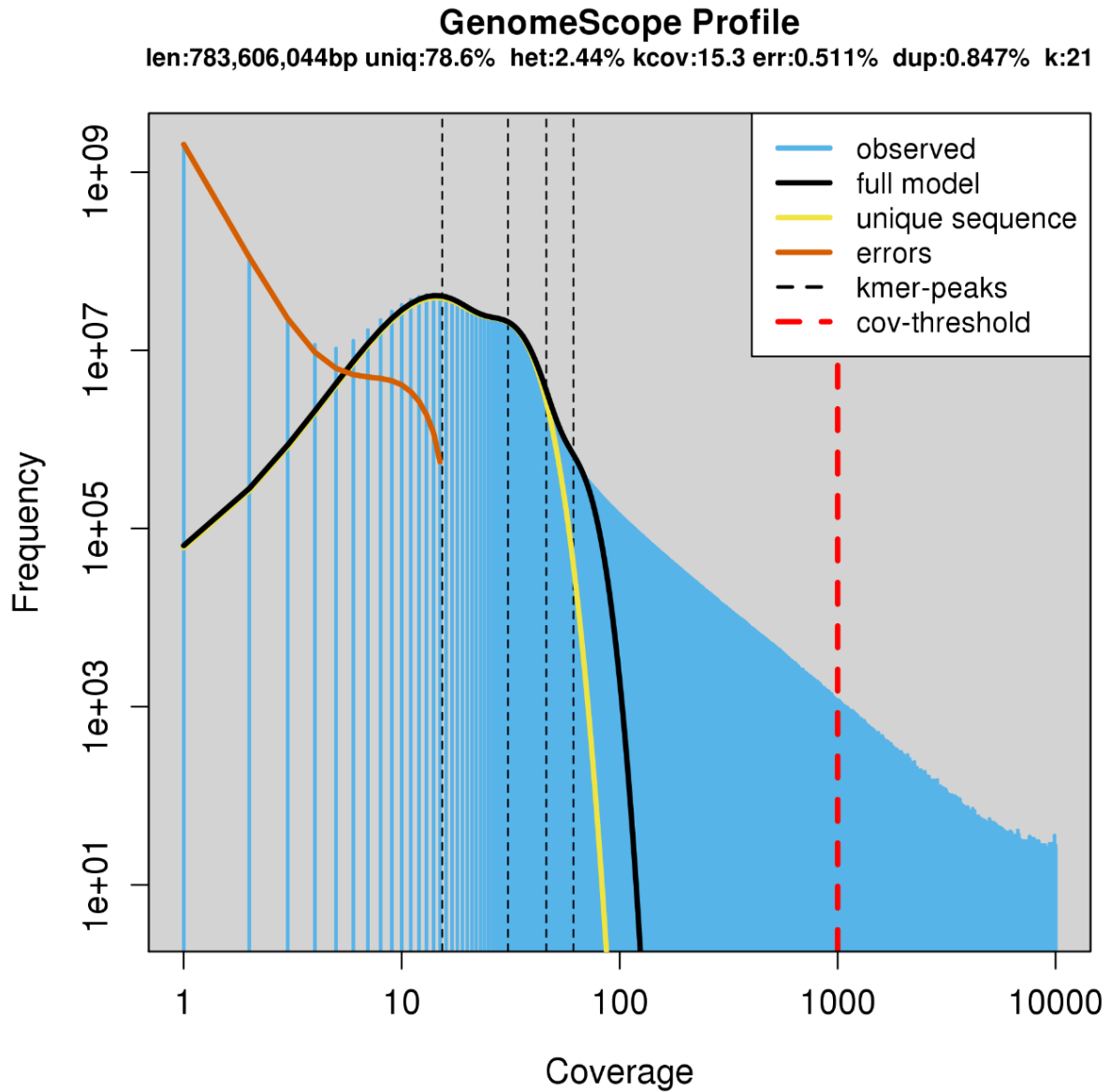


Figure 3 : Profile GenomeScope des 21-mer dans les données génomiques d'*Hermetia illucens*. La longueur a été estimée à environ 0,8 Gigabase avec un taux d'hétérozygotie à 2,44%.

Les Kmers de tailles 21 sont couramment utilisée pour estimer la taille des génomes. En revanche au cours de leur étude, l'équipe de Generalovic a utilisée des Kmers de taille 31 (Annexe 1.01 – p220). Afin de comparer nos résultats, nous avons refait l'expérience en utilisant également des Kmer de taille 31 et avec le même outil qu'eux (GenomeScope). En utilisant 80X de données nos résultats indique une taille de génome estimée à 843 Mb contre 1,01 GB pour l'autre équipe soit un écart d'environ 16,5%.

Cet écart pourrait être une preuve de variation importante entre les génomes des différentes *Hermetia illucens*.

Nous avons ensuite essayé différentes contraintes sur les données MinION. Notamment, ne garder que les 90% des séquences de meilleure qualité, ou encore différentes associations d'assembleurs. Finalement les meilleurs résultats ont été obtenus en utilisant l'ensemble de nos données illumina et MinION et l'assembleur Masurca associé à Cabog (Annexe 1.02 - p221).

Les métriques de cet assemblage étaient un N50 à 302 kb pour un total de 21000 contigs.

En revanche, la taille de génome estimée de cet assemblage était au-dessus des premières prédictions et se situait à 1,6 Gigabase.

Par la suite nous avons estimé la complétude du génome avec le logiciel Busco en utilisant la base de données diptères :

C :90.0%	[S :63.1%, D :26.9%], F :5.1%, M :4.9%, n :2799
2517	Nombre total de gènes retrouvés (C)
1765	Parmi les gènes retrouvés ceux présents en simple copie (S)
752	Parmi les gènes retrouvés ceux présents en plusieurs copies (D)
144	Gènes retrouvés mais incomplets - fragmentés (F)
138	Gènes absents (M)
2799	Gènes recherchés au total (n)

Un score Busco de 90 est correct mais ce score ne prend réellement sens qu'en le comparant à d'autres assemblages d'une même espèce.

Ayant essayé différentes contraintes et différents logiciels d'assemblage, ce score de complétude est également le plus haut que j'ai pu obtenir.

## Conclusion

Le travail sur ce génome n'était pas terminé, en effet il restait encore l'annotation à faire. Mais le génome a été publié par une autre équipe ([Zhan et al 2020](#)), avec de bons résultats (figure 4).

Genome size	1.1 Gb (1.7 Mb N50; 100% completeness)
Coding content	16,770 genes (2470 novel genes)
Repeat content	65.8%

Figure 4 : Résultat de l'assemblage de Zhan et al 2020

Nous avons donc deux choix qui s'offraient à nous, soit obtenir de nouvelles séquences et assembler un génome de meilleure qualité encore, c'est-à-dire arriver au niveau du chromosome. Soit poursuivre les travaux sur la diversité génétique des mouches soldat noires. C'est cette seconde option qui fut choisie et bien nous en a pris, car quelques mois plus tard un nouveau génome fut publié et cette fois-ci à l'échelle chromosomique ([Generalovic et al 2021](#)).

Ce n'est donc pas sans une pointe d'amertume que j'ai délaissé ce travail sur le génome de référence. Celui-ci m'avait pris une dizaine de mois. Mais j'ai pu progresser sur des aspects moins compétitifs que la publication d'un génome de référence d'une espèce dont les scientifiques et les industriels étaient en train de s'emparer. De plus, certaines données sur les éléments transposables issus de ces travaux seront intégrées dans le chapitre 3.

D'autre part, les données générées au cours de cette partie de ma thèse trouveront une utilité dans la suite de mes projets de recherche. En effet, nous allons dans le futur, à la suite de cette thèse, nous concentrer sur des comparaisons de génomes entiers entre les différentes populations de BSF.

# Chapitre 2 : Ce que les génomes mitochondriaux complets nous apprennent sur l'histoire évolutive de la mouche soldat noire, *Hermetia illucens*.

Guilliet et al. *BMC Ecology and Evolution* (2022) 22:72  
<https://doi.org/10.1186/s12862-022-02025-6>

BMC Ecology and Evolution

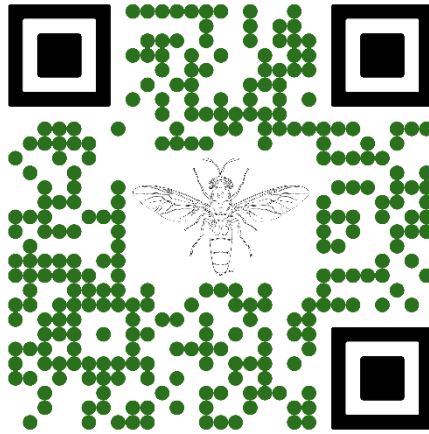
RESEARCH

Open Access

What complete mitochondrial genomes tell us about the evolutionary history of the black soldier fly, *Hermetia illucens*



J. Guilliet<sup>1,2,3\*</sup>, G. Baudouin<sup>2</sup>, N. Pollet<sup>1,3</sup> and J. Filée<sup>1,3\*</sup>



Le second chapitre de cette thèse a fait l'objet d'une publication dans le journal *BMC Ecology and Evolution* en mai 2022. Il traite de l'histoire évolutive de la mouche soldat noire. Pour cette étude, nous avons analysé 677 séquences du gène mitochondrial CO1 provenant de l'ensemble des zones de présence de la BSF avec un effort particulier sur la France. Ceci nous a permis de définir 52 haplotypes différents. Nous avons par la suite séquencé et assemblé 60 génomes mitochondriaux complets. Avec ce dernier jeu de donnée, nous avons pu déterminer les relations phylogénétiques entre les différents haplotypes et également dater la séparation entre les deux principales populations de BSF. Nous avons trouvé une séparation remontant à plus de deux millions d'années. En se basant sur l'histoire invasive de la BSF, nous avons finalement émis l'hypothèse que cette séparation de population serait concomitante avec la fermeture de l'isthme de Panama et donc l'invasion par la BSF de l'Amérique du Nord.

L'article est en annexe (p238) et accessible via le QR code ci-dessus.

Abstract

*Background*

The Black Soldier Fly (BSF) *Hermetia illucens* is a cosmopolitan fly massively used by industrial companies to reduce biowaste and produce protein and fat for poultry and aquaculture feed. However, the natural history and the genetic diversity of the BSF are poorly known. Here, we present a comprehensive phylogeny and time tree based on a large dataset of complete mitochondrial genomes better to understand the evolution and timing of the BSF.

*Results*

In this study, we analyzed 677 CO1 sequences derived from samples found all over the five continents, leading us to discover 52 haplotypes, including ten major haplotypes. This worldwide cryptic genetic and genomic diversity is mirrored at a local scale in France, in which we found five major haplotypes sometimes in sympatry. Phylogenetic analyses of 60 complete mitochondrial genomes robustly resolved the phylogeny of the major BSF haplotypes. We estimate the separation events of the different haplotypes at more than two million years for the oldest branches characterizing the ancestral split between present North American lineages and the other highly diverse south-central American clades, possibly the following radiation beyond the isthmus of Panama northwards. Our data confirm that this North American lineage ultimately gave birth to almost all commercial

BSF stocks that participated in the worldwide BSF dissemination through farm escapements.

### Conclusions

Our data resolve the phylogenetic relationships between the major lineages and give insights into the BSF's short and long-term evolution. Our results indicate that commercial BSF stock's genetic and genomic diversity is very low. These results call for a better understanding of the genomic diversity of the BSF to unravel possible specific adaptations of the different lineages for industrial needs and to initiate the selection process.

Keywords: *Hermetia illucens*; phylogeny; mitochondrial genome; phylogeography

## Background

The black soldier fly (BSF) *Hermetia illucens* (Linnaeus, 1758) is a Diptera of the Stratiomyidae family. The native area of the BSF is considered to be Central America and the northern part of South America, then spread to the entire continent over the last thousands of years (Hauser and Woodley, 2015; Kaya et al., 2021). These geographical areas are now considered their natural area (Hauser and Woodley, 2015; Rozkosný, 1983). The BSF is now a cosmopolitan fly found in tropical, subtropical, and tempered regions, from the 40th parallel north to the 45th parallel south 2, 3, where it is considered an exotic, non-invasive species, according to the current regulations (CBD, 2010; European Commission, 2021; Wang and Shelomi, 2017). This fly likely

spread through shipping routes (Picker et al., 2002). While studies based on naturalistic observations show that it was first discovered in Malta in 1926 (Lindner, 1936), it has been suggested that the earliest record in Europe was in Italy about 500 years ago (Benelli et al., 2014). The BSF appeared in the early 1950s in the South-East of France (Barbier, 1952) before spreading through the South-West (Dauphin, 2017) along the Rhône valley (Richoux, 2009) and finally along the Atlantic coast (Maquart et al., 2020). Currently, the BSF can be found on the whole European territory, with few exceptions. It is reasonable to assume that the rise of international traffic and the industrial use of the BSF may have led to a steady flow of introductions in France and worldwide (Craig Sheppard et al., 1994; Leclercq, 1969).

The BSF was initially used in forensic entomology to date the post-mortem time (Lord et al., 1994). Industrial companies heavily use it to reduce biowastes (Gold et al., 2018; Hoc et al., 2020), to produce protein (Ewald et al., 2020; Giannetto et al., 2019; Hoc et al., 2020), and fat (Ewald et al., 2020; Hoc et al., 2020; Spranghers et al., 2017). These products are used to make animal feed (Huis, 2013) or biodiesel (Leong et al., 2016; Surendra et al., 2020).

BSF industries face similar issues as other sectors: knowing how to manage their species and genetic aspects to get the best out of it. To this end, it is crucial to know the genetic diversity of BSF existing internationally, nationally, and locally. Many articles exist about the nutritional aspect of



the BSF (Biasato et al., 2019; Heuel et al., 2021; Neumann et al., 2018; Schiavone et al., 2019; Spranghers et al., 2017) or its microbiota (Erickson et al., 2004; Khamis et al., 2020; Liu et al., 2008), but few have focused on the genetic diversity of the BSF. We have some evidence of BSF genetic diversity from previous studies. Ten haplotypes based on the CO1 gene were found in South Korea with 245 individuals (Park Soyeon et al., 2017). Still, on the CO1 gene and a worldwide sampling, 56 haplotypes were found with a divergence rate up to 4.9%, and it has been shown that individuals belonging to divergent lineages are still interfertile (Ståhls et al., 2020). Finally, based on 15 microsatellite markers, 16 genetic clusters were found with hot spots in South America (Kaya et al., 2021). Other works focused on more fundamental aspects of BSF and assembled a reference mitochondrial genome and a chromosomal-scale nuclear genome (Generalovic et al., 2020; Qi et al., 2017; Zhan et al., 2020).

Within the framework of the BSF industry, phylogeography can allow on the one side to account for genetic diversity at different scales (global, national, and local) and, on the other side, to have keys to the geographical origin of the species present in a territory or used by companies. This study may suggest local adaptations or some phenotypic plasticity due to habituation to different biomes.

Our study quantified genetic diversity at different scales: at the global level by studying the variation within the CO1 gene and more locally at the French level. Our report is based on data obtained by sequencing 114 new

individuals for the CO1 gene focusing specifically on France (90 French individuals were analyzed) and extracting 563 CO1 sequences from databases. We confirm the hidden genetic diversity within *Hermetia illucens* at all scales of observation, allowing us to find nine different haplotypes. We sequenced and assembled 56 mitochondrial genomes from four different continents and 40 different locations. In addition, we collected sequencing data from databases allowing us to assemble three additional mitochondrial genomes. We also collected a mitochondrial genome already assembled from databases, resulting in 60 mitochondrial genomes. These data allowed us to resolve the phylogenetic relationships between the major haplotypes robustly. In addition, we also provide here the date of the separation events of the different haplotypes found. This separation turns out to be more than two million years old for the most distant haplotypes.

## Materials and methods

### Sampling

We decided to conduct a global sampling followed by a specific focus on France. A detailed table of samples and associated metadata is provided in (Table S01 – supplementary p222). The origin of all CO1 sequences used in our study is summarised in Table 1, and a summary of sample numbers by continent is provided in Table 2.

ORIGIN	NUMBER OF INDIVIDUALS (CO1)	NUMBER OF INDIVIDUALS (MITOCHONDRIA)
COLLECTION FROM THE WILD	91	43
COMMERCIAL ORIGIN	23	13
MINED FROM DATABASES (GENBANK & BOLD)	560	1
ASSEMBLED FROM SRA (RAW SEQUENCES)	3	3
TOTAL	677	60

Table 1: Origin of CO1 sequences and mitochondrial genomes

GEOGRAPHIC ORIGIN	NUMBER OF INDIVIDUALS – CO1	NUMBER OF INDIVIDUALS – MITOCHONDRIA
SOUTH AMERICA	55	4
NORTH AMERICA	43	2
ASIA	277	4
OCEANIA	106	0
AFRICA	78	11
EUROPE (INCLUDING FRANCE)	118 (90)	39 (38)
TOTAL	677	60

Table 2: Geographic origin of co1 sequences and mitochondrial genomes

Regarding the samples, one adult individual was found in Mexico in the wild in 2016. Individuals coming from Asia (Taiwan), Africa (Ghana and Kenya), and South America (French Guiana) were collected during the spring and summer of 2019. They come either from the wild or breeding farms (See supplementary S01 p222). In France, the collection of BSF was done with the help of the citizen compost network (RCC, n.d.) during the summer of 2020. We recovered mainly L5 larvae exclusively from shared and personal composts and some nearby adults. Between one and ten individuals were collected per area. In some French localities, collections in the wild have been made near BSF breeding areas (about five kilometers). We obtained

larvae in the areas of Blois, Bordeaux, Toulouse, Montpellier, Lyon, Paris. All these samples were stored in 100% ethanol at -20°C prior to DNA extraction.

In the NCBI Sequence Read Archive (SRA), we gathered all the CO1 sequences on NCBI and Bold and whole-genome sequencing data and reconstructed the mitochondrial genome. We were surprised in the Bayonne area to have BSF-like larvae, which turned out to be *Exaireta spinigera* larvae. These larvae were analyzed in the same way as the BSF.

## DNA extraction

We extracted DNA from larvae, pupae, and imago under the same conditions by grinding individuals in liquid nitrogen using a mortar and pestle. The ground samples were then processed according to the protocol of the Macherey-Nagel AXG 100 kit for the isolation of genomic DNA from tissue with some modification: we added a second purification step with 3.5 ml of freshly prepared ethanol and centrifugation at 15,000 rpm at 4 ° C for 15 min just before the final re-suspension. The DNA pellet was finally re-suspended in 100 µL of Tris 10 mM EDTA 1 mM pH 8.0. The purity was evaluated by spectrophotometry using a Nanodrop (Desjardins and Conklin, 2010), and the concentration was quantified using the Qubit kit DNA Broad range. DNA integrity was monitored using routine agarose gel electrophoresis.

## CO1 Sequencing

We designed specific CO1 primers from the reference mitochondrial DNA (Qi et al., 2017 - Table S02 – supplementary p227), then a mix was realized using

NEB 5X standard buffer One Taq. The mix is then amplified in ONETAQ HOT 45 PCR. We purified the PCR products using the Illustra Exostar 1-step GE Healthcare kit before Sanger cycle sequencing. Cycle sequencing reactions were performed using the BigDye 3.1 chemistry, purified by precipitation, and sequenced using an ABI 3130 sequencer (Applied Biosystems). Base-calling was performed using the sequencing analysis software version 5.3 (Applied Biosystems). We assembled forward and reverse reads using the Geneious V R11 (Geneious, n.d.) assembler with the sensitivity set at medium/fast. The regions containing the primers were excised, and we kept only the sequences having more than 75% of high-quality base calls.

## **Multiple sequence alignment and haplotype network reconstruction**

We aligned CO1 sequences obtained in the laboratory and those recovered from the databases using the MAFFT (Kato and Standley, 2013) aligner included as a plug-in in the Geneious software (Geneious, n.d.). We used the algorithm G-INS-i with a scoring matrix of 200 PAM / k=2, a gap open penalty at 1.53, and an offset value at 0.123. We exported the multiple sequence alignment in nexus format (Maddison et al., 1997) for haplotype network reconstruction and visualization using POPART (Leigh and Bryant, 2015). We modified the output file to add geographical origins to the sequences.

## Shotgun sequencing and mitochondrial DNA assembly

A total of 56 BSF individuals were used for whole-genome shotgun sequencing using an Illumina method by Novogene UK (Cock et al., 2010; Erlich et al., 2008; Hansen et al., 2010; Jiang et al., 2011; Yan et al., 2013). The technique used is pair-end with a medium coverage rate of 4.68X (SD 0,50). The resulting sequences were checked for quality and counted for read parity. We take advantage of the fact that mtDNA is highly repeated in insect DNA. In insects, mtDNA represents on average 0.42% of the total DNA in the genome sequence project (Meng et al., 2019). As stipulated, we used 5000 kb of data/sample (4-5x coverage) for the mitochondrial assembly, which on average contains  $(5,000 \times 0.42) / 100 = 21,000$  kb of mtDNA. As the mitochondrial genome is about 15 kb, we obtain coverage of 1400x for the mitochondrial genome, which is mainly sufficient to obtain a complete genome in one contig. The sequences were assembled de novo with the MEGAHIT software (Li et al., 2015) (based options); the mitochondrial genomes were thus recovered from the assemblies by blast (Altschul et al., 1990) using the reference BSF mitochondrial genome sequence (NC\_035232). The mitochondrial DNAs were annotated using MitoS online software (Bernt et al., 2013), and the D-Loop zone was removed. We added four complete mitochondrial sequences from databases or reconstructed them via reading sequences available on the SRA (Leinonen et al., 2011) using the same assembly procedure and then aligned them under the same conditions.

## CO1 and whole mitochondrial genome phylogeny

We use the MEGAX (Kumar et al., 2018) software to determine the best model and compute an ML Tree with 500 Bootstraps using a complete deletion parameter. We visualized the data with the online tool ITOL (Letunic and Bork, 2021), and we rooted the tree with the Stratiomyidae *Exaireta spinigera*. We chose *Exaireta spinigera* because it was the closest species for which we have a complete mitochondrial genome (the blastn result of the complete mitochondrial genome of *Hermetia illucens* vs. *Exaireta spinigera* gives us 81.28% of similarity). To simplify the interpretation of the phylogeny, we have kept only a single identical sequence per haplotype.

## Time tree

We used the BEAST software (Suchard et al., 2018) to estimate the divergence time of the different BSF groups. The analysis was performed with an expansion growth model and an uncorrelated lognormal relaxed clock with the proposed insect molecular clock. Ten million iterations were performed, and then with the TreeAnnotator (Abrami et al., 2019), the obtained trees were clustered with a burn-in of 25% of the trees. The visualization of the final tree was done with the online software ITOL (Letunic and Bork, 2021).

# Results

## Haplotype Network

We reconstructed a haplotype network from the analysis of 677 sequences spanning 658 nucleotides of the CO1 mitochondrial gene. Using the minimum spanning method (Paradis, 2018), 52 distinct haplotypes were delineated from the analysis of this CO1 fragment. These different haplotypes group up to 230 sequences (Figure 1).

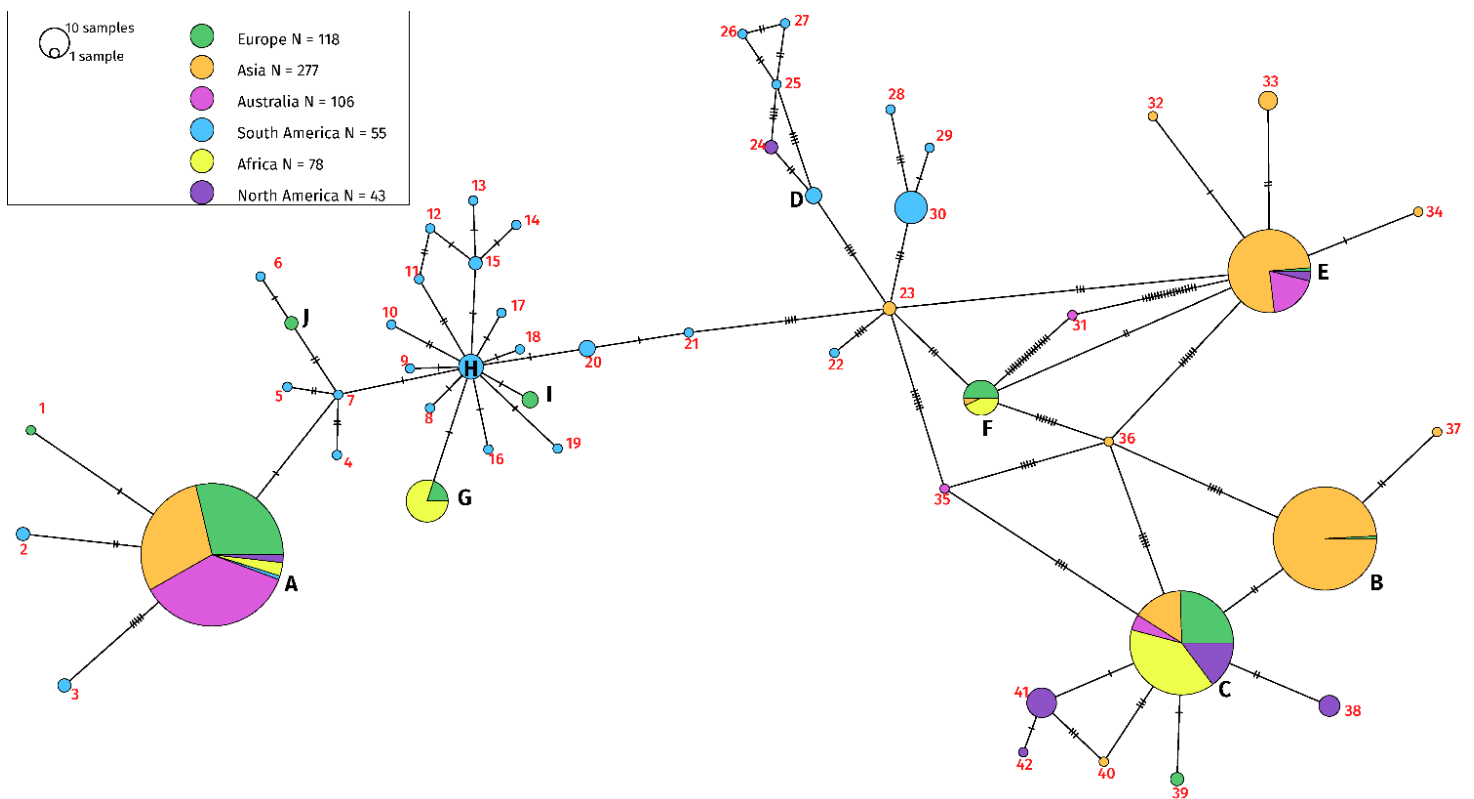


Figure 1: CO1 haplotype network of *Hermetia illucens*

THE HAPLOTYPE NETWORK GENERATED BY POPART AFTER MULTIPLE ALIGNMENTS ARE BASED ON THE CO1 GENE OF HERMETIA ILLUCENS. ANALYSIS OF THE DIFFERENCES BETWEEN THE CO1 GENE SEQUENCES OF 677 INDIVIDUALS (WE DID NOT SEPARATE COMMERCIAL AND WILD INDIVIDUALS). THE PIE CHARTS REPRESENT THE SAME HAPLOTYPE, AND THEIR RELATIVE SIZE INDICATES THE NUMBER OF INDIVIDUALS SHARING THIS HAPLOTYPE. THE NUMBER OF BARS BETWEEN EACH SEGMENT INDICATES THE DIFFERENCES BETWEEN THE TWO SEQUENCES. WITHIN A PIE CHART, THE COLOURS REPRESENT THE GEOGRAPHICAL ORIGIN OF THE INDIVIDUALS. THE MAJOR HAPLOTYPES INCLUDE MORE THAN FIVE SEQUENCES (A, B, C, E, F, G, H, 30, 38 & 41). WE GROUPED INDIVIDUALS BY CONTINENT, SEPARATING NORTH AMERICA FROM CENTRAL AND SOUTH AMERICA. THE LIST OF SEQUENCES AND THEIR EXACT ORIGIN IS AVAILABLE AS AN ADDITIONAL DOCUMENT (S01).



There are 31 haplotypes containing only one sequence and 11 haplotypes with between 2 and 4 sequences from the same continent. We have named these 42 haplotypes “minor haplotypes.” We observed ten haplotypes between five and 230 sequences, including six multi-continental haplotypes that we will refer to as major haplotypes. We found an important diversity both within and between continents using the CO1 phylogenetic mitochondrial gene. Minor haplotypes represent 8.5% of all sequences (58 sequences) and 80.8% of haplotypes. Major haplotypes represent 91.5% of the sequences (619 sequences) and 19.2% of the haplotypes.

In Europe, we found ten distinct haplotypes from 118 sequences, nine for Asia from 277 sequences, five for Oceania from 106 sequences, 30 for South America from 55 sequences, four in Africa from 78 sequences, and seven in North America from 43 sequences. Some major haplotypes are represented in a more restricted area. For example, haplotype B is mainly in Asia (figure 1). We observed a high haplotype diversity in South America with 28 minor haplotypes that contain between one and four sequences. The haplotype C (we choose C for “commercial”) contains all the sequences derived from farmed strains of the industry working on BSF.

### **Global distribution analysis**

We visualized the actual distribution of the different haplotypes found previously on a distribution map (figure 2).

Each haplotype from A to J is now represented with a color (figure 2); haplotypes numbered in the haplotype network (figure 1) are now represented in grey.

We observed great variability in South America, highlighted on the map by the grey pie charts representing mostly minor haplotypes. We observe 22 singletons in South America out of the 52 haplotypes found, more than 40% of the total diversity. On the other hand, we found six out of ten major haplotypes in Europe.

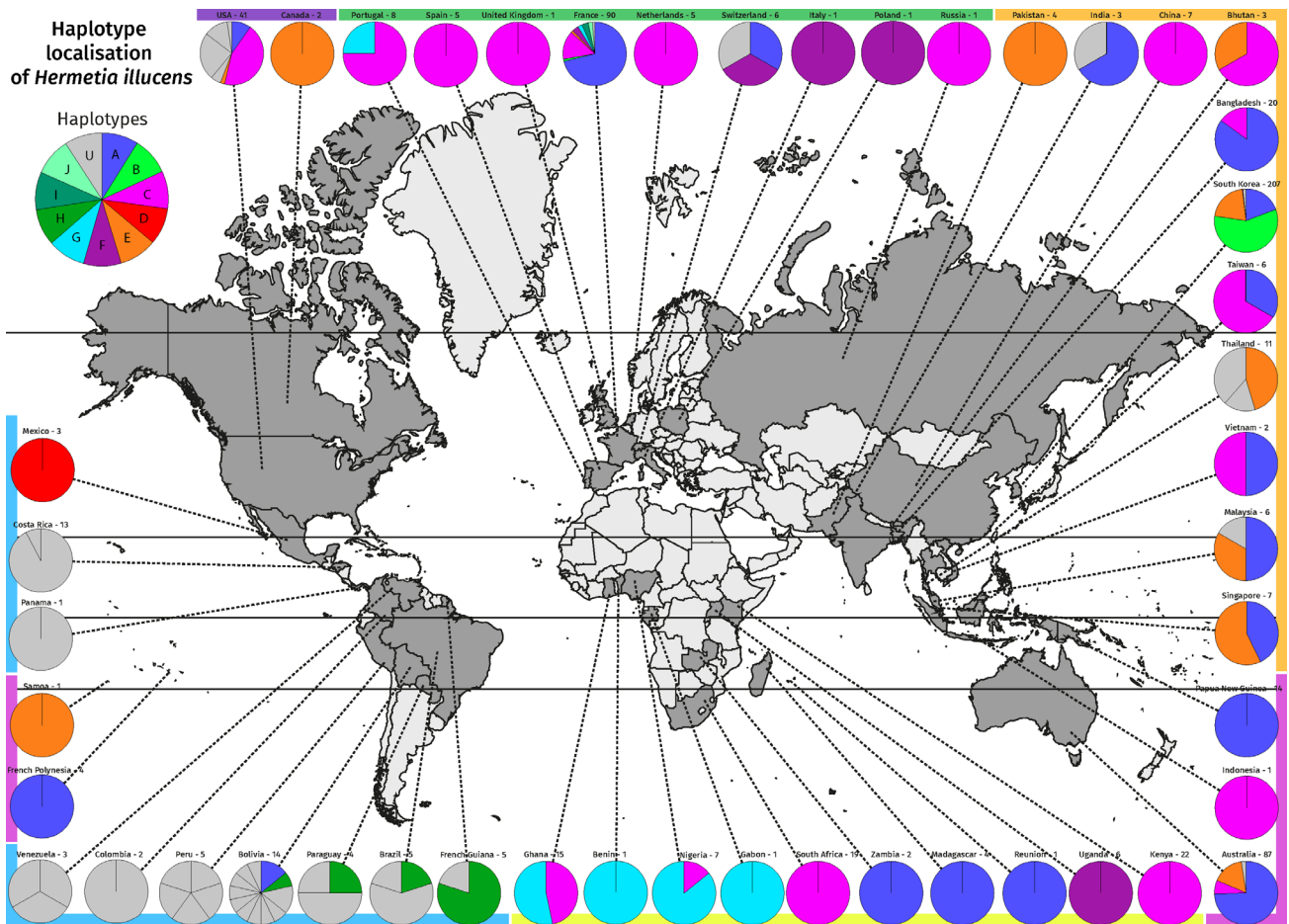


Figure 2 : World CO1 haplotype distribution

GLOBAL DISTRIBUTION OF THE DIFFERENT *HERMETIA ILLUCENS* HAPLOTYPES BASED ON THE CO1 GENE. THE PIE CHARTS REPRESENT THE PROPORTION OF EACH HAPLOTYPE WITHIN A COUNTRY. THE COUNTRY'S NAME IS FOLLOWED BY THE NUMBER OF INDIVIDUALS ANALYSED. A COLOURED BAND REPRESENTING THE CONTINENTS IS NEXT TO EACH PIE CHART. THE COLOURS

REPRESENT THE HAPLOTYPES DETERMINED IN FIGURE 1. THE HAPLOTYPES REPRESENTED WITH GREY COLOURS, INDICATED WITH THE LETTER U, ARE REPRESENTED WITH A NUMBER (1-42) IN FIGURE 1. SEPARATION IS MADE IN THE PIE CHART WHEN THE HAPLOTYPES CHANGE FOR THE SAME COUNTRY. FOR EXAMPLE, VENEZUELA CONTAINS THREE UNIQUE HAPLOTYPES, MEANING THAT THESE THREE INDIVIDUALS ARE DIFFERENT. A SUMMARY TABLE CAN BE FOUND IN SUPPLEMENTARY TABLE 1 (S01 p222).

Haplotype A and C were the most represented haplotypes in our sampling: Haplotype A was present on all continents and represented 34% of all sequences, and haplotype C was only absent from South America and represented 18% of all sequences. Some haplotypes seemed to be restricted to some world areas, such as haplotype G, which was present in all the localities in West Africa and surprisingly in a Portuguese sample, or the haplotype H in South America. Haplotype F was present in Europe, Uganda, and South Korea. Haplotype E was more present in Asia, Oceania, and North America. Finally, some haplotypes such as B, I, and J present rare sequences found in more restricted areas such as B in Korea or H and J in France (figure 2).

## CO1 Phylogeny

During our CO1 gene analyses, some BSF-like larvae were identified as *Exaireta spinigera* larvae. The sequences were amplified and sequenced with the same primers as for BSF.

By taking only one sequence for each haplotype, we have reconstructed a phylogeny of the CO1 gene sequence. The CO1 phylogeny was rooted with the *Exaireta spinigera* (figure 3). A sample from Cameroon was revealed to belong to a closely unidentified species of BSF and was termed *Hermetia* sp. Cameroon in our subsequent analysis (Blastn score between *Hermetia*

sp. from Cameroon and BSF gives us 85.01% identity with an E Value of 2e-174).

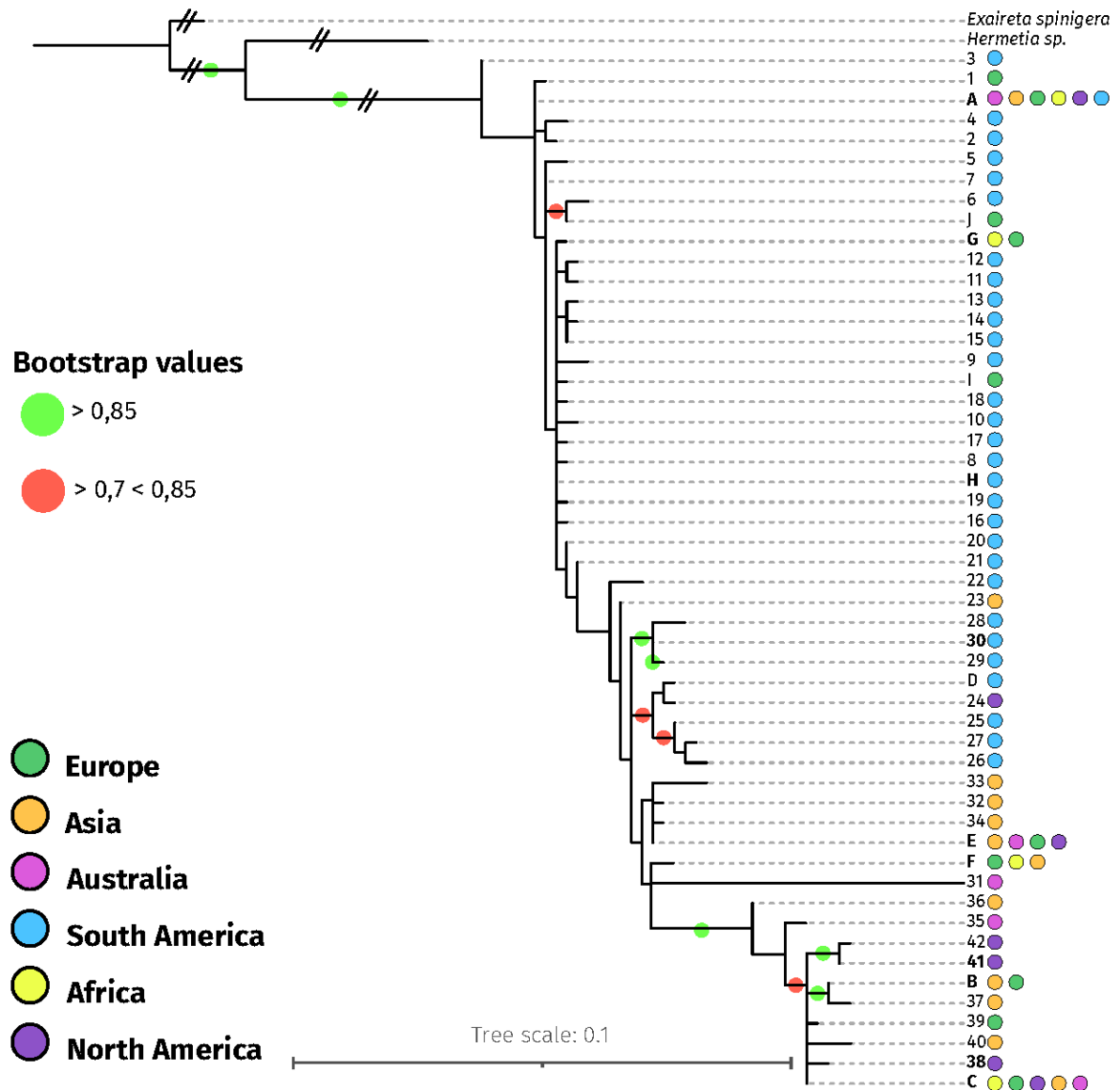


Figure 3 : CO1-based phylogenetic tree

ML PHYLOGENETIC TREE. ROOTED PHYLOGENY WITH EXAIRETA SPINIGERA (STRATIOMYIDAE). BOOTSTRAP VALUES ARE INDICATED WITH COLOURED CIRCLES: GREEN FOR VALUES > 0.85 AND RED FOR VALUES BETWEEN 0.7 AND 0.85. WE ONLY KEPT ONE SEQUENCE PER HAPLOTYP TO VISUALISE THE RELATIONSHIP BETWEEN THEM. NEXT TO EACH HAPLOTYP, COLOURED CIRCLES REPRESENT THE GEOGRAPHICAL ORIGIN OF THE SPECIES PRESENT IN THIS HAPLOTYP. THE ORDER INDICATES THE RELATIVE ABUNDANCE OF THE INDIVIDUALS IN EACH HAPLOTYP. THE STATISTICAL SUPPORT WAS ESTIMATED WITH 500 BOOTSTRAPS WITH MEGA X SOFTWARE USING THE TAMURA 92 G+I MODEL.

The CO1 gene phylogeny supports the analyses performed with the haplotype network (Figure 1), showing remarkable sequence diversity among the CO1 gene. Interestingly, all the commercial strains are found in the C haplotype (a detailed phylogenetic tree is given as a supplementary with an asterisk placed on individuals from farms (Figure S03 – Supplementary p228).

A group containing haplotypes 35 to 42 and haplotypes B and C forms a divergent group from the others, with strong statistical support (bootstrap value of 0.99). Some individuals captured in the wild have been found near BSF commercial breeding sites. For example, we have captured C haplotype individuals in Avignon close by companies working with the BSF.

Although the global resolution of the deep branches of the tree is low, it is also possible to better understand the relationship between the haplotypes: haplotypes B and C are close to each other and well separated from other haplotypes. The other haplotypes are not very well resolved with the CO1 sequences. Finally, the samples coming from Latin America are scattered all over the tree, showing their great diversity and supporting the hypothesis of the geographical origin of the BSF.

### **France displays vast BSF genetic diversity at local scales**

Among the ten major haplotypes, our sampling effort in France showed that five are present in France, and a unique haplotype is only found in France (Figure 4) in La Rochelle. Haplotype A remains the dominant haplotype in

France (71.1% of all sequences) in seven of the eight localities (excluding Beaufort en Vallée, where the flies come from a company). Regarding the haplotype C, BSF belonging to this group have been found in four locations in the regions of Poitiers, Bordeaux, Beaufort-En-Vallée and Montpellier. Except for Beaufort-En-Vallée, where the flies come from a company, the other BSF come from natural environments. BSF industries exist in all these localities. More surprisingly, different haplotypes coexist in some localities. For example, we found three different haplotypes in the region of Paris and four in the region of Montpellier. Finally, we found a singular haplotype in Colomiers (J) and La Rochelle (1).

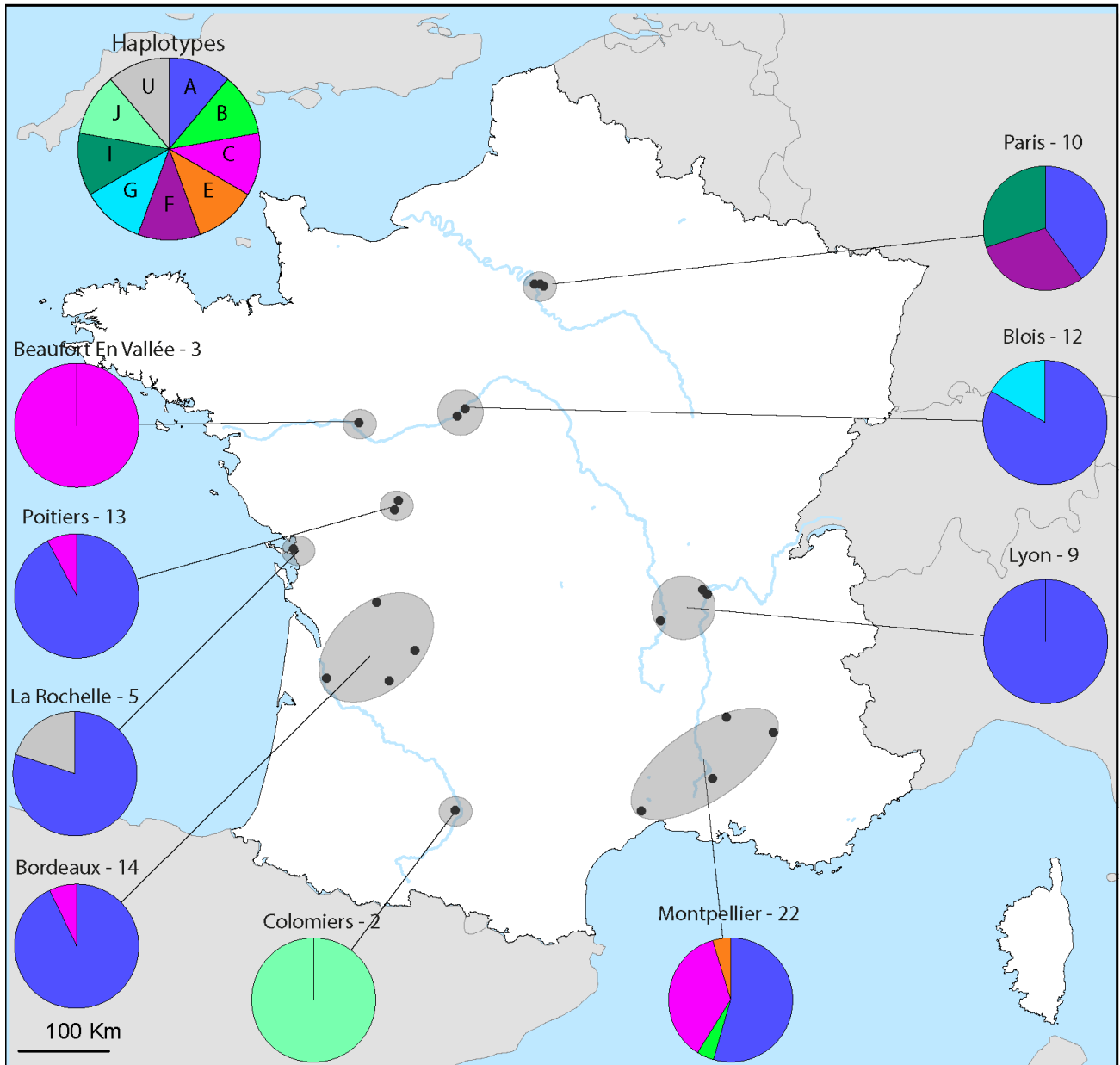


Figure 4 : France CO1 haplotype distribution

DISTRIBUTION OF THE DIFFERENT HAPLOTYPES OF *HERMETIA ILLUCENS* IN FRANCE. THE DOTS REPRESENT THE SPECIFIC AREA OF COLLECTION. THE PIE CHARTS REPRESENT THE DISTRIBUTION OF HAPLOTYPES IN THE CORRESPONDING AREA. THE NAMES OF THE MAIN CITIES IN EACH AREA ARE INDICATED ABOVE, AND THE NUMBER OF SEQUENCED INDIVIDUALS.

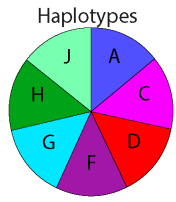
PARIS INCLUDES VINCENNES AND NOGENT SUR MARNES. BLOIS INCLUDES VALLOIRE SUR CISSE. LYON INCLUDES SAINT MARCELLIN EN FOREZ, CORBAS, CHAMPAGNE AU MONT D'OR AND DARDILLY. MONTPELLIER INCLUDES AVIGNON, TRESCLEOUX AND DIEULEFIT. BORDEAUX INCLUDES LORMONT LA FORCE AND PERIGUEUX. POITIERS INCLUDES SAINT GEORGES LES BAILLARGEUX.

## Complete mitochondrial genome analyses

To better understand BSF phylogeny and estimate the timing of haplotype divergence, we analyzed the mitochondrial genomes of 60 individuals (56 that have been sequenced and assembled, three from which sequences have been recovered and assembled and one recovered), including 3 *Exaireta spinigera* genomes (used as an outgroup). The 60 reconstructed mitochondrial genomes have the typical structure of insect mtDNA: 13 protein-coding genes, 22 tRNA, two rRNA-coding genes, and non-coding regions and control regions (an example is given as a supplementary (Figure S04 – Supplementary p229). The genetic differences are homogeneously distributed overall mtDNA and are not localized on specific genes (Figure 5A). A great genetic diversity at the mitochondrial genome level is observed. Two very homogeneous groups, A and C, are well represented, and sequences of the C haplotype display many mutations all along the genomes from the others. The two sequences with the most differences are 55-F-Angouleme and 21-Ghana, with a similarity percentage of 96.373% (all results are in supplementary S05 (p230). Shorter alignment lengths result from comparisons with the outgroup sequences that split the alignments into two fragments.) The mitochondrial genomes of the three *Exaireta spinigera* were extremely close to each other (100% by BLAST - supplemental S05 p230), so we kept only one to serve as an outgroup.



We also wanted to know if the mutations are equally (or not) distributed between the different genes. Therefore, we compared the similarity of the 13 mitochondrial genes of the 57 BSFs (fig 5B) to those of *Exaireta spinigera*. The results are presented as percent sequence similarity.



**A**

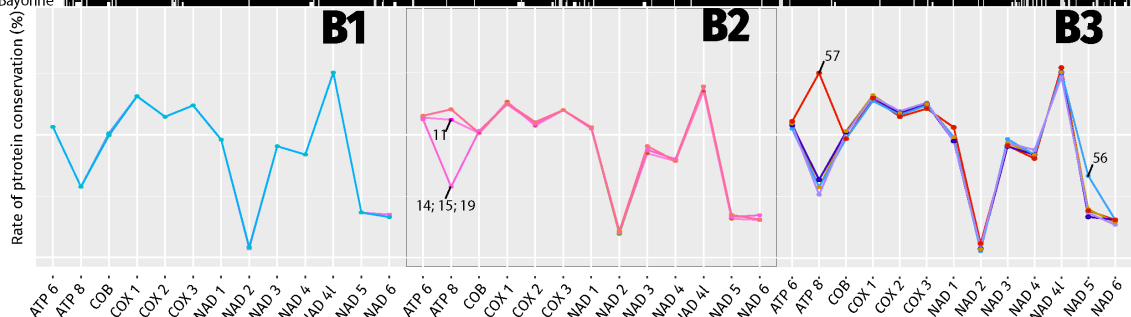
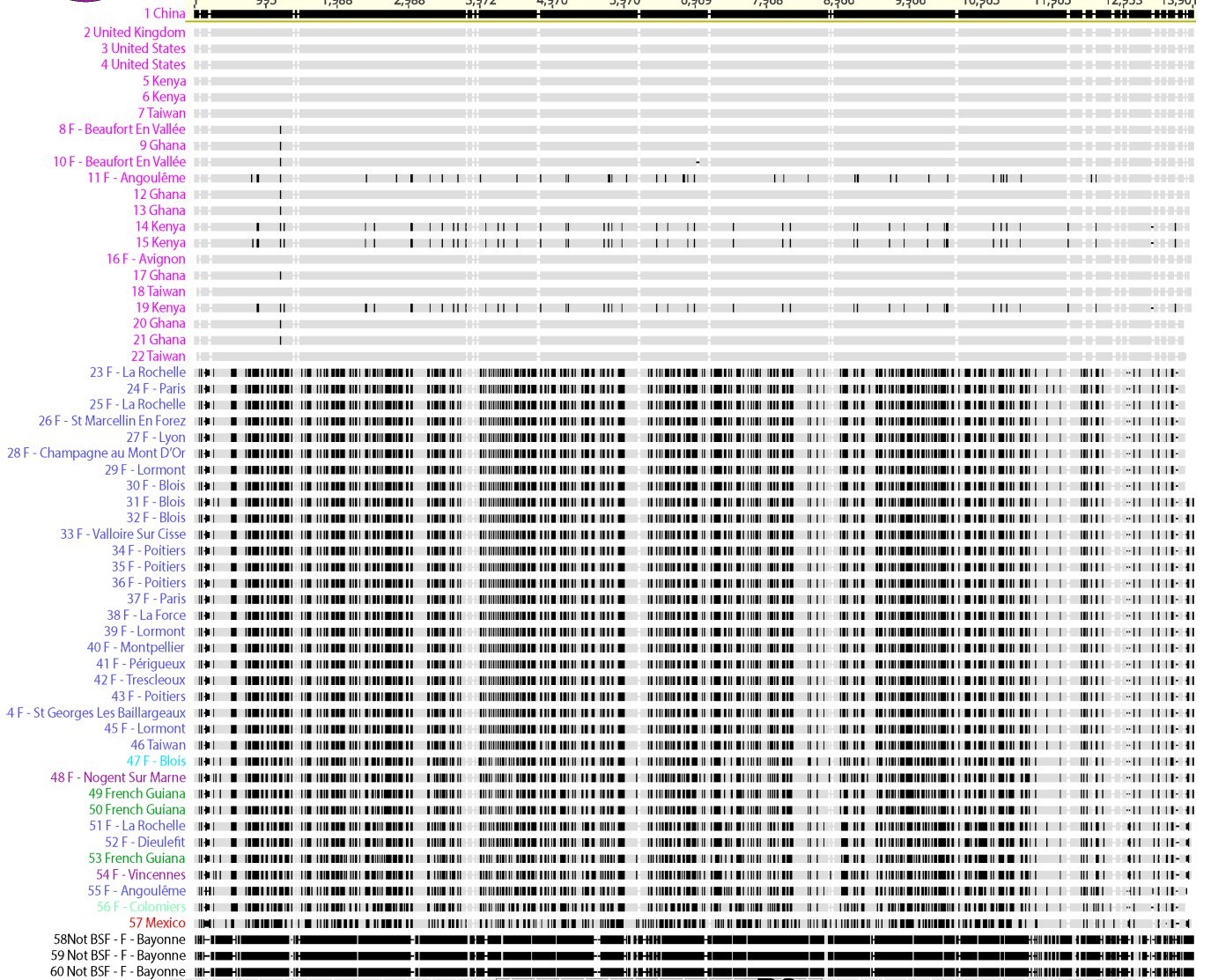
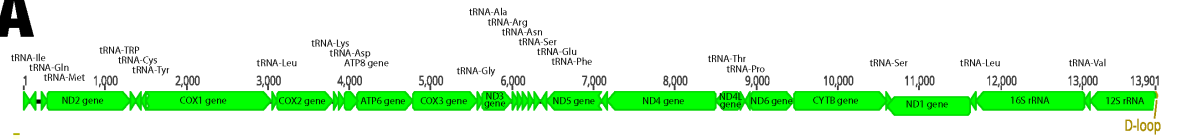


Figure 5 : Mitochondrial DNA alignment and sequence conservation between *Hermetia illucens* and *Exaireta spinigera*.

PANEL A: MULTIPLE ALIGNMENTS OF MITOCHONDRIAL DNA SEQUENCES FROM *HERMETIA ILLUCENS*. ALIGNMENT WAS PERFORMED WITH THE MAFFT ALIGNER. THE COLOURS OF THE NAMES COME FROM THEIR HAPLOTYPES DETERMINED WITH THE CO1 SEQUENCES. EACH BLACK MARK REPRESENTS A DIFFERENCE WITH THE REFERENCE (1 – CHINA). ABOVE THE ALIGNMENT, THERE IS A LINEAR REPRESENTATION OF THE MITOCHONDRIAL GENES (A CIRCULAR ONE CAN BE SEEN IN SUPPLEMENTARY S04 P229). THE THREE LAST SEQUENCES ARE THOSE OF *EXAIRETA SPINIGERA* ANALYSED IN THE SAME WAY AS *HERMETIA ILLUCENS* (WE KEPT ONE THAT WE USED AS AN OUTGROUP BECAUSE IT WAS THE CLOSEST COMPLETE MITOCHONDRIAL SEQUENCE PHYLOGENETICALLY THAT WE HAD). THE SEQUENCES ARE NUMBERED TO BE FOUND IN FIGURE 6.

PANEL B: SEQUENCE CONSERVATION OF MITOCHONDRIAL GENES OF *HERMETIA ILLUCENS* TO *EXAIRETA SPINIGERA*. THE SEQUENCES ARE GROUPED ACCORDING TO HAPLOTYPES. THE A HAPLOTYPE (B1), THE C HAPLOTYPE (B2), AND THE OTHER HAPLOTYPES (B3). THE SEQUENCES INSIDE EACH GRAPH ARE REPRESENTED IN THE ALIGNMENT (5A). SEQUENCES FROM 23 TO 46 PLUS SEQUENCES 51, 52, AND 55 FOR HAPLOTYPE A (5-B1); FROM 1 TO 22 FOR HAPLOTYPE C (5-B2); AND FROM 47 TO 57 (EXCEPT SEQUENCES 51, 52, AND 55) FOR THE OTHER HAPLOTYPES (5-B3).

We grouped the individuals into three different categories:

Haplotype A (Figure 5 B1) represents the percentages of differences of the 27 individuals of haplotype A compared with *Exaireta spinigera*. All individuals show the same level of gene difference as those of *Exaireta spinigera*. Only the NAD6 gene presents slight differences.

Haplotype C (Figure 5 B2) represents the percentages of differences of the 22 individuals of haplotype C compared to *Exaireta spinigera*. It shows that most of the mitochondrial genes are identical except for the ATP8 gene, which is different for four individuals (11 Angouleme; 14 Kenya; 19 Kenya, and 15 Kenya). Slight differences are also present in the NAD3 and NAD6 genes.

Haplotype D-F-G-H-I (Figure 5 B3) represents the percentages of differences of the eight other individuals of haplotype C compared with *Exaireta spinigera*. The profiles are slightly different, with a more substantial disparity for the ATP8, NAD1, and NAD5 genes, especially the Mexican one

(57). The profiles show the same pattern except for the ATP8 gene of the individual from Mexico (57), which is more conserved.

### **Complete mitochondrial genome phylogeny**

The complete mitochondrial genome phylogeny demonstrates the very high genetic divergence among BSF populations found with the CO1 gene analysis (Figure S07). The increased sensitivity of the analysis, due to the size of the sequences (13.96 kb, 13 genes), robustly resolves the relationships between different haplotypes identified previously (nodes >95%). In addition, it allows us to distinguish subgroups within the previously determined haplotypes. Two major divisions emerge within the tree: group C (which includes all commercial strains) and others. Two subgroups emerge within the C group (11 – F – Angouleme; 14 Kenya; 19 Kenya; 15 Kenya and the others).

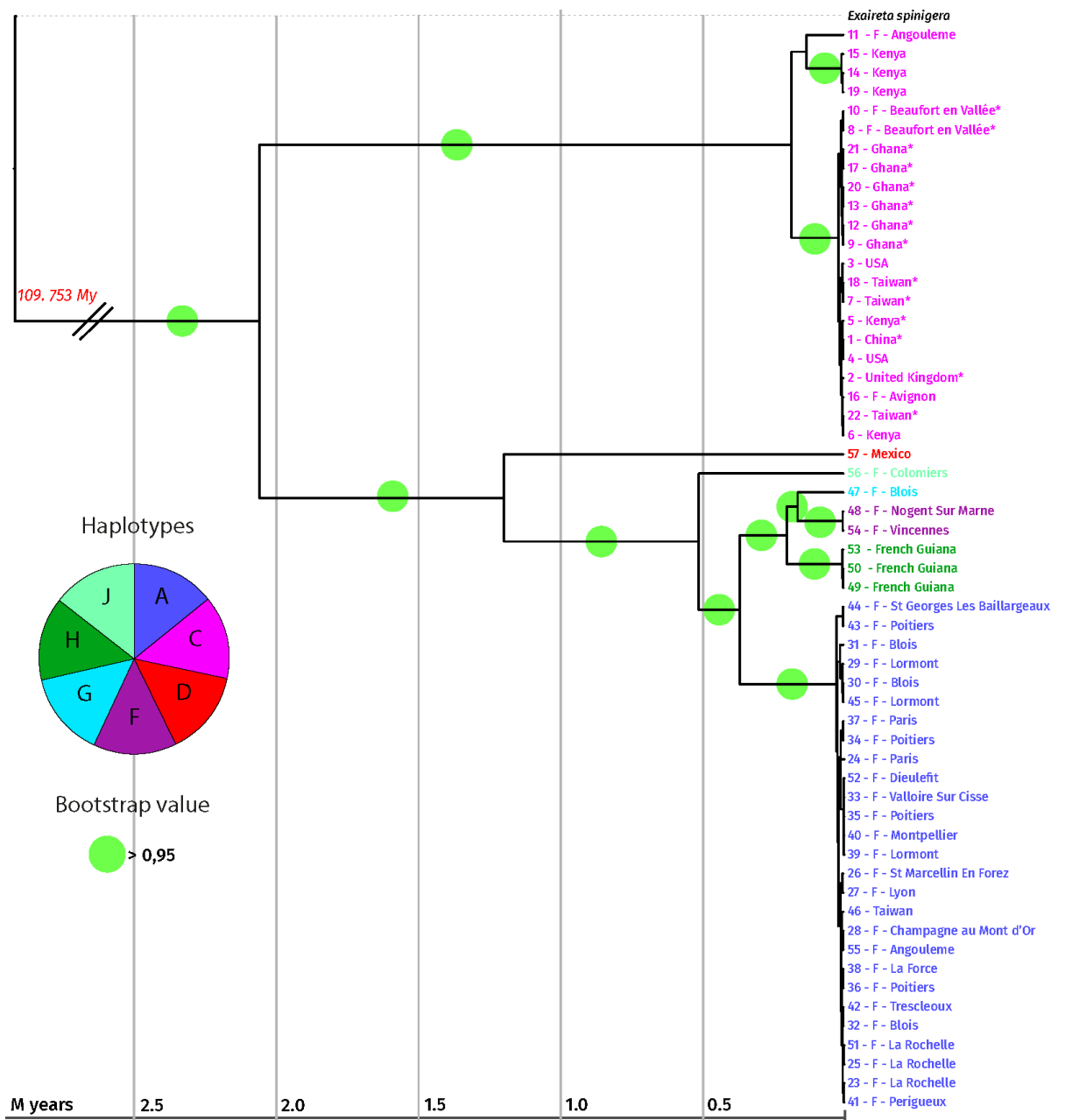


Figure 6 : Mitochondrial DNA phylogenetic and time tree

BAYESIAN TIME TREE OBTAINED WITH BEAST OF 57 MITOCHONDRIAL GENOMES OF HERMETIA ILLUCENS DEPRIVED OF D-LOOP REGIONS, ALIGNED WITH MAFFT AND ROOTED WITH EXAIRETA SPINIGERA. THE SCALE IS IN MILLIONS OF YEARS. A BREAK IN SCALE IS MADE AT THE ROOT LEVEL. THE COLOURS OF THE NAMES CORRESPOND TO THE HAPLOTYPES DETERMINED WITH THE CO1 SEQUENCES (FIGURE 1). COMMERCIAL INDIVIDUALS ARE MARKED WITH AN ASTERISK.

THE GREEN CIRCLES COME FROM THE ML PHYLOGENETIC TREE OF THE 57 MITOCHONDRIAL DNA SEQUENCES OF HERMETIA ILLUCENS, DEPRIVED OF D-LOOP REGIONS. MAFFT WAS USED TO ALIGN THE SEQUENCES, 500 BOOTSTRAP REPLICAS WERE PERFORMED. THE OUTGROUP USED IS EXAIRETA SPINIGERA (STRAHIOMYIDAE). BOOTSTRAP VALUES ARE INDICATED BY A GREEN CIRCLE WHEN THEY ARE > 0.95 (THE SEPARATE TREES CAN BE SEEN IN THE SUPPLEMENTARY S07 P232).

The other large group contains sequences of A, D, I, G, and H haplotypes. The I, G, and H haplotypes are related and close to each other. The very cosmopolitan A haplotype is well conserved with limited mitochondrial sequence divergences.

Moreover, although phylogenetically close, the I, G, and H haplotypes are found in very disparate areas (the G haplotype is mainly in Africa, the I in Europe, and the H in Latin America), supporting multiple worldwide introductions.

Strikingly the high level of the conservation of the ATP8 sequence described above in the Mexican sample (haplotype D) does not explain the basal position of the sequence in the phylogeny; indeed, we tested the phylogeny by removing the ATP8 gene, and no change was found (Figure S06 – Supplementary p231).

The ATP8 gene alone is not sufficient to explain the appearance of the observed subgroups. However, unlike the CO1 gene, which varies only slightly within this haplotype, the NAD2, NAD3, NAD5, and NAD6 genes are more variable.

The average percentage of identity between haplotypes A and C is about 96.39%. Between haplotypes C and D, 96.54%. Between haplotype C and H, 96.40% and 96.43% between C and I. Between A and D, 97.82% (all results are in Figure S05 – Supplementary p230).

## The divergence time of the different groups

The time tree (figure 6) gives us information about the divergence time of the different mitochondrial genomes. The phylogenetic tree presented in figure 6 contains both the Bayesian and the ML trees. The two separate trees are presented in supplementary (S07 p232). We noted a relatively long divergence time between haplotype C and the others (2.2 million years). The others have diverged more recently (about 1.2 million years). Within the C haplotype, a divergence has been about 300,000 years, while in the other groups, the divergence times are much lower, about 20,000 years for the A haplotype.

The time of divergence with *Exaireta spinigera* has been estimated at 110 million years, confirmed by the literature that gives us 112 million years with the Timetree tool (Hedges et al., 2015).

## Discussion

*Hermetia illucens* is now considered a species of major interest for the feed industry (Tomberlin and van Huis, 2020). Thus, it became urgent to study the genetic diversity within the different populations at different scales. Analyzing the different haplotypes in the same geographical area is a powerful tool for studying the recent history of BSF. Hence, an in-depth study of the genetic variations between the different haplotypes may allow us to resolve better the scenario of the arrival of BSF in a territory.

Our study uncovered a high genetic variability of mitogenomes within the species *Hermetia illucens*. We found that haplotype C is the most divergent. The dating of the division between the C haplotype and the others goes back more than 2 million years. We found a well-resolved separation within haplotype C (>95%); the other haplotypes are also well resolved (>95%).

Our study confirms the genetic diversity found in previous papers (Kaya et al., 2021; Park Soyeon et al., 2017; Ståhls et al., 2020). We found a comparable number of haplotypes, although we added many more sequences of diverse origin but mainly from non-indigenous areas. For example, Ståhls et al. (Ståhls et al., 2020) found 56 distinct haplotypes with 418 CO1 sequences, and we found a comparable number of haplotypes based on the analysis of 677 CO1 sequences. These results suggest we have now captured a major part of the extant BSF worldwide diversity outside of the genetic hotspots identified previously (Kaya et al., 2021; Ståhls et al., 2020), namely Latin America. Indeed, in South America, we found 30 haplotypes among the 55 sequences sampled that open up the possibility of a still unknown genetic diversity in this part of the world.

With the maximum diversity found in Latin America, the hypothesis of the geographical origin of BSF<sup>1</sup> is supported by our study. Analysis of the complete mitochondrial genome supports the hypothesis of several distinct introductions on all non-native continents. Considering the estimated divergence times between the different haplotypes, it seems impossible



that this is due to a single introduction followed by diversification in the territory.

On the other hand, according to Kaya et al. (Kaya et al., 2021), an alternative but complementary scenario is also possible: bridgehead populations, consisting of a mixture of haplotypes, have been introduced on different continents. Indeed, taking Europe as an example, the eight haplotypes found in Europe can be explained by the admixture scenario proposed by Kaya et al. (Kaya et al., 2021), who argue that central European BSF populations originated from admixture between Australian and Southeast African populations, which themselves can be traced back to earlier independent admixture events. Furthermore, Kaya et al. (Kaya et al., 2021) suggested some gene flows with Mediterranean (altogether originating from Latin America) and particularly East African populations, which interestingly seems to be corroborated in our study, as exemplified by the F haplotype. It would be extremely useful in the future to combine nuclear data and mitochondrial data to better resolve the potentially diverging patterns of the different markers and to precisely estimate relative impacts of historic genetic admixtures *versus* recent introductions on regional population structures. France contains on its territory a great diversity of haplotypes, whose origin can be explained either by the historical introduction around the year 1950 (Barbier, 1952) in addition to more recent and repetitive introductions. The haplotype A, which is by far the most common haplotype in wild BSF found in France (71.1%) and in Europe, may correspond to the

initial introduction. On the other hand, our data suggest that the presence of haplotype C in France may be due to a more recent introduction caused by BSF industries.

As suggested in a recent article on the structure and demography of the BSF based on microsatellite data ([Kaya et al., 2021](#)), the captive populations used for breeding in Europe and North America derive from a common, genetically related strain. Our analysis based on CO1 and complete mitochondrial genome sequences confirm this result: the haplotype we named C contains all the sequences of industrial origin with very little genetic diversity. Interestingly, this commercial group does not have any South American relatives. Based on nuclear genetic data, Kaya et al. report wild North American populations being most closely related to the most prevalent managed strains, which we infer to be dominated by haplotype C. This result reinforces the hypothesis of a North American origin of the BSF commercial stock.

Moreover, specific individuals in the wild belonging to this group were collected in the vicinity (around 5 km) of industries breeding BSF. For example, if we focus on the commercial haplotype analysis in France, the wild BSF belonging to haplotype C may have escaped from industries in Poitiers, Bordeaux, Avignon, and Montpellier. This last point raises the question of the biosecurity of BSF farms worldwide. We can assume that the priority of companies involved in BSF farming is not biosecurity and escape control. Indeed, costs in terms of security to prevent insects from escaping

are high, and the regulation is late at this level. In that case, the continuous flow of escaping BSF of industrial origin in large numbers could either result in an establishment in territories where it was not yet present or trigger losses of genetic integrity of locally resident wild populations via continuous one-directional introgression. This would lead inexorably to a loss of the local genetic diversity that may prevent future adaptations to changing environmental conditions.

However, it is interesting to note that in Ståhls et al. 2020, in contrast to our study, some genetic diversity in the breeding strains in Europe and North America has been evidenced (10 haplotypes). This may be due to a broader sampling effort of commercial strains. Indeed, Ståhls et al. had the opportunity to sample many small farms leading to 292 BSF sequences from rearing cultures. This allowed them to access a broader genetic diversity not found in our study. It is also possible that smaller farms did not obtain BSF from the dealers who constituted most of the more prominent industries but directly from the wild. Nevertheless, our results confirm that the industrial actors work with the same and unique haplotype.

We also evidence that the mitochondrial genome is ideally suited to resolve long-term BSF phylogenetic relationships. It allows us to date the natural history of the BSF, bringing a sum of details that were impossible to distinguish based only on the CO1. For example, we report puzzling variations in the level of conservation of the ATP8 genes for some samples (haplotypes C and D). Interestingly, these variations have already been

found in other insect species (for example, in Coleoptera, see [Zhang et al. 2019](#), and in grasshoppers, see [Li et al. 2018](#)). The ATP8 gene is small (160 bp on average in insects) and encodes for a subunit of the complex V of the ATP synthase. In *Hermetia illucens*, the initiator codon of the ATP8 gene is ATT, is 167 bp long, and it overlaps with the ATP6 gene. It has been assumed in other species that the modification of the sequence of the ATP8 gene can affect the structure of mitochondrial complex V and thus its function.

Furthermore, it has been shown that changes within ATP synthase can be explained by adaptation to different ecological environments ([Gu et al., 2016](#); [Peng et al., 2012](#); [Zhang et al., 2019](#)). This indicates that some local adaptation driven by the ATP8 genes may exist in BSF, as it has already been shown in other insects. This can lead to an adaptation to survive in high-altitude environments, i.e., at a lower oxygen level and low temperature ([Luo et al., 2013](#)). Even if these variations did not change the phylogeny of the BSF, it could be a clue to some local adaptation signal. The fact that all commercial strains (plus USA plus Mexico) show a high level of conservation of the ATP8 gene compared to all other wild strains may suggest some adaptation to rearing.

Complete mitochondrial genome data allows us to make a robust and well-resolved time tree of the species. The divergence times between the different haplotypes are surprisingly long (more than two million years for the most divergent lineages), and the differences observed between the sequences prove that there is still unsuspected and hidden genomic

diversity. Comparing it to *Drosophila*, two million years correspond to many speciation events leading to different reproductively isolated species (Russo et al., 2013). By contrast, Stahls et al. indicated that BSF belonging to phylogenetically divergent haplotypes can still cross and generate offspring, indicating no apparent reproductive isolation in laboratory conditions. Interestingly, our data evidence that the cohabitation of different haplotypes at the same locality could lead to crosses between haplotypes. According to our results, the two million years of divergence between the C haplotypes and the other haplotypes could correspond to the North American lineages separating from the ancestral BSF population in South America upon their northwards range expansion beyond the Isthmus of Panama, which appears 2.8 million years ago (O’Dea et al., 2016). This seems especially true since the C haplotype has not been found in South America but only in North America. This ancestral and natural colonization of North America more than 2 million years ago may also explain why the C haplotype group displays a higher degree of mitochondrial genetic diversity than the other groups, a process already observed in other insects as ants for example (Winston et al., 2017). Further investigations are thus needed to understand the possible genetic barriers between highly divergent BSF groups but also the possible patterns of introgression in natural BSF populations. This phenomenon might considerably confuse the elucidation of the population structure and the natural history of these species.

Finally, it is interesting to recall that, if indeed, most industrial actors use the same haplotype for the industry, the other haplotypes present in France are highly divergent from this one. Strikingly, we found in one compost bin in the region of Paris flies belonging to two different haplotypes (A and F). Thus, crossbreeding between these two individuals belonging to these two haplotypes would bring a solid genetic mix that could benefit breeding. Associated with this significant genetic diversity, BSF may also have some significant phenotypic diversity that might be useful for specific industrial usage and selection processes to increase performances or ameliorate peculiar traits or behaviors (Zhou et al., 2013).

## Conclusion

Through the analysis of 60 mitochondrial genomes during this study, we proposed a robust and well-resolved time tree of the BSF that confirms the vast genetic diversity within *Hermetia illucens*. Our result elucidates the phylogenetic relationship and time divergence between the BSF lineages. By adding almost 20% of CO1 sequences, mainly from non-native areas, we did not significantly increase the number of haplotypes previously found (Ståhls et al., 2020). This indicates that a significant part of the diversity from non-native areas has already been found and that we should now concentrate our efforts on native areas. Furthermore, the combination of our results with the nuclear analysis already performed (Kaya et al., 2021) gives an interesting insight into the invasion routes of the BSF and the divergence times between the different groups. We have also shown that

many BSF industrial companies work with flies derived from a unique haplotype. We would also like to recall that the apparent biosecurity deficiency within some BSF industrial companies probably led to a continuous flow of escapement and possible local introduction/replacement of natural populations of BSF. Finally, our data indicated that it becomes fascinating to look at the nuclear genomes' evolution of the BSF to better understand the specific adaptations of the different lineages that might be useful for industrial needs and initiate the selection process for peculiar traits.

## **Declaration**

### **Ethics approval and consent to participate**

Not applicable

### **Consent for publication**

Not Applicable

### **Availability of data and materials**

The dataset generated during the study has been deposited in the European Nucleotide Archive database under Accession Numbers PRJEB48031 Bioproject. Alignment and raw data can be downloaded following this link

[https://figshare.com/articles/dataset/sequences\\_and\\_trees\\_zip/19589335](https://figshare.com/articles/dataset/sequences_and_trees_zip/19589335)

The supplementary can be downloaded following this link:

[https://figshare.com/articles/dataset/Supplementary\\_zip/19589332](https://figshare.com/articles/dataset/Supplementary_zip/19589332).

## **Competing interests**

The private company Cycle Farms partially granted this project. The company has not influenced the design and the analysis of the results and has not influenced the writing of the manuscript and its conclusions. No commercial products have been developed based on the outputs of this article. None of the authors has any competing interests.

## **Funding**

This work was funded in the framework of a partnership between the CNRS and the CycleFarms company. A CIFRE PhD supports JG. Grant from the Agence Nationale pour la Recherche et la Technologie (ANRT)

## **Author's Contributions**

JG carried out the sampling, performed data curation, and wrote the original draft. NP performed data curation and reviewed the manuscript. GB helped with the data curation and reviewed the manuscript. JF conceptualized and supervised the study, performed data curation, and reviewed and finalized the manuscript.

## **Acknowledgments**

We want to thank Floran Laville and Frederic Marion-Poll for their continuous support and help during this project. We warmly thank the members of the Réseau Compost Citoyen for their precious help during the sampling in France. We also thank Matan Shelomi (National Taiwan University), Meng-Kun Wu, Shu-Min Chen, Jing-Jiun Huang (Yi Mi Community College, Taiwan) for flies from Taiwan. We thank the International Centre of



Insect Physiology and Ecology (ICIPE) from Nairobi City, Kenya. We thank Mathieu Chouteau from CNRS LEESIA, in French Guiana. We finally thank Benoit Gilles for his help at the beginning of this project and the two anonymous reviewers for their constructive comments.

# Chapitre 3 : Une diversité nucléaire pleine de surprises

## Résumé

La diversité génétique d'une espèce peut s'exprimer de bien des manières et à bien des niveaux. Dans ce chapitre, je vous présente un aspect de cette diversité en utilisant les éléments transposables. Ces séquences nucléotidiques capables de se déplacer dans les génomes, sont responsables d'un accroissement de la diversité génétique. Nous avons voulu savoir si les contenus en éléments transposables étaient les mêmes pour toutes les populations et si l'on retrouvait des constantes entre les génomes d'*Hermetia illucens*.

En étudiant les éléments transposables de 59 individus, nous avons pu remarquer que les éléments transposables présents étaient majoritairement des LINE et plus spécifiquement quelques familles. Mais nous avons également pu montrer qu'il y avait des différences fortes entre les différents individus quant au contenu global en éléments transposables.

## Introduction

Les éléments répétés, comprenant notamment les éléments transposables (TE), sont une composante majeure des génomes eucaryotes (Lodish et al., 2005). La proportion de TE dans les génomes d'insectes peut être extrêmement variable allant de 1% chez le diptère *Belgica antarctica* (Kelley et al 2014) à 65% chez le criquet migrateur (Wang et al 2014). Il existe de très fortes variations entre les espèces et les ordres au regard de la composition en éléments transposables (Petersen et al., 2019). Les dynamiques d'invasion des TE au sein des génomes peuvent être différentes chez deux espèces proches. C'est le cas notamment chez *Drosophila simulans* et *D. melanogaster* où les dynamiques d'invasion sont régulées en partie par les ARNpi (Saint-Leandre et al., 2020). De même qu'il y a également des variations entre individus de la même espèce comme c'est le cas chez *Drosophila melanogaster* (Barrón et al., 2014 ; Lee and Langley, 2010).

Les TE, parfois appelés "gènes sauteurs" sont considérés comme des séquences nucléotidiques égoïstes, se propageant dans les génomes en ayant des effets neutres ou délétères (Barrón et al., 2014). Cependant, malgré ces effets négatifs, les TE sont aussi capables d'apporter un avantage sélectif ; que ce soit par de l'adaptation à un nouveau biotope (González et al 2008), la résistance aux insecticides (Chen et Li 2007), ou encore une adaptation au climat (Kim et al 2014). De manière générale les TE sont impliqués dans l'évolution structurelle et fonctionnelle des génomes (Gogvadze and Buzdin, 2009). Il a été montré chez la souris, qu'une partie des éléments transposables exerçaient des changements détectables dans l'expression des gènes (Nellåker et al., 2012). De même chez la phalène du bouleau (*Biston betularia*), alors que le papillon est naturellement blanc, le morphotype noir est dû à la présence d'un élément transposable, inséré dans l'intron d'un gène (Hof et al., 2016).

En fonction de leur mode de déplacement les TE sont rangés dans deux classes (Goodier and Kazazian, 2008). La classe 1 qui se déplace dans les génomes par un intermédiaire à ARN. Ces éléments appelés rétrotransposons peuvent avoir de longues répétitions terminales (LTR). Les TE de la classe 1 qui n'en possèdent pas sont également séparés en deux différents TE. Les LINE (longs éléments nucléaires intercalés) et les SINE (petits éléments nucléaires intercalés). La classe 2 dans les éléments transposables n'a pas d'intermédiaire à ARN, ils sont appelés transposons.

*Hermetia illucens*, la mouche soldat noire (BSF) est, depuis une dizaine d'années, utilisée pour réduire les biodéchets (Gold et al., 2018 ; Amrul et al., 2022) et produire des protéines et des acides gras utilisables en tant que nourriture animale ou comme source de biocarburant (Wang and Shelomi, 2017). Une filière autour des différentes utilisations de la BSF s'est créée et est en cours d'industrialisation.

La BSF est originaire du nord de l'Amérique du Sud et s'est propagée il y a environ 2 millions d'années vers l'Amérique du Nord (Guilliet et al 2022 – Chapitre 2), au cours des derniers siècles des populations, issues des Amériques, se sont propagées sur l'ensemble des continents (Kaya et al 2021). La BSF s'est adaptée de façon remarquable à l'ensemble des biotopes et conditions climatiques. Toutefois les déterminants génomiques et génétiques de ces adaptations ne sont pas connus, en particulier l'adaptation à des biotopes différents de ceux d'origine. Notamment le rôle des éléments transposables via des modifications génomiques structurales et en qualité de moteur d'adaptation, ne sont pas connus chez la BSF.

Dans cette étude, nous nous proposons de définir quelles sont les grandes superfamilles d'éléments transposables présentes chez les mouches soldat noires, et en quelle quantité. Nous allons regarder ces différentes familles d'éléments transposables, d'un point de vue populationnel pour les étudier dans une optique de génomique comparative.

Puis nous essaierons également de voir si certaines familles d'éléments transposables sont présentes dans toutes les populations ou au contraire dans une seule. Ce premier panorama des éléments transposables est une base solide pour de futures investigations sur les interactions génotypes / phénotypes.

Ce travail a débuté dans les derniers mois de ma thèse et sera poursuivi par la suite. Il s'agit donc ici de résultats préliminaires qui ne demandent qu'à être poursuivis.

## Matériel et méthode

### Matériel biologique et génétique

Les mouches soldat noires étudiées dans ce présent document, proviennent majoritairement de France et ont été récupérées durant l'été 2020 en France avec l'aide du réseau compost citoyen (RCC, n.d.). Un imago

a été collecté au Mexique en 2016, et les autres proviennent d'Asie (Taïwan), d'Afrique (Ghana et Kenya) et d'Amérique du Sud (Guyane Française). Nous avons eu dans nos échantillons des larves semblables à celles de la BSF qui se sont avérées appartenir à l'espèce *Exaireta spinigera*. Ces individus, au nombre de trois ont été traités dans les mêmes conditions que les BSF.

*Exaireta spinigera*, est également un Stratiomyidae dont l'écologie est similaire à la BSF (Mackillop et al., 2022).

En plus de ces individus, nous avons également téléchargé des données de séquençage qui étaient disponibles sur SRA (Sequence Read Archive – NCBI). Cela nous a permis d'avoir les séquences de trois nouveaux individus. Deux d'entre eux proviennent des Etats-Unis (bio-projet PRJNA268391) et l'une du royaume uni (bio-projet PRJEB37575). Les données sont résumées dans les tableaux 1 et 2.

ORIGINE	NOMBRE D'INDIVIDUS
COLLECTE DANS LA NATURE	43
ORIGINE COMMERCIALE	13
ASSEMBLES DEPUIS SRA (SEQUENCES BRUTES)	3
TOTAL	59

Tableau 1 : Origine des individus étudiés

ORIGINE GEOGRAPHIQUE	NOMBRE D'INDIVIDUS
AMERIQUE DU SUD	4
AMERIQUE DU NORD	2
ASIE	3
OCEANIE	0
AFRIQUE	11
EUROPE (FRANCE)	39 (38)
TOTAL	59

Tableau 2 : Origine géographique des individus étudiés

## Séquençage

Un total de 56 individus ont été séquencés, en génome complet, en utilisant la technique illumina par la société Novogene UK (Cock et al., 2010 ; Erlich et al., 2008; Hansen et al., 2010; Jiang et al., 2011; Yan et al., 2013). La technique utilisée pour le séquençage s'est faite sans PCR, en paired end avec des fragments de 150 paires de bases. Nous avons obtenu un taux de couverture moyen de 4,68X ( $\pm$  0,5X).

Un total de 59 individus sont étudiés. Parmi eux trois ne sont pas des mouches soldat noires mais des *Exaireta spinigera*, traitées dans les mêmes conditions ; et trois n'ont pas été séquencées dans les mêmes conditions, mais proviennent des données génétiques présentes sur les bases de données du NCBI (deux des Etats-Unis et une du Royaume-Unis).

## Analyse

Afin d'avoir une première idée du contenu en éléments transposables, nous avons utilisé RepeatModeler et RepeatMasker v4.05 (<http://www.repeatmasker.org>) à la fois sur le génome de référence (Generalovich et al 2021) ainsi que sur le génome que nous avons assemblé à partir de BSF fournis par la société Cycle Farms (Chapitre 1 du présent manuscrit). RepeatModeler permet de constituer une base d'éléments transposables *de novo*, tandis que RepeatMasker va masquer dans l'assemblage les zones contenant des éléments répétés.

Les différents jeux de données des individus séquencés ont, quant à eux été analysés via dnaPipeTE V1.3.1 (Goubert et al., 2015). Ce dernier est un pipeline basé sur le caractère très répété des TE. Il est conçu pour trouver, classer et quantifier les éléments transposables et autres répétitions dans un ensemble de données de séquençage. On utilise des reads à faible niveau de couverture et sans étape d'assemblage. En utilisant dnaPipeTE, nous avons, analysé pour chaque individu, dont nous avons des séquences, la proportion de chaque classe d'éléments transposables (LINE ; SINE ; LTR ; DNA éléments ; Hélitrons). Dans un second temps, nous avons analysé les données des différentes superfamilles, en regardant cette fois le nombre de bases alignées. Nous avons calibré dnaPipeTe en utilisant les reads du génome de référence. Après une série d'essai entre 0,1 et 1X, le taux de couverture où la taille moyenne des consensus était la plus longue, était de 0,2X.

Enfin, après les analyses par dnaPipeTE, et en regardant le nombre de bases alignées, nous avons analysé pour chaque classe les familles d'éléments transposables. La visualisation des résultats s'est faite via R (<https://www.r-project.org/> - les scripts python et R sont disponibles en annexe 3.01 et 3.02 page 233 à 236).

## Résultats et discussion

### Proportion d'éléments transposables dans les génomes complets.

Le génome d'*Hermetia illucens* est un grand génome pour un diptère. En utilisant consécutivement RepeatModeler et RepeatMasker, nous avons identifié, que 57,4% du génome était constitué d'éléments répétés (figure 1). L'étude a été faite sur le génome de référence (Generalovich et al 2021).

31,8% du génome est composé de LINE, ce qui en fait l'élément le plus abondant au sein du génome de la BSF. Les LTR représentent 3,4%, les hélitrons 0,6%, les DNA éléments représentent 3,5%, les SINE représentent 0,8% et les simples répétitions 0,3%. 17% du génome est masqué sans être classé dans l'un des groupes précédemment cités.

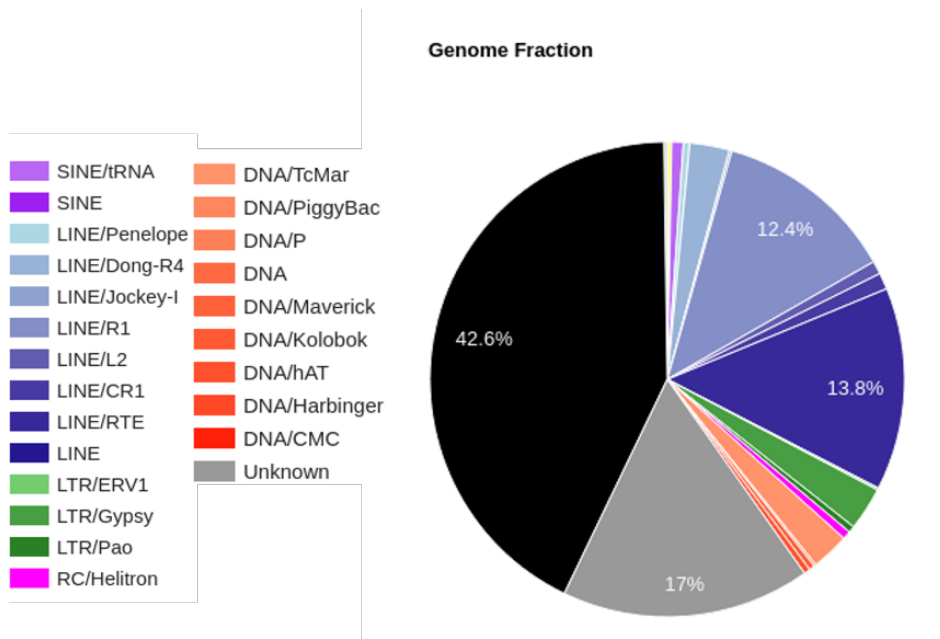


Figure 1 : Proportion des différentes superfamilles d'éléments transposables chez la BSF – génome de référence - réalisé avec RepeatMasker

Nous avons comparé les proportions des différentes superfamilles du génome de référence avec celles du génome que nous avons nous même assemblé précédemment (chapitre 1). Les résultats sont donnés dans la figure 2.



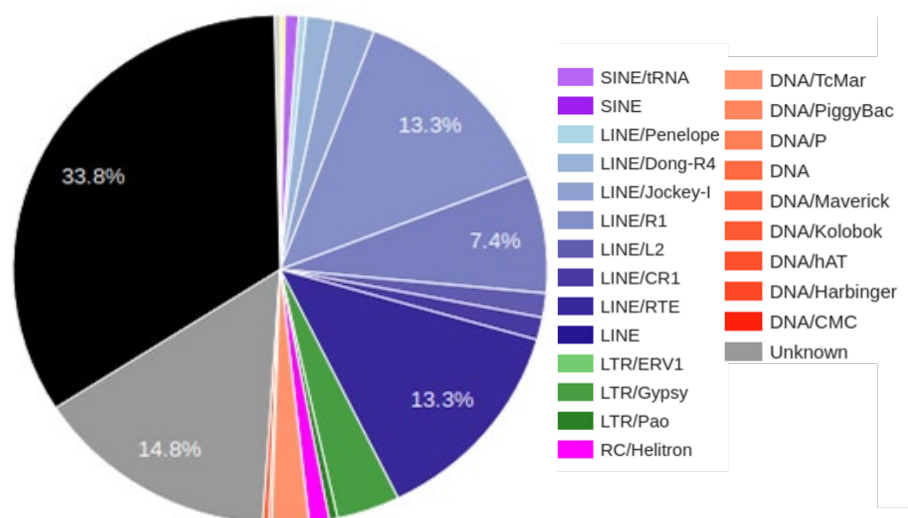
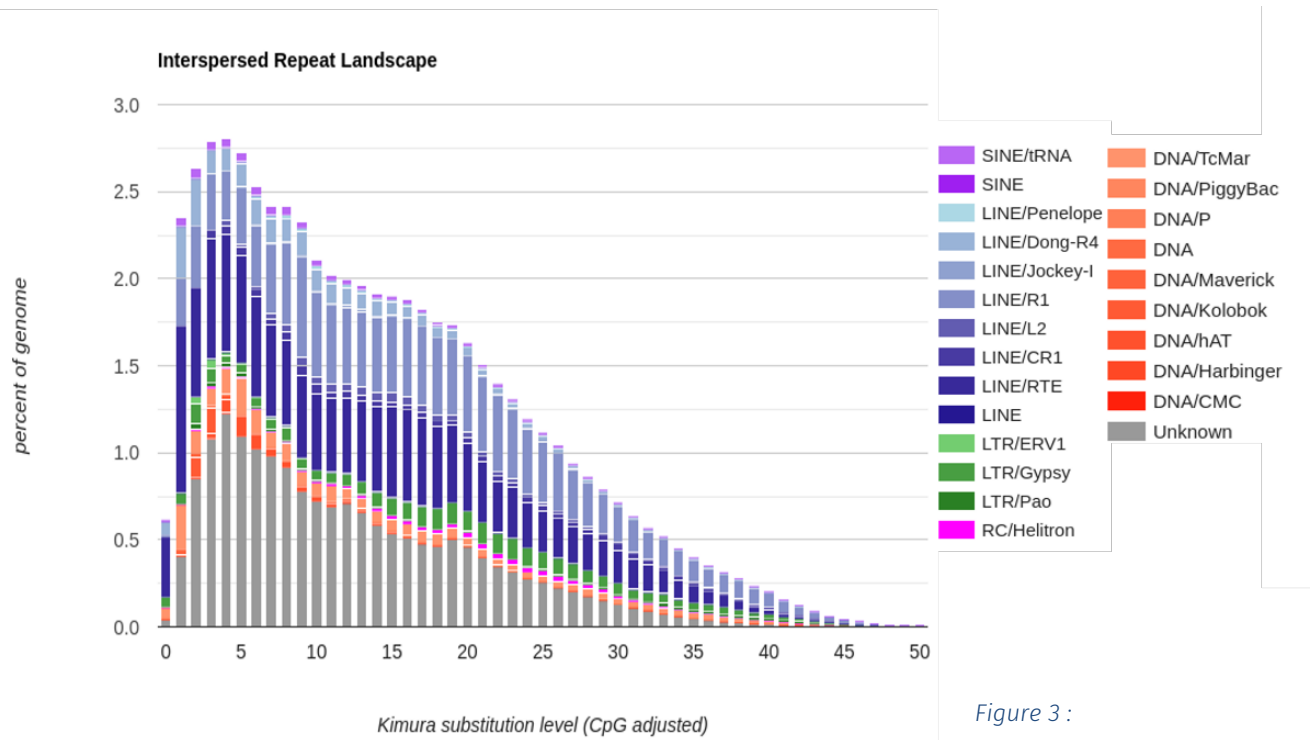


Figure 2 : Proportion des différentes superfamilles d'éléments transposables chez la BSF – génome Cycle Farms -réalisé avec RepeatMasker

Les proportions de superfamilles d'éléments transposables entre les deux assemblages sont similaires, bien qu'il semble y avoir une plus grande proportion d'éléments transposables dans le génome des BSF de Cycle Farms comparé à celle du génome de référence.

D'autre part, RepeatModeler permet également, de nous donner une indication sur le niveau de similitude des séquences d'éléments transposables d'une même famille. Concrètement un consensus de toutes les séquences appartenant à la même famille est généré. Puis une comparaison de chaque séquence homologue à la séquence consensus est faite. L'idée étant de comparer le nombre de mutations acquises par l'élément, en appliquant un poids différent entre les transitions et les transversions. On obtient un graphique, une substitution de kimura (figure 3). Ce genre de représentation visuel permet d'avoir une idée de l'ancienneté des séquences d'éléments transposables dans les génomes. Plus les séquences seront proches du consensus, plus elles seront récentes et inversement.



*Substitution de Kimura des différentes superfamilles d'éléments transposables chez la BSF – génome de référence.*

Beaucoup de séquences sont récentes (<5%) et ce, chez toutes les superfamilles à l'exception des LTR qui elles, semblent plus anciennes. Il y a donc une forte activité récente des TE qui montre que la dynamique au sein des génomes de BSF est très importante.

Toujours dans un objectif de comparaison avec le génome que nous avons assemblé (chapitre 1), nous avons également utilisé RepeatModeler afin d'observer et de comparer la dynamique des différentes familles d'éléments transposables (figure 4)

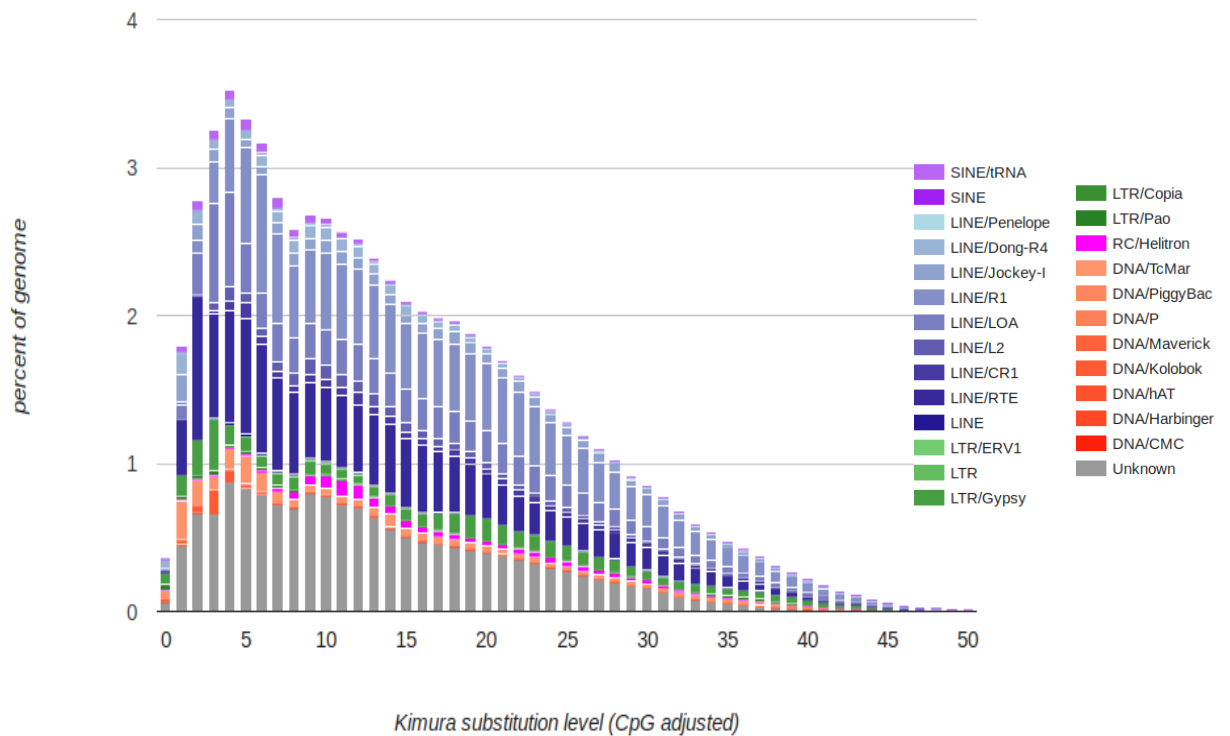


Figure 4 : Substitution de Kimura des différentes superfamilles d'éléments transposables chez la BSF – génome Cycle Farms.

Les résultats sont de nouveau similaires entre le génome de référence et celui des BSF issus de Cycle Farms.

## Comparaison des résultats entre RepeatModeler et dnaPipeTE

Afin de valider la méthode que nous voulons utiliser, nous avons pris des séquences brutes issues du séquençage du génome de référence, et nous avons utilisé dnaPipeTE pour comparer les résultats avec RepeatModeler (figure 5).

Le but était de voir si les résultats étaient comparables entre les deux logiciels qui utilisent des données en entrées différentes.

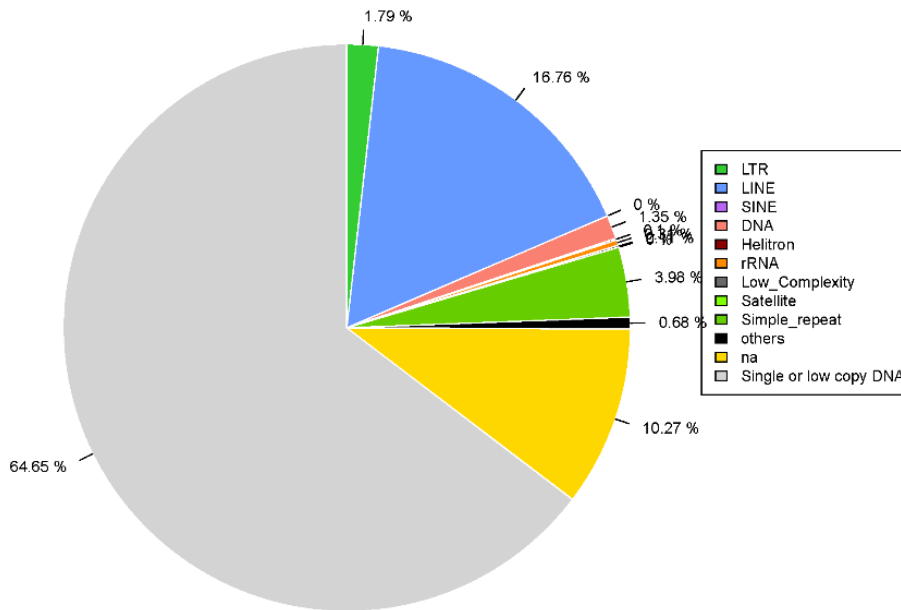


Figure 5 : Proportion des différents groupes d'éléments transposables chez la BSF réalisé avec dnaPipeTE

DnaPipeTE a trouvé 35,35% d'éléments répétés au sein du génome de référence (Generalovich et al 2021) contre 57,4% avec RepeatModeler.

GENOME REFERENCE	D'ELEMENTS REPETES	DE REPEATMODELER (%)	DNAPIPETE (%)
TOTAL		57,4	35,35
LINE		31,5	16,76
SINE		0,8	0
DNA		3,5	1,35
LTR		3,3	1,79
HELITRON		0,6	0,1

Tableau 3 : Comparaison des pourcentages des superfamilles d'éléments transposables entre RepeatModeler et dnaPipeTE

De manière générale, dnaPipeTE sous-estime la quantité d'éléments répétés dans le jeu de données que nous avons mis à sa disposition. En revanche, les proportions de chaque famille sont conservées.

Nous avons également voulu savoir combien de familles étaient retrouvées par RepeatModeler et par dnaPipeTE pour les deux génomes (tableau 4).

Génome	Génome de référence		Génome Cycle Farms	
	RepeatModeler	dnaPipeTE	RepeatModeler	dnaPipeTE
LINE	450	78	240	84
SINE	10	0	9	0
DNA	158	44	97	51
LTR	127	46	70	58
Hélitrons	7	4	8	3
Unknown	570	7	275	12
<b>Nombre total de famille (LINE ; SINE ; DNA ; LTR ; Hélitrons ; Unknown)</b>	<b>1322</b>	<b>187</b>	<b>699</b>	<b>216</b>

Tableau 4 : Comparaison du nombre total de familles de TE. Comparaison génomes et logiciels

De nouveau dnaPipeTE, sous-estime le nombre de TE dans les génomes. En revanche les écarts peuvent s'expliquer par le nombre de TE inconnus, trouvés par RepeatModeler. Les écarts entre les deux génomes, peuvent être le reflet des différences de tailles que nous avons identifiées au cours du premier chapitre.

Tous ces éléments pris en compte, nous avons validé dnaPipeTE, comme étant pertinent pour analyser les familles d'éléments transposables dans les génomes. Nous allons pouvoir le faire sur tous les individus que nous avons à notre disposition.

## Comparaison des quantités d'éléments transposables au niveau populationnel

Les différents individus étudiés dans cette partie étant ceux qui ont servi à établir la phylogénie mitochondriale présentée au chapitre 2 (Guilliet et al 2022). Nous avons classé les résultats en accord avec cette phylogénie (figure 6).

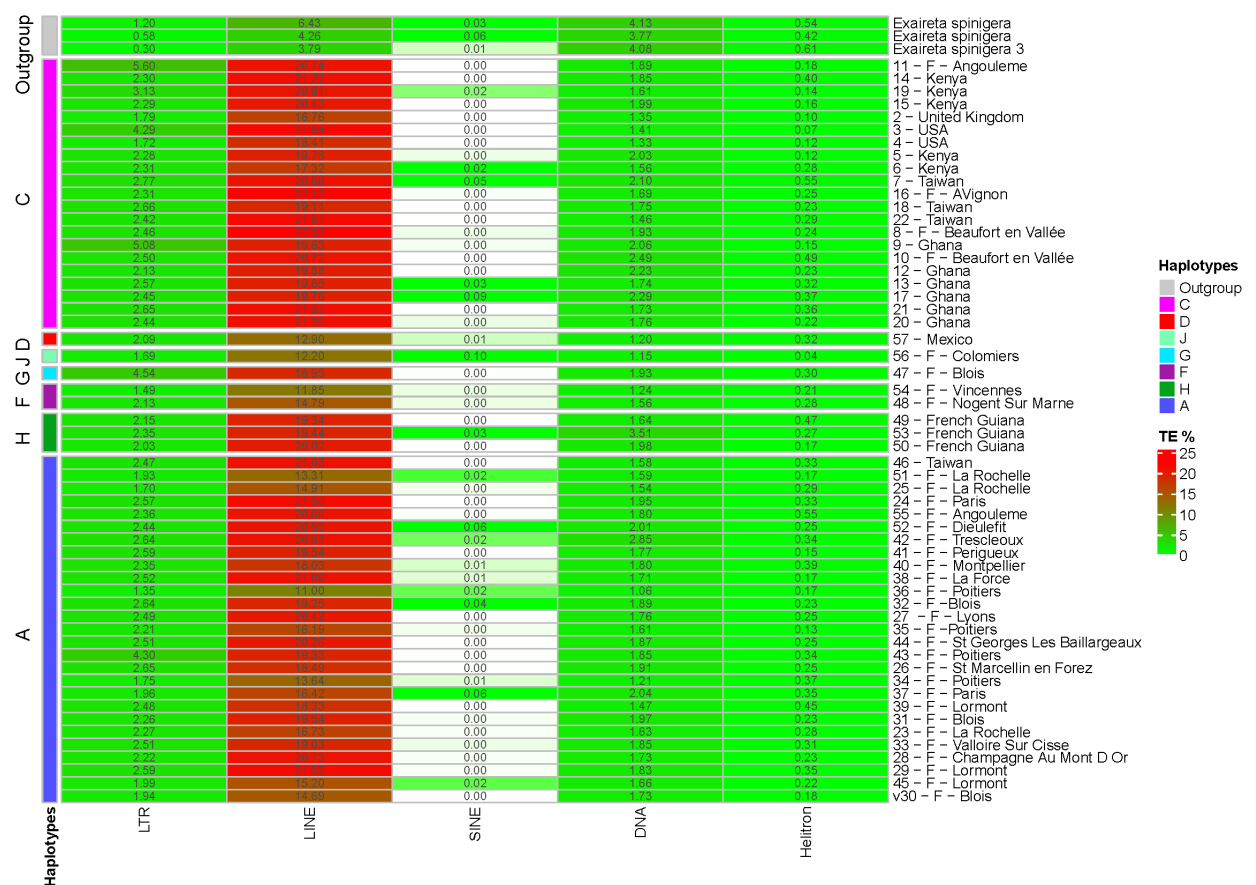


Figure 6 : Pourcentage de chaque superfamilles d'éléments transposables dans les données de séquençage des différentes populations de BSF

Chez les BSF, ce sont toujours les LINE qui sont les TE les plus représentés et ce peu importe la population. Ils varient de 11% à plus de 22% (moyenne  $18,7 \pm 2,3$ ), en comparaison, les LINE sont peu représentés chez *Exaireta spinigera*, qui nous avait servi d'outgroup pour la phylogénie.

Les éléments à ADN varient de 1,06 à 3,51% (moyenne à 1,8 ± 0,3), les hélitrons varient de 0,04 à 0,55% (moyenne à 0,3% ± 0,1), les LTR varient de 1,35% à 5,60% (moyenne à 2,5% ± 0,5) et les SINE sont toujours très peu représentés dans l'analyse dnaPipeTE où l'on a un maximum de 0,1%.

Les fortes variations qui apparaissent au sein des superfamilles, sont par ailleurs reflétées chez les différents individus, sans lien, entre les localités de capture, ni avec l'haplotype mitochondrial (figure 7). Certains individus de la même population et provenant de la même localité (Guyane Française) présentent des contenus en éléments transposables identiques. Tandis que dans d'autres populations (Blois de l'Haplotype A) ils varient entre 18,5 et 25%. De façon générale, les individus de l'haplotype C (en rose) sont plus homogènes au regard de leur contenu en TE. Tandis que ceux de l'haplotype A (bleu foncé) présentent d'avantage de variations quant à leur proportion en TE. Cela remet en question la pertinence d'utiliser une phylogénie mitochondriale, comme proxy de l'histoire évolutive du génome nucléaire. Afin de le vérifier il faudrait refaire une phylogénie basé sur des gènes nucléaires. Par ailleurs, comme nous l'avons vu au cours du premier chapitre, il pourrait y avoir des variations de taille de génome. Ces dernières pourraient biaiser les résultats d'analyses par dnaPipeTE ; celles-ci étant

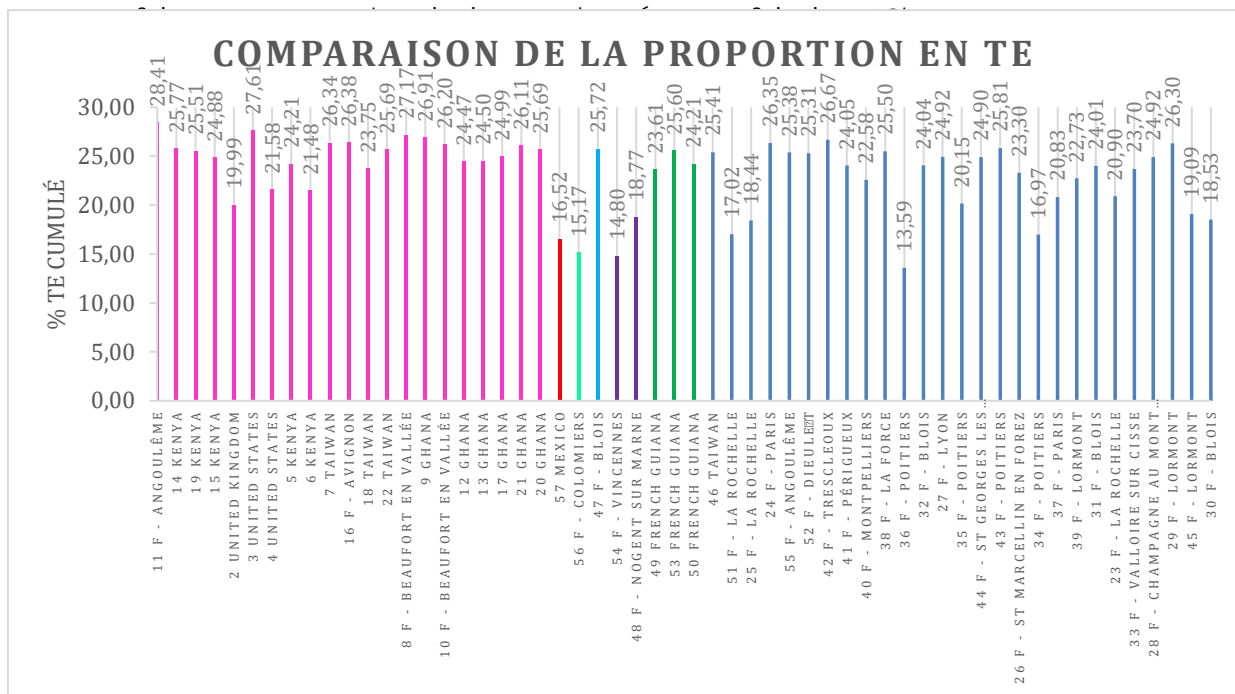


Figure 7 : Comparaison des proportions des principales superfamilles d'éléments transposables (LINE ; SINE ; DNA ; LTR ; Hélitrons)

Nous avons par la suite regardé dans les superfamilles de chaque groupe d'éléments transposables, celles qui étaient les plus représentées (figure 8).

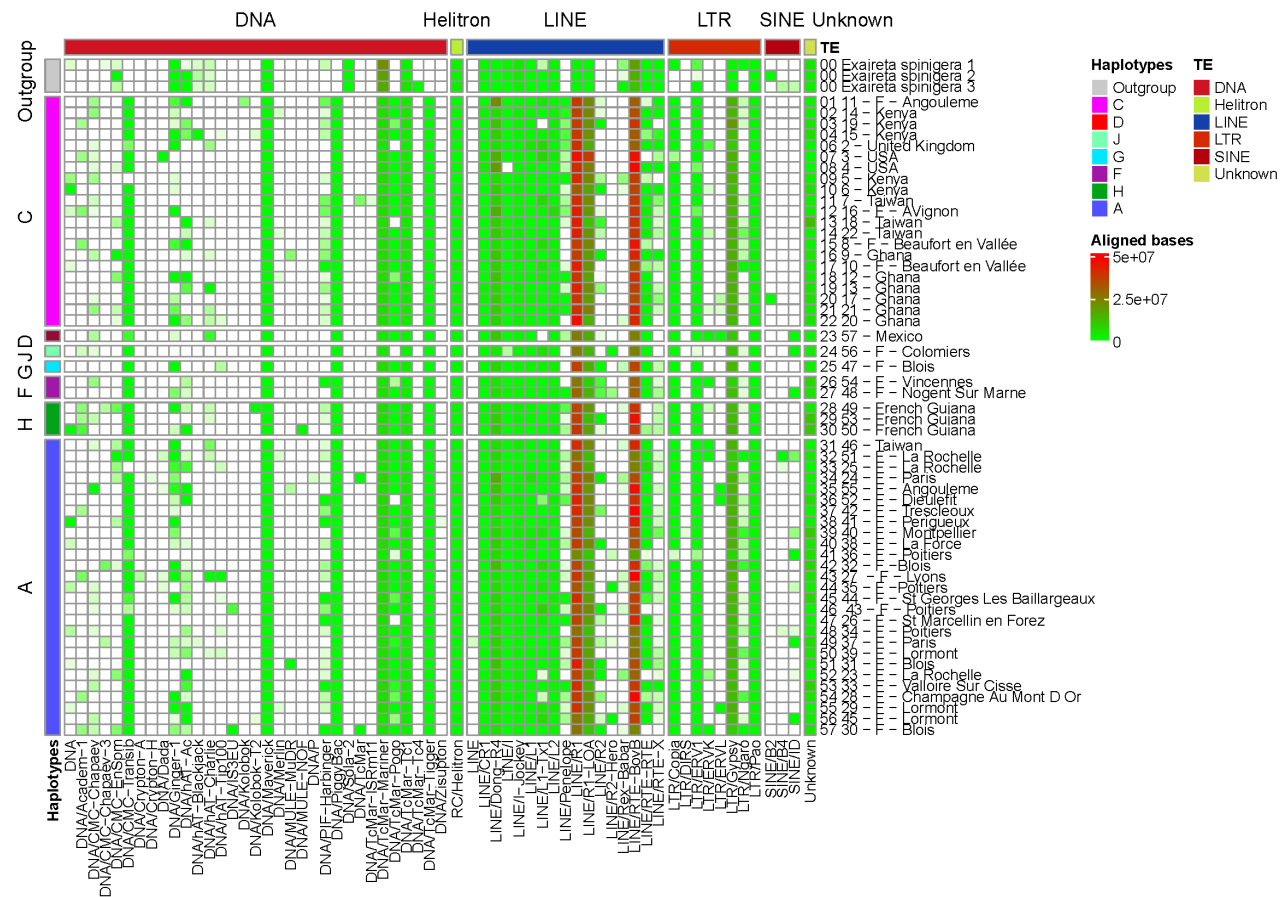


Figure 8 : Nombre de bases alignées pour chaque superfamille d'éléments transposable chez différents haplotypes de BSF

La plupart des superfamilles de LINE sont bien représentées chez tous les individus avec une surreprésentation des LINE R1 et LINE RTE/BovB. Il ne semble pas y avoir de structuration spécifique des éléments transposables, en fonction de l'haplotype considéré. Certaines superfamilles sont présentes chez tous les individus c'est le cas notamment des LINE CR1. Au contraire, quelques-unes ne sont présentes que chez certains individus tels que les DNA Tcmar, sans relation évidente avec les haplotypes mitochondriaux précédemment identifiés.

Nous avons par la suite regardé pour chaque famille au sein des superfamilles afin d'affiner encore nos résultats (figure 9 à 12). Les résultats sont à nouveau disparates et certaines familles sont présentes à nouveau chez tous les individus.



De manière générale un nombre limité de famille sont responsables de l'écrasante majorité des TE dans toutes les populations. C'est le cas notamment des LINE R1 et R1-LOA (figure 9) et de certaines famille de DNA TcMar (Figure 10). De même chez les LTR (figure 11) deux familles de la superfamille Gypsy présente des différences notables entre les individus.

Les variations de quantité observées auparavant (figure 7) ne sont pas dues à de nouvelles familles présentes dans quelques population mais à une variation de l'abondance relative des même familles. Toutefois il existe de nombreuses familles orphelines retrouvées dans un nombre limité de populations. Tel que l'élément P (figure 10), mais toujours en faible abondance. Deux familles d'éléments transposables inconnues (figure 12) sont aussi présentes et grandement représentées dans tous les génomes avec de fortes disparités entre les individus.



Figure 10 : Nombre de bases alignées par familles de DNA elements

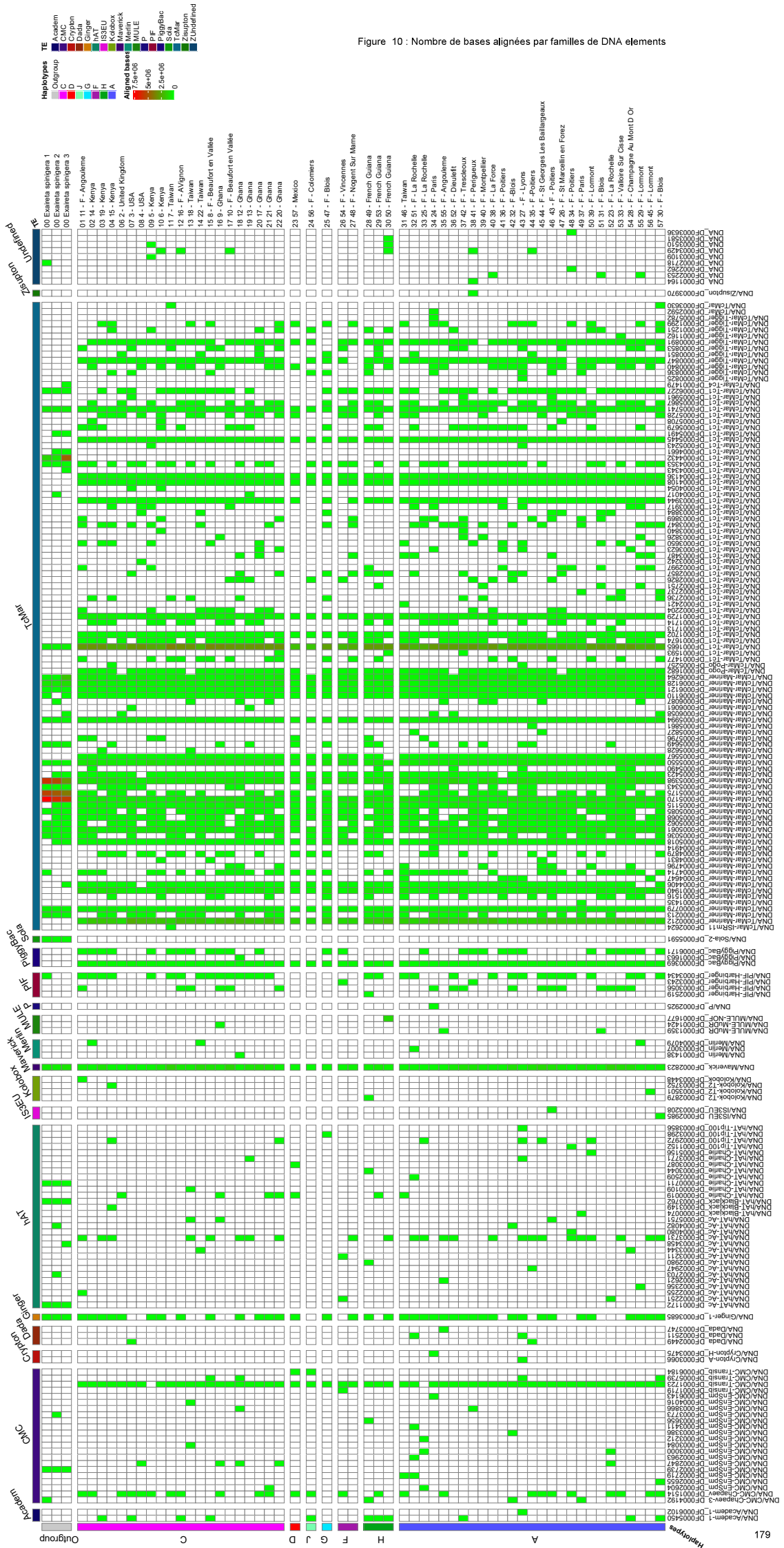


Figure 11 : Nombre de bases alignées par familles de LTR

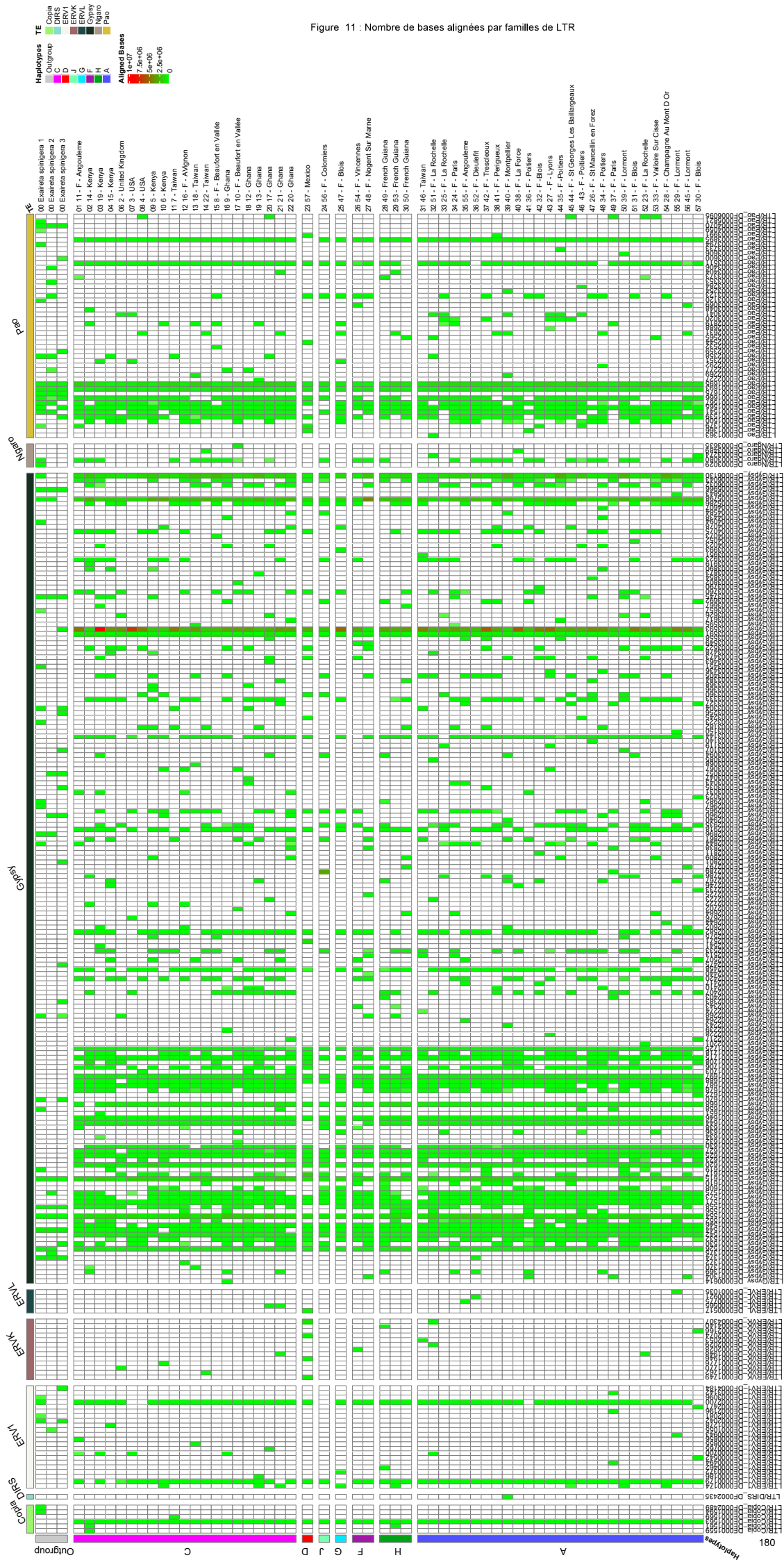
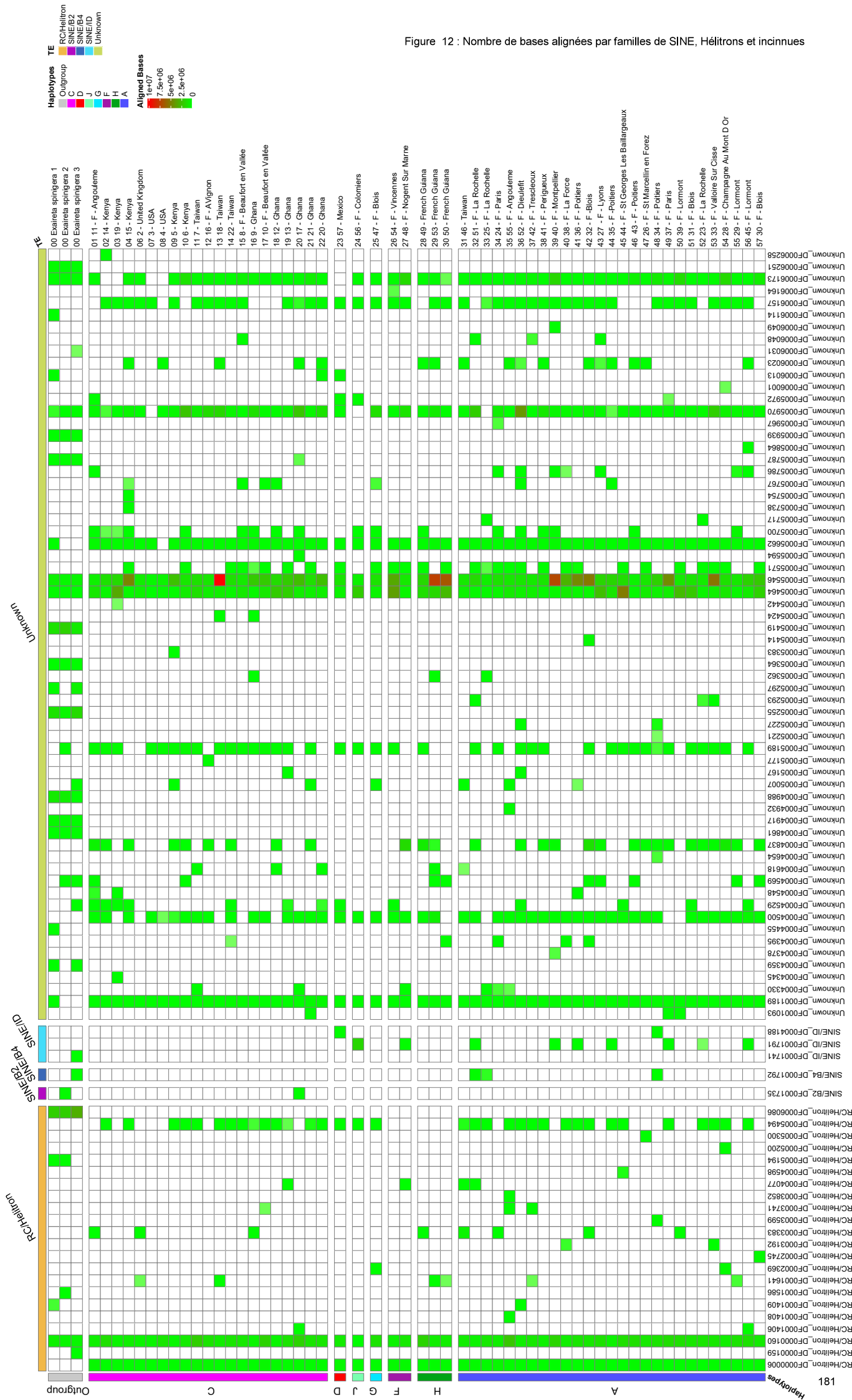


Figure 12 : Nombre de bases alignées par familles de SINE, Hélicons et incinues



## Conclusion

Les éléments transposables jouent un rôle considérable dans l'évolution des génomes (Ayarpadikannan and Kim, 2014). Au cours de cette étude préliminaire, nous constatons que des différences majeures apparaissent entre les génomes, quant à la composition en TE. Nous allons continuer notre investigation sur les différentes familles de TE présentes dans tous les génomes, ainsi que ceux qui ne sont présents, que dans un nombre réduit de génomes. Nous allons également nous intéresser à la dynamique des éléments transposables, en comparant la présence ou l'absence de certains éléments transposables pour un locus précis. Nous voulons faire cela au regard du génome de référence, ce qui nous donnera une idée de la dynamique des éléments transposables. Enfin, nous allons également rechercher d'éventuels transferts horizontaux chez les TE orphelins. Ces travaux conduiront à la publication d'un article de recherche, sur la dynamique des éléments transposables appliquée à de la génomique des populations.

# Discussion générale

Il y a peu de domaines, à mon sens, qui soient aussi plaisants que la biologie. L'étude du vivant est à la fois d'une telle complexité et d'une telle évidence que chaque aspect est un pur régal.

J'ai eu la chance au cours de ma thèse d'être parmi les pionniers d'un sujet émergent. Bien évidemment, là je ne parle ni de biologie, ni de génétique, mais bien de l'étude des *Hermetia illucens*. L'élevage des BSF, est de par sa nature profonde, rien de plus qu'un énième élevage d'animaux, mais comme nous allons le voir il est différent par bien des aspects.

## Rappel des principaux résultats

Nous sommes désormais mieux armés (comme les *Stratiomyidae*) pour aborder les futures étapes en sélection et en amélioration des BSF d'élevage.

Nous disposons d'un génome assemblé à l'échelle chromosomique ainsi que d'un catalogue de gènes (Chapitre 1 - [Generalovic et al., 2021](#)). Nous avons également relevé de potentielles différences de tailles importantes entre les génomes qui font partie d'un même haplotype mitochondrial.

Nous avons désormais une bonne idée de la diversité génétique de la BSF, à savoir qu'il existe une grande richesse génétique dans la nature, alors qu'au contraire, les élevages semblent se concentrer sur une seule souche. Nous avons également présenté plusieurs scénarii possibles quant à la dispersion de la BSF, sur son continent d'origine, ainsi que sur le reste de la planète ([Guilliet et al 2022](#) - Chapitre 2). Nous avons levé des inquiétudes quant à la biosécurité de ces élevages et mis en relief le temps de divergence entre les principaux haplotypes ; le reliant potentiellement avec un événement géologique majeur, à savoir la fermeture de l'isthme de Panama. Enfin, nous avons un premier catalogue des éléments transposables qui sont présents et actifs au sein des génomes d'*Hermetia illucens* (Chapitre 3). Nous y avons découvert de fortes variations entre les populations, sur la proportion totale d'éléments transposables. Il en est de même pour la présence systématique de certaines familles, alors que d'autres ne sont présentes que dans certains génomes sans lien avec la phylogénie mitochondriale.

## Retour sur les articles

Je le disais un peu avant, il y a des distinctions entre l'élevage de BSF et les autres élevages. Le premier aspect sur lequel ce domaine est différent est lié à sa récence : il s'agit d'une tentative originale et pionnière de domestication d'une espèce d'insectes pour l'agro-alimentaire. La plupart des élevages sont très anciens dans l'histoire humaine, à l'image du cheval en -3500 ([Outram et al., 2009](#)) ou encore - 6000 pour les poules ([West and Zhou, 1989](#)) et même la domestication des abeilles qui remonte à 7000 ans avant notre ère ([Oldroyd, 2012](#)). *A contrario*, les premiers élevages de



mouches soldat noires, n'ont vu le jour qu'au cours des dernières décennies (Sheppard et al., 1994). C'est une occasion unique de pouvoir appliquer les nouvelles méthodes apprises de la génétique et de la biologie des populations. Nous avons le choix de ne pas recommencer les mêmes processus de sélection que ceux qui ont été fait sur les autres élevages et de gagner un temps considérable, que ce soit dans l'élevage ou dans la sélection. Nous pouvons également éviter de reproduire des erreurs qui ont été commise dans d'autres élevages, comme la race de vache blanc bleu belge incapable de vêler par elle-même (Klein, 2022).

C'est avec cette volonté de gagner du temps (et donc de l'argent ne nous mentons pas) que les entreprises de l'insecte se projettent rapidement dans des aspects génétiques et de sélection.

Comme nous l'avons vu au cours du premier chapitre, avoir un génome de référence pour l'espèce, ainsi qu'un catalogue des gènes sont des prérequis indispensables pour la sélection. Ces génomes de haute qualité sont la pierre angulaire de la génomique moderne ("A reference standard for genome biology," 2018). De plus, dès qu'un génome de référence est disponible, les reads issus des nouveaux séquençages, seront alignés sur celui-ci plutôt que d'être assemblés *de novo*. Ce genre de technique représente un gain de temps (Hari and Parthasarathy, 2019).

Aussi dans le monde de la BSF, nous avons un génome fait à partir d'une seule population (Generalovic et al 2021). Ce génome ne peut donc pas, à proprement parler, correspondre à un génome de référence. En effet, il faudrait que celui-ci prenne en compte les différentes populations. Qu'il soit un assemblage chimérique de l'ensemble de ceux-ci pour être une séquence nucléotidique consensus de l'ensemble de l'espèce. L'idée est d'avoir une vision pan-génomique de l'espèce pour éviter les effets négatifs liés à l'utilisation d'un génome basé sur un seul individu (Morneau, 2021). Les travaux à mener dans les prochaines années auront, entre autres, pour objectif de donner un aperçu global du génome de l'espèce.

Les récentes avancées en matière de séquençage ont rendu les coûts de plus en plus abordables ("DNA Sequencing Costs," 2021) et par conséquent le reséquençage de génomes entiers devient une alternative extrêmement intéressante. Elle permet ainsi d'aller observer finement les variations génétiques entre les individus (Ng and Kirkness, 2010). Cela permet d'aller chercher au cœur des génomes les variations génétiques associées à un trait phénotypique ou encore à une maladie (Ikegawa, 2012). Ce genre d'analyse appelée GWAS est de plus en plus utilisé sur les animaux d'élevage et notamment chez les insectes tel que l'abeille (Avalos et al., 2020). Il est également possible d'aller rechercher les fréquences de

certains allèles mineurs (MAF) qui jouent un rôle important dans la l'héritabilité de certains traits (Hernandez et al., 2019).

Par ailleurs, une fois un génome de référence établi, il convient de s'intéresser à la diversité génétique des populations qui composent l'espèce.

En effet, lors d'un processus de sélection, les possibilités sont limitées par la diversité génétique présente (exception faite de la mutagenèse).

Il est possible, tout d'abord, d'estimer la diversité génétique au sein des élevages, même si nous savons que la mise en élevage de la BSF entraîne une diminution en quelques générations, de la diversité génétique (Hoffmann et al., 2021 ; Rhode et al., 2020). Cependant, nous ne savons pas actuellement si cette diminution de la diversité génétique au cours des générations se stabilise au cours du temps ni quel est l'impact de la richesse initiale sur la diminution. Nous en sommes réduits à constater la richesse allélique après un certain nombre de génération d'élevage.

Une fois cette diversité constatée, il devient alors possible d'appliquer une pression de sélection afin d'améliorer certains aspects phénotypiques de la mouche soldat noire. Certains de ces aspects ont déjà été explorés, que ce soient les effets de la consanguinité qui augmenterait le nombre d'œufs pondus (Cai et al., 2022), ou encore sur de la sélection massale qui augmente le poids moyen des individus au cours des générations (Facchini et al., 2022).

Malheureusement, sans moyen de réintroduire de la diversité génétique, les possibilités d'amélioration d'un élevage sont limitées à sa richesse allélique avant les étapes de sélection.

Il est également envisageable de réaliser des mutants, que ce soit via de la mutagenèse aléatoire ou dirigée, comme cela a déjà été fait chez la BSF (Zhan et al., 2020). Toutefois, les implications légales quant à la production d'OGM sont lourdes ("OGM," 2018), et les demandes d'autorisation de mise sur le marché et de traçabilité (la traçabilité n'est pas exigée dans le cadre de mutagenèse aléatoire) prennent du temps et sont contraignantes pour les professionnels.

Une autre solution serait d'arriver avec des méthodes de sélection massale et sans ajout de diversité génétique à obtenir des animaux dont les phénotypes correspondent aux attentes. Il faudrait dans ce cas, contrôler l'environnement pour qu'il soit le plus stable possible. Cependant les limites de ce genre de modèles sont rapidement atteintes (Magne et al., 2019).

Il y a effectivement des risques, quel que soit l'élevage, d'avoir des changements environnementaux, une indisponibilité des intrants couramment utilisés, ou encore, l'arrivée d'un pathogène. Et pis, il a été montré qu'une amélioration génétique à court terme de la productivité est souvent corrélée avec une diminution sur le long terme (Weiner, 2003).

Il est donc, à un certain point, indispensable de rajouter de la diversité dans les élevages.

Pour se faire, il est possible de réintroduire des individus génétiquement différents, que l'on aura sélectionné (ou non) afin d'améliorer les possibilités génétiques de l'espèce. Cela est d'autant plus intéressant chez les insectes, notamment les BSF où il reste dans le monde une très grande quantité de populations sauvages (Rhode et al., 2020). Ainsi, il est indispensable d'estimer la richesse génétique des populations, avant de se lancer dans ces processus de croisements.

La diversité s'exprime de différentes manières dans les élevages :

Il y a évidemment la diversité génétique dont nous parlons plus spécifiquement ici. C'est-à-dire le polymorphisme génétique entre les individus de l'espèce. Repérer des variants génétiques d'intérêt (résistance aux pathogènes, amélioration des capacités de production) est indispensable (Phocas et al., 2017).

Il y a également la diversité phénotypique, qui sont les traits observables de l'animal, tels que la masse, le comportement, la production d'œufs ... Ces caractères sont souvent la combinaison de caractéristiques génétiques et environnementales. Ce sont ces critères là que les éleveurs tentent d'améliorer.

La diversité spécifique, ne dépend pas en elle-même de l'animal élevé mais de l'interaction avec d'autres êtres vivants, comme un ajout de probiotique.

Enfin, toutes ces diversités entrent en compte dans ce que l'on appelle la diversité fonctionnelle (figure 1). Celle-ci se définit comme étant celle qui est intéressante pour l'élevage (Tichit et al., 2011).

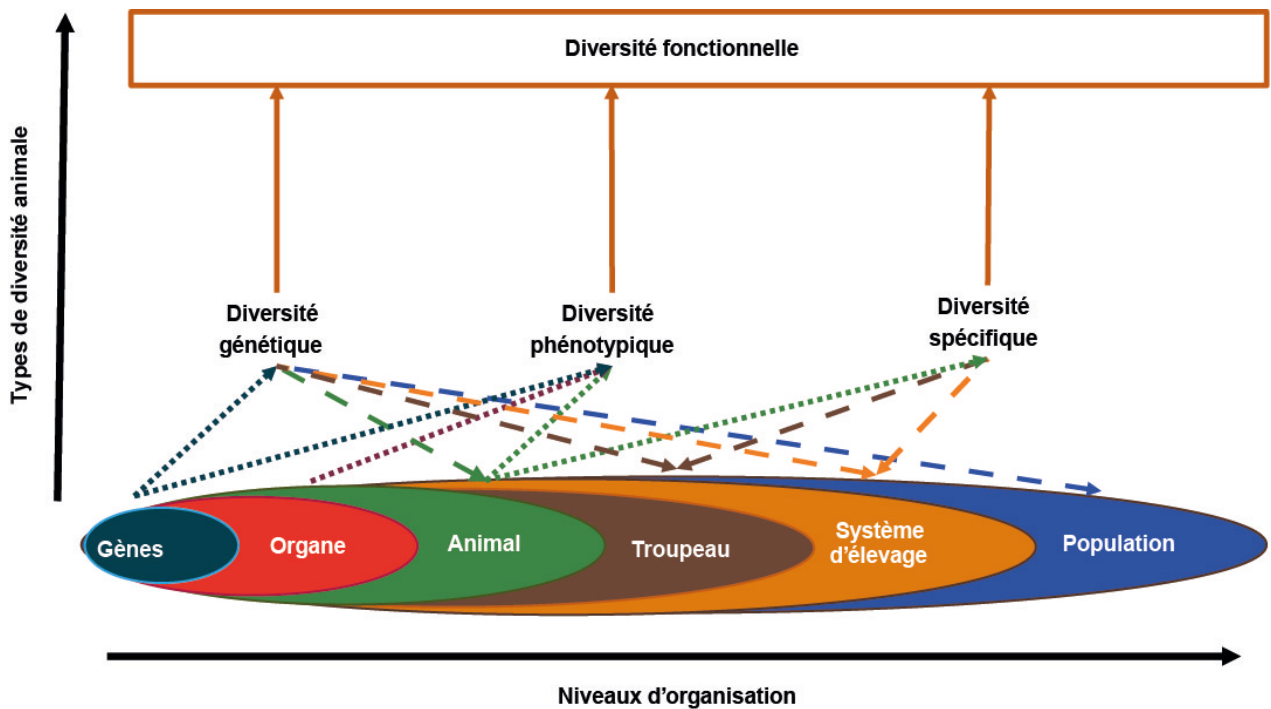
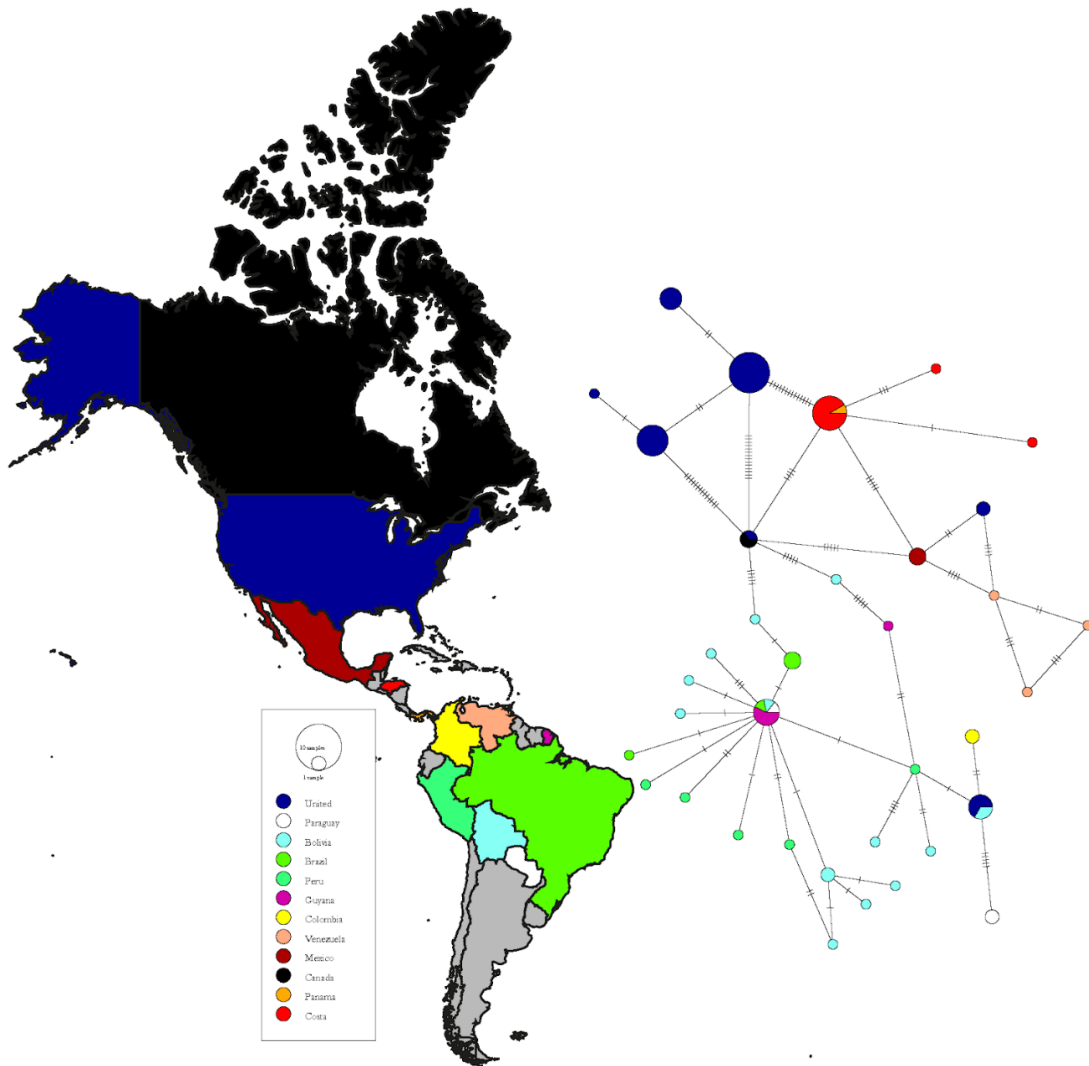


Figure 1 : Les différentes diversités et leurs interactions. D'après Magne et al 2019

Car il convient de gérer au mieux cette diversité, l'étudier est un prérequis indispensable. Nous l'avons vu au cours du deuxième chapitre de cette thèse, la mouche soldat noire présente une diversité génétique étonnante et riche. Cette diversité trouvée au sein du génome mitochondrial, nous permet de dater une séparation des principales populations à plus de deux millions d'années. Ceci promet une richesse importante d'un point de vue nucléaire. En effet bien que le ratio de mutation entre l'ADN mitochondrial et l'ADN nucléaire varie de 2 à 6 chez les invertébrés (Allio et al., 2017), l'extraordinaire variation trouvée entre les haplotypes mitochondriaux chez la BSF nous permet d'espérer de grandes variations génotypiques et potentiellement phénotypiques.

Par ailleurs, l'étonnant gradient nord-sud trouvé au sein des haplotypes CO1 (figure 2) du continent d'origine, tend à prouver la progression vers le nord de la mouche au cours de son évolution. Ces résultats préliminaires font de la BSF, un modèle unique pour l'étude de la dispersion naturelle d'une part et médié par l'homme d'autre part (Kaya et al 2021). D'autre part, cela est également une preuve de son adaptabilité à différents climats, et de fait,

l'étude de sa diversité génétique peut nous apporter des réponse sur les améliorations possible dans les élevages.



En effet la BSF a connu une dispersion naturelle entre sa zone native et l'Amérique du Nord ; et plus récemment une dispersion rapide sur l'ensemble de la planète. A l'image des variations au sein des gènes mitochondriaux, que nous avons trouvés et présentés au cours du chapitre 2, la mouche soldat noire peut devenir dans les prochaines années un modèle d'étude de l'invasion et de l'adaptation.

Enfin, un autre aspect de la diversité génétique pour lequel l'étude des génomes peut apporter des réponses est la diversité des éléments transposables. En effet, il a été prouvé que les éléments transposables étaient corrélés à l'apparition d'un stress chez les drosophiles (Rech et al., 2019). De même chez les plantes certains éléments transposables sont réactivés en cas de stress. Ces derniers peuvent être de différentes natures : en cas de sécheresse (Kalendar et al., 2000), en réponse au froid (Jiang et al., 2003) ou à la chaleur (Ito et al., 2011). Au vu de la dispersion rapide des BSF autour de la terre et leurs remarquables adaptations (Kaya et al 2021) il est probable qu'une part de cette adaptation soit causée par certaines familles d'éléments transposables. Comme nous l'avons vu au cours du troisième chapitre, les proportions d'éléments transposables peuvent varier assez drastiquement entre des individus phylogénétiquement proches.

## Les particularités du domaine

Une dizaine d'animaux sont élevés pour nourrir d'autres animaux, mais ils ont tous une application directe pour l'homme, exception faite de la BSF.

Animaux	Autres utilisations qu'en nourriture animale
Ver à soie	Sériculture
Souris domestique	Animal de recherche, animal de compagnie
Rat domestique	Animal de recherche, animal de compagnie
Grillon domestique	Nourriture humaine
Mouche du vinaigre ( <i>D. melanogaster</i> )	Animal de recherche
<i>Daphnia magna</i>	Animal de recherche
<i>Blaptica dubia</i>	Animal de recherche, animal de compagnie
Ténébrion meunier	Nourriture humaine
Vers de cire	Appât, animal de recherche
<b>Mouche soldat noire</b>	∅

Tableau 1 : Liste des animaux utilisés comme nourriture pour l'élevage et leurs autres utilisations direct pour Homo sapiens

Voici donc bien le premier élevage industriel, exclusivement dédié à l'élevage. C'est cet aspect qui fait que l'élevage de BSF est unique en son genre.

On pourrait ici objecter que les mouches soldat noires vont également diminuer les biodéchets, et cela est vrai, si elles sont nourries avec des déchets.

En effet, comme l'indique Seyedalmoosavi et al (2022), la BSF est une espèce qui n'entre pas en concurrence direct avec les élevages traditionnels (pour les ressources) mais plutôt comme une aide à ces élevages tant qu'elles sont élevées avec des aliments de qualités inférieures et inutilisables par ailleurs.

Cependant, il y a tout lieu de penser que ce n'est pas encore le cas. Les industriels du domaine cherchent pour l'instant, à comprendre et à maximiser les capacités des BSF en élevage. Aussi, il y a fort à parier que la quantité de biodéchets réellement incorporés dans les rations alimentaires est en dessous des capacités réelles de la BSF. Si tel est bien le cas, l'élevage de BSF n'est pas encore une véritable prouesse écologique, simplement une autre manière de produire de la protéine à destination des élevages avicoles et aquacoles.

Tant que les capacités exactes des BSF ne seront pas pleinement connues, il paraît impossible de les exploiter de la manière la plus optimale. Ce qui veut dire qu'après une décennie d'études poussées sur la BSF, nous réalisons quel est le potentiel de cet insecte (Siddiqui et al., 2022 ; Mertenat et al., 2019), mais, nous ne bénéficions pas encore de tout ce qu'elle a à offrir.

Il faut également prendre en compte l'aspect économique des élevages. Il y a en effet un véritable enjeu pour améliorer la rentabilité des élevages de BSF. La farine d'insectes (vendue entre 3500 et 5000 \$ la tonne en 2021) est en ce moment en concurrence avec la farine de poisson qui en Juillet 2022 était à un peu moins de 1600 \$ la tonne et le tourteau de soja qui lui avoisine les 540 \$ la tonne ("IndexMundi - Country Facts," n.d.). Alors, bien que l'impact écologique de la farine de BSF soit très probablement plus faible que les deux autres, l'enjeu de rentabilité des élevages reste une priorité pour les producteurs. Le salut de l'aliment à base d'insectes peut venir de deux façons différentes :

- L'augmentation des coûts des autres matières premières, due à leur impact écologique (ce qui implique une taxe des matières premières en réponse à leurs impacts écologiques).

- L'amélioration autour de l'insecte en lui-même qui comprend, à la fois, des améliorations techniques sur les procédés (élevages, extraction de protéines ...) et via la sélection d'animaux plus performants dans la transformation de biodéchets en aliment pour l'élevage.

Si le premier point paraît utopique, le second en revanche est une réalité que les acteurs industriels et chercheurs académiques tentent de mettre au point. En revanche, il faut aussi se méfier de ne pas perdre l'essence écologique des élevages de BSF. En effet dans un souci de diminution des coûts, la délocalisation est une option tentante mais qui entraînera une augmentation du transport de matière.

Un autre aspect, potentiellement problématique a également été levé dans le deuxième chapitre de cette thèse. Il s'agit des mesures de biosécurité qui devraient être mises en place pour éviter l'échappement des BSF issus des élevages. Nos travaux ont notamment montré la présence du génotype commercial dans la nature à proximité des élevages de BSF en France. En effet, jusqu'à présent aucune étude portant sur la dispersion des adultes n'a été réalisée. Il sera possible de le faire prochainement, car nous savons que les adultes ne craignent pas un marquage à la peinture acrylique (Jones and Tomberlin, 2020), et qu'à priori dans un environnement contrôlé, les BSF ne semblent pas s'éloigner beaucoup de leur zone d'émergence (Axtell and Arends, 1990 ; Tomberlin and Sheppard, 2002). En attendant que ces études soient réalisées et publiées, les éleveurs de BSF doivent faire tout ce qui est en leurs compétences et connaissances pour prévenir des échappements de leurs élevages. En effet, comme nous le savons, il existe d'une part une très grande homogénéité génétique dans les élevages et d'autre part, une grande diversité dans la nature (Chapitre 2 du présent manuscrit). Que ce soit dans une visée de conservation de cette biodiversité ou dans une optique d'amélioration des souches d'élevage, la préservation de la diversité est un élément d'une importance capitale. Des élevages laissant échapper des BSF dans des zones où il n'y a pas encore de BSF pourraient conduire à l'installation de mouches soldat noires dans ces zones et potentiellement perturber l'écologie locale. De même qu'un échappement continu dans des zones déjà occupées par une autre génétique pourrait conduire *in fine* au remplacement de la population source par celle de l'élevage.

Les échappements de mouches soldat noires peuvent intervenir à deux moments au cours de l'élevage. Le premier est lors de la pupaison. En effet, à ce moment précis du cycle de vie les larves vont quitter le substrat pour



trouver un endroit plus sec pour leurs pupaison (Sheppard et al., 1994). Si les conditions d'humidité sont adéquates alors les larves vont pouvoir s'échapper des bacs d'élevage ce qui peut créer de l'échappement. Il est donc primordial de contrôler cette humidité. L'autre période où il peut y avoir de l'échappement est au stade adulte. Les BSF volent lentement et ont tendance à se poser rapidement (Axtell and Arends, 1990). Chaque entrée et sortie des cages peut être l'occasion pour des imagos de se poser sur les manipulateurs et ainsi risquer de s'échapper. Il convient donc pour éviter tout risque d'échappement, d'une part, de multiplier le nombre de sas entre les cages et l'air libre et d'autre part, de renforcer les vérifications de toutes les personnes manipulant au niveau des serres.

L'élevage de BSF se veut écologique et économique. Cependant les gains peuvent être limités à cause des coûts de transport et l'achat de matière organique pour l'élevage (Beesigamukama et al., 2022). Pour ce qui est de la production de protéines, l'élevage de BSF dans de petites structures avicoles ou aquacoles est bénéfique aussi bien économiquement qu'écologiquement (Nyakeri et al., 2017).

Récemment des études ont également montré qu'un hydrolysate des protéines extraites des BSF avaient de meilleures propriétés fonctionnelles et une plus grande activité anti oxydante (Batish et al., 2020). Cependant, produire ces hydrolysats de protéines va entraîner une consommation démesurée d'énergie fossile, d'eau et participer au réchauffement climatique (Colantoni et al., 2017). Par exemple, produire un kilogramme d'hydrolysate de protéine à partir de co-produits animaux entraîne là l'émission de 1,5 kg d'équivalent CO<sub>2</sub>, la consommation de 25 MJ d'énergie fossile et la consommation de 116 litres d'eau. Ces hydrolysats, bien qu'étant potentiellement plus efficaces que les protéines brutes extraites des BSF, sont à mon sens une telle catastrophe écologique, que leurs productions viendraient en contradiction avec le *let motiv* des entreprises de l'insecte.

Par ailleurs, avec plus de 200 milliards de BSF élevées par an (Rowe, 2020), la notion de bien-être animal va très certainement être l'une des problématiques émergentes du domaine. A l'heure actuelle, aucune association de grande envergure ne s'est positionnée sur les élevages d'insectes. Sans aucun doute, cela viendra avec la croissance du secteur. Nous pouvons donc, à l'instar de Barrett et al. (2022) commencer à travailler sur ces questions pour que les élevages soient le plus respectueux possible des animaux. Les raisons pour lesquelles il n'y a pas encore de mouvement

important pour la conservation des *Hermetia illucens* (ou des autres insectes d'élevage qui sont sacrifiés au cours du processus) sont doubles. Le premier est lié à la récence des élevages, le second est quant-à-lui, un biais psychologique lié à l'anthropomorphisme (Urquiza-Haas and Kotrschal, 2015). Ce dernier point est similaire à celui qui fait que les espèces phylogénétiquement plus proches d'*Homo sapiens*, reçoivent plus d'attention quand il s'agit de conservation animale (Martín-Forés et al., 2013).

## Les perspectives

Le génome de référence que nous avons ne représente qu'un seul individu. Vouloir comparer les génomes de différentes populations peut également nous permettre d'en apprendre beaucoup plus sur l'espèce. Nous allons désormais concentrer une partie de nos efforts sur le séquençage des génomes complets des autres lignées que nous avons trouvé et présenté au chapitre 2.

Cela nous permet d'élucider la question des différences de taille de génome que nous avons trouvé au cours de ma première année de thèse (chapitre 1).

En effet, il y a un écart d'environ 150 Mb entre le génome de référence et celui que nous avons assemblé. Les génomes mitochondriaux présentent un pourcentage de similarité de 99.993% (chapitre 2). Par ailleurs, la proportion d'éléments transposables est d'environ 20% pour le génome de référence et de 25% pour celui de Cycle Farms (Chapitre 3). Toutes ces données, font que les différences de tailles de génomes, deviennent une nouvelle problématique passionnante. Nous avons également utilisé les données de séquençage que nous avons pour estimer la taille de deux autres génomes (tableau 2)

HAPLOTYPE	LOCALITE	TAILLE DE GENOME ESTIMEE (MPB)
C	Référence (UK)	1010
C	Cycle Farms	840
C	Ghana	936
A	Montpellier	878

Tableau 2 : Différences de tailles des génomes estimées

Les tailles ont été estimées en utilisant des tailles de Kmer de 21. Les graphiques associés à la souche Ghanéenne et Montpelliéraine sont donnés en annexe 4.01 (page 237).

Avec les résultats des génomes complets, nous pourrions commencer des études de type GWAS et rechercher des QTL impliqués dans des traits phénotypiques d'intérêt industriel. Cette approche de génomique comparative, se place dans la suite logique des travaux de ma thèse. Ils impliqueront les mêmes partenariats entre le laboratoire EGCE et l'entreprise Cycle Farms. Par ailleurs, nous allons également continuer à explorer la diversité des éléments transposables des génomes d'*Hermetia illucens*. Nous voulons déterminer les familles les plus actives, celles qui ne sont présentes que dans certaines populations et potentiellement celles qui seront responsables d'un trait phénotypique.

Ces perspectives pour les trois prochaines années amèneront à la publication d'au moins deux autres articles de recherches. Par ailleurs, comme de nouvelles lignées ont été découvertes au cours de cette thèse, toujours en lien avec l'entreprise Cycle Farms, celles-ci vont être mises en culture afin de les caractériser phénotypiquement. Notre but est à la fois de vérifier que les autres populations de BSF peuvent être mises en élevage, dans les mêmes conditions, de comparer les optimums de croissance de chacune des lignées, ainsi que de les caractériser dans des conditions similaires. L'idée est aussi de vérifier l'inter-fertilité des différentes populations. En effet, il se peut que des barrières comportementales, génétiques ou physiques ne le permettent pas. Une fois fait, le catalogue de gènes de chaque lignée pourra être réalisé afin de relier un phénotype à un génotype et potentiellement commencer des croisements entre les lignées.

Avec comme objectifs d'une part, d'améliorer les caractéristiques générales de l'espèce et d'autre part, d'essayer de spécialiser les lignées sur des substrats différents ou pour des applications différentes.

# Conclusion générale

De grands groupes industriels se sont articulés autour de l'élevage de ces mouches soldat noires, et la filière, encore jeune, commence à se structurer. Chacun des protagonistes trouvant petit à petit sa place, tant au niveau production de mouches que sur des aspects plus fondamentaux (tels que la reproduction et la sélection). Les enjeux écologiques et économiques sont prédominants dans ce domaine qui peut permettre de nourrir, indirectement, mais de façon durable une population humaine grandissante et souffrant d'un dérèglement climatique toujours plus fort. Devenir compétitif, pour les industries, passe par des améliorations des procédés et de la BSF. Améliorer les qualités d'*Hermetia illucens*, est un défi que les travaux de cette thèse permettent d'appréhender.

La sélection par les industries est un processus au cours duquel les individus les plus compétitifs vont être gardés pour se reproduire. Je dis ici compétitif et non pas adapté, car les enjeux humains et la sélection naturel ne vont pas de concert dans les élevages. En effet les industriel recherchent d'avantage un compromis adaptabilité / compétitivité. Comme un nombre d'œufs pondus plus important tout en diminuant la durée du cycle. Il peut s'agir également d'une résistance aux pathogènes ou à des températures plus basse.

Afin que cette sélection soit optimale, il faut que les industries puissent sélectionner les traits d'intérêts pour leurs besoins. Cela passe par une étape de définition des phénotypes à sélectionner, une seconde étape qui est de définir quels sont les génotypes associés et enfin comment ces traits peuvent être sélectionnés.

D'un point de vue biologique, ces traits sont la plupart du temps le résultat d'une interaction entre la génétique et l'environnement. Si les conditions environnementales peuvent être optimisées et stabilisées, la génétique quant à elle nécessite d'être étudiée et appréhendée de manière à pouvoir relier un génotype à un phénotype.

Au cours de cette thèse, j'ai commencé à ouvrir certaines portes autour de la génétique des populations chez *Hermetia illucens*. Ce diptère dont nous avons vanté les mérites au cours de cette thèse, et dont on peut, si l'on reste malin, tirer beaucoup en termes économique et écologique. Cependant, comme nous pouvons le constater encore une fois,

*Hermetia illucens* reste pleine de surprise et des questions fondamentales et appliquées continueront de se bousculer à nos portes.

La génétique de la BSF est loin d'être élucidée, et de nombreuses questions restent en suspens sur les capacités d'adaptation de la BSF et sur les résultats de mélange des populations.

La génomique de la mouche soldat noire est, au fur et à mesure que l'on se plonge dedans, de plus en plus passionnante, avec des génomes qui semblent avoir des tailles et des contenus en éléments transposables extrêmement diverses.

Ce travail, a répondu à des questions, nous avons désormais une bonne idée de la diversité génétique dans la nature (elle est riche et très diverse) et dans les élevages (où là elle très réduite). Mais nous avons également soulevé des questions :

- Pouvons-nous réduire encore le nombre d'hypothèses quant au scénario de dispersion et d'évolution de la BSF ?
- Quels sont les phénotypes associés aux génotypes que nous avons identifiés ?
- Peut-on mettre en élevage les autres populations ? Et présentent-elles les mêmes caractéristique que notre souche d'élevage ?
- A quoi sont dues ces différences de taille de génome, et dans quelle mesure les éléments transposables jouent un rôle dans les différences génomiques ?
- Dans quelle mesure peut-on améliorer les capacités de l'espèce ?
- Comment se prévenir des échappements de mouches soldat noires dans les élevages ?

# Bibliographie

## A

A reference standard for genome biology, 2018. . Nat Biotechnol 36, 1121–1121.

<https://doi.org/10.1038/nbt.4318>

Abrami, G., Mehler, A., Lücking, A., Rieb, E., Helfrich, P., 2019. TextAnnotator: A flexible framework for semantic annotations.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D., Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacleb, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Sidén-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wassarman, D.A., Weinstock, G.M., Weissenbach, J., Williams, S.M., WoodageT, null, Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M., Venter, J.C., 2000. The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195. <https://doi.org/10.1126/science.287.5461.2185>

ADEME, 2019. Publication ADEME : « Déchets chiffres-clés – L'essentiel 2019 » | FNCCR. FNCCR | Les services publics en réseau. URL <https://www.fnccr.asso.fr/article/publication-ademe-dechets-chiffres-cles-lessentiel-2019/>

Ait Aabd, N., Bouharroud, R., Tahiri, A., Wifaya, A., Mimouni, A., El Mousadik, A., 2019. Genetic Diversity and Breeding of Argan Tree (*Argania spinosa* L. Skeels), in: Al-Khayri, J.M., Jain, S.M., Johnson, D.V. (Eds.), Advances in Plant Breeding Strategies: Nut and Beverage Crops: Volume 4. Springer International Publishing, Cham, pp. 31–56. [https://doi.org/10.1007/978-3-030-23112-5\\_2](https://doi.org/10.1007/978-3-030-23112-5_2)

Alagappan, S., Rowland, D., Barwell, R., Mantilla, S. m. o., Mikkelsen, D., James, P., Yarger, O., Hoffman, L. c., 2022. Legislative landscape of black soldier fly (*Hermetia illucens*) as feed. Journal of Insects as Food and Feed 8, 343–355. <https://doi.org/10.3920/JIFF2021.0111>

All about feed, 2017. Update: A-Z of insect protein/oil companies [WWW Document]. All About Feed. URL <https://www.allaboutfeed.net/all-about/new-proteins/update-a-z-of-insect-protein-oil-companies/> (accessed 7.29.22).

Allio, R., Donega, S., Galtier, N., Nabholz, B., 2017. Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as

a Molecular Marker. *Molecular Biology and Evolution* 34, 2762–2772. <https://doi.org/10.1093/molbev/msx197>

Almeida, C., Murta, D., Nunes, R., Baby, A.R., Fernandes, Â., Barros, L., Rijo, P., Rosado, C., 2022. Characterization of lipid extracts from the *Hermetia illucens* larvae and their bioactivities for potential use as pharmaceutical and cosmetic ingredients. *Heliyon* 8, e09455. <https://doi.org/10.1016/j.heliyon.2022.e09455>

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)

Amrul, N.F., Kabir Ahmad, I., Ahmad Basri, N.E., Suja, F., Abdul Jalil, N.A., Azman, N.A., 2022. A Review of Organic Waste Treatment Using Black Soldier Fly (*Hermetia illucens*). *Sustainability* 14, 4565. <https://doi.org/10.3390/su14084565>

Arru, B., Furesi, R., Gasco, L., Madau, F.A., Pulina, P., 2019. The Introduction of Insect Meal into Fish Diet: The First Economic Analysis on European Sea Bass Farming. *Sustainability* 11, 1697. <https://doi.org/10.3390/su11061697>

Astuti, D.A., Wiryawan, K.G., 2022. Black soldier fly as feed ingredient for ruminants. *Anim Biosci* 35, 356–363. <https://doi.org/10.5713/ab.21.0460>

Avalos, A., Fang, M., Pan, H., Ramirez Lluch, A., Lipka, A.E., Zhao, S.D., Giray, T., Robinson, G.E., Zhang, G., Hudson, M.E., 2020. Genomic regions influencing aggressive behavior in honey bees are defined by colony allele frequencies. *Proceedings of the National Academy of Sciences* 117, 17135–17141. <https://doi.org/10.1073/pnas.1922927117>

Awal Rabiul, Rahman Masudur, Choudhury, A.R., Hasan Mehedi, Rahman Towfiqur, Mondal Fuad, 2022. Influences of artificial light on mating of black soldier fly (*Hermetia illucens*) - a review. *International Journal of Tropical Insect Science*. <https://doi.org/10.1007/s42690-022-00786-7>

Axtell, R.C., Arends, J.J., 1990. Ecology and Management of Arthropod Pests of Poultry. *Annual Review of Entomology* 35, 101–126. <https://doi.org/10.1146/annurev.en.35.010190.000533>

Ayarpadikannan, S., Kim, H.-S., 2014. The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. *Genomics Inform* 12, 98–104. <https://doi.org/10.5808/GI.2014.12.3.98>

## B

Barbier, J., 1952. Introduction en France d'un Diptère Stratiomyide américain. *Bulletin de la Société entomologique de France* 57, 108–108.

Barker, J.S., 2011. Conservation and management of genetic diversity: a domestic animal perspective. *Canadian Journal of Forest Research*. <https://doi.org/10.1139/x00-180>

Barrett, M., Chia, S. y., Fischer, B., Tomberlin, J. k., 2022. Welfare considerations for farming black soldier flies, *Hermetia illucens* (Diptera: Stratiomyidae): a model for the insects as food and feed industry. *Journal of Insects as Food and Feed* 1–30. <https://doi.org/10.3920/IJFF2022.0041>

Barrón, M.G., Fiston-Lavier, A.-S., Petrov, D.A., González, J., 2014. Population Genomics of Transposable Elements in *Drosophila*. *Annu. Rev. Genet.* 48, 561–581. <https://doi.org/10.1146/annurev-genet-120213-092359>

Barros, L.M., Gutjahr, A.L.N., Ferreira- Keppler, R.L., Martins, R.T., 2019. Morphological description of the immature stages of *Hermetia illucens* (Linnaeus, 1758) (Diptera: Stratiomyidae). *Microsc. Res. Tech.* 82, 178–189. <https://doi.org/10.1002/jemt.23127>

- Batish, I., Brits, D., Valencia, P., Miyai, C., Rafeeq, S., Xu, Y., Galanopoulos, M., Sismour, E., Ovissipour, R., 2020. Effects of Enzymatic Hydrolysis on the Functional Properties, Antioxidant Activity and Protein Structure of Black Soldier Fly (*Hermetia illucens*) Protein. *Insects* 11, 876. <https://doi.org/10.3390/insects11120876>
- Beesigamukama, D., Mochoge, B., Korir, N., Menale, K., Muriithi, B., Kidoido, M., Kirscht, H., Diiro, G., Ghemoh, C.J., Sevgan, S., Nakimbugwe, D., Musyoka, M.W., Ekesi, S., Tanga, C.M., 2022. Economic and ecological values of frass fertiliser from black soldier fly agro-industrial waste processing. *Journal of Insects as Food and Feed* 8, 245–254. <https://doi.org/10.3920/JIFF2021.0013>
- Belghit, I., Liland, N.S., Gjesdal, P., Biancarosa, I., Menchetti, E., Li, Y., Waagbø, R., Krogdahl, Å., Lock, E.-J., 2019. Black soldier fly larvae meal can replace fish meal in diets of sea-water phase Atlantic salmon (*Salmo salar*). *Aquaculture* 503, 609–619. <https://doi.org/10.1016/j.aquaculture.2018.12.032>
- Benelli, G., Canale, A., Raspi, A., Fornaciari, G., 2014. The death scenario of an Italian Renaissance princess can shed light on a zoological dilemma: did the black soldier fly reach Europe with Columbus? *Journal of Archaeological Science* 49, 203–205. <https://doi.org/10.1016/j.jas.2014.05.015>
- Berlocher, S.H., Feder, J.L., 2002. Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu Rev Entomol* 47, 773–815. <https://doi.org/10.1146/annurev.ento.47.091201.145312>
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* 69, 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- Bertinetti, C., Samayoa, A.C., Hwang, S.-Y., 2019. Effects of Feeding Adults of *Hermetia illucens* (Diptera: Stratiomyidae) on Longevity, Oviposition, and Egg Hatchability: Insights Into Optimizing Egg Production. *Journal of Insect Science* 19. <https://doi.org/10.1093/jisesa/iez001>
- Beskin, K.V., Holcomb, C.D., Cammack, J.A., Crippen, T.L., Knap, A.H., Sweet, S.T., Tomberlin, J.K., 2018. Larval digestion of different manure types by the black soldier fly (Diptera: Stratiomyidae) impacts associated volatile emissions. *Waste Management* 74, 213–220. <https://doi.org/10.1016/j.wasman.2018.01.019>
- Biasato, I., Renna, M., Gai, F., Dabbou, S., Meneguz, M., Perona, G., Martinez, S., Lajusticia, A.C.B., Bergagna, S., Sardi, L., Capucchio, M.T., Bressan, E., Dama, A., Schiavone, A., Gasco, L., 2019a. Partially defatted black soldier fly larva meal inclusion in piglet diets: effects on the growth performance, nutrient digestibility, blood profile, gut morphology and histological features. *J Anim Sci Biotechnol* 10, 12. <https://doi.org/10.1186/s40104-019-0325-x>
- Bondari, K., Sheppard, D.C., 1981. Soldier fly larvae as feed in commercial fish production. *Aquaculture* 24, 103–109. [https://doi.org/10.1016/0044-8486\(81\)90047-8](https://doi.org/10.1016/0044-8486(81)90047-8)
- Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S., Mager, D.L., Feschotte, C., 2018. Ten things you should know about transposable elements. *Genome Biology* 19, 199. <https://doi.org/10.1186/s13059-018-1577-z>
- Boyes, J.W., 1973. CHROMOSOMES OF RHAGIONIDAE, STRATIOMYIDAE AND XYLOMYIDAE (DIPTERA). *Canadian Journal of Genetics and Cytology*. <https://doi.org/10.1139/g73-028>
- Bradley, S.W., Sheppard, D.C., 1984. HOUSE FLY OVIPOSITION INHIBITION BY LARVAE 7.
- Brammer, C.A., von Dohlen, C.D., 2007. Evolutionary history of Stratiomyidae (Insecta: Diptera): The molecular phylogeny of a diverse family of flies. *Molecular Phylogenetics and Evolution* 43, 660–673. <https://doi.org/10.1016/j.ympev.2006.09.006>
- Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S.B., Krstic, P.S., Lindsay, S., Ling, X.S., Mastrangelo, C.H., Meller, A., Oliver, J.S., Pershin, Y.V., Ramsey, J.M., Riehn, R., Soni, G.V., Tabard-Cossa, V., Wanunu, M., Wiggin, M., Schloss, J.A., 2008. The potential and challenges of nanopore sequencing. *Nat Biotechnol* 26, 1146–1153. <https://doi.org/10.1038/nbt.1495>



- Britt, B.B., Dalla Vecchia, F.M., Chure, D.J., Engelmann, G.F., Whiting, M.F., Scheetz, R.D., 2018. *Caelestiventus hanseni* gen. et sp. nov. extends the desert-dwelling pterosaur record back 65 million years. *Nat Ecol Evol* 2, 1386–1392. <https://doi.org/10.1038/s41559-018-0627-y>
- Bruno, D., Bonelli, M., Cadamuro, A.G., Reguzzoni, M., Grimaldi, A., Casartelli, M., Tettamanti, G., 2019a. The digestive system of the adult *Hermetia illucens* (Diptera: Stratiomyidae): morphological features and functional properties. *Cell Tissue Res* 378, 221–238. <https://doi.org/10.1007/s00441-019-03025-7>
- Bruno, D., Bonelli, M., De Filippis, F., Di Lelio, I., Tettamanti, G., Casartelli, M., Ercolini, D., Caccia, S., 2019b. The Intestinal Microbiota of *Hermetia illucens* Larvae Is Affected by Diet and Shows a Diverse Composition in the Different Midgut Regions. *Appl Environ Microbiol* 85, e01864-18. <https://doi.org/10.1128/AEM.01864-18>
- Bull, R.M., 1976. The larval stages of the pasture and yellow soldier flies, *Inopus rubriceps* (Macq.) and *I. flavus* (James) (Diptera, Stratiomyidae) 6.

## C

- Cai, M., Li, L., Zhao, Z., Zhang, K., Li, F., Yu, C., Yuan, R., Zhou, B., Ren, Z., Yu, Z., Zhang, J., 2022. Morphometric Characteristic of Black Soldier Fly (*Hermetia illucens*) · Wuhan Strain and Its Egg Production Improved by Selectively Inbreeding. *Life* 12, 873. <https://doi.org/10.3390/life12060873>
- Calleri, D.V., McGrail Reid, E., Rosengaus, R.B., Vargo, E.L., Traniello, J.F.A., 2006. Inbreeding and disease resistance in a social insect: effects of heterozygosity on immunocompetence in the termite *Zootermopsis angusticollis*. *Proc Biol Sci* 273, 2633–2640. <https://doi.org/10.1098/rspb.2006.3622>
- Calvert, C.C., Morgan, N.O., Martin, R.D., 1970. House Fly Larvae: Biodegradation of Hen Excreta to Useful Products. *Poultry Science* 49, 588–589. <https://doi.org/10.3382/ps.0490588>
- Camenzuli, L., Van Dam, R., de Rijk, T., Andriessen, R., Van Schelt, J., Van der Fels-Klerx, H., 2018. Tolerance and Excretion of the Mycotoxins Aflatoxin B1, Zearalenone, Deoxynivalenol, and Ochratoxin A by *Alphitobius diaperinus* and *Hermetia illucens* from Contaminated Substrates. *Toxins* 10, 91. <https://doi.org/10.3390/toxins10020091>
- Caruso, D., Devic, E., Talamond, P., Subamia, I., Baras, E., 2020. 2014 Caruso et al., pdf technical+handbook BSF.
- Catts, E.P., Goff, M.L., 1992. *Forensic Entomology in Criminal Investigations* 20.
- CBD, 2010. What are Invasive Alien Species? [WWW Document]. URL <https://www.cbd.int/invasive/WhatareIAS.shtml> (accessed 8.10.22).
- Chen, S., Li, X., 2007. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol Biol* 7, 46. <https://doi.org/10.1186/1471-2148-7-46>
- Cock, P.J.A., Fields, C.J., Goto, N., Heuer, M.L., Rice, P.M., 2010. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res* 38, 1767–1771. <https://doi.org/10.1093/nar/gkp1137>
- Cohen, A., Duchemin, E., 2020. Fiche économique : élevage d’insectes comestibles Agri-Réseau | Documents [WWW Document]. Agri-Réseau. URL <https://www.agrireseau.net/documents/102715/fiche-economique-elevage-d-insectes-comestibles> (accessed 7.29.22).
- Cohen, K.M., Finney, S.C., Gibbard, P.L., Fan, J.-X., 2013. The ICS International Chronostratigraphic Chart. *Episodes* 36, 199–204. <https://doi.org/10.18814/epiiugs/2013/v36i3/002>
- COL [WWW Document], n.d. . COL. URL <https://www.catalogueoflife.org/> (accessed 7.28.22).

- Colantoni, A., Recchia, L., Bernabei, G., Cardarelli, M., Roupael, Y., Colla, G., 2017. Analyzing the Environmental Impact of Chemically-Produced Protein Hydrolysate from Leather Waste vs. Enzymatically-Produced Protein Hydrolysate from Legume Grains. *Agriculture* 7, 62. <https://doi.org/10.3390/agriculture7080062>
- Copello, A., 1926. *Biología de Hermetia illucens* Latr. (La mosca de nuestras colmenas). *Revista de la Sociedad Entomológica Argentina* 1.
- Craig Sheppard, D., Larry Newton, G., Thompson, S.A., Savage, S., 1994. A value added manure management system using the black soldier fly. *Bioresource Technology* 50, 275–279. [https://doi.org/10.1016/0960-8524\(94\)90102-3](https://doi.org/10.1016/0960-8524(94)90102-3)
- Cranston, P.S., Gullan, P.J., 2009. Phylogeny of Insects, in: *Encyclopedia of Insects*. Elsevier, pp. 780–793. <https://doi.org/10.1016/B978-0-12-374144-8.00208-3>
- Cullingham, C.I., Peery, R.M., Fortier, C.E., Mahon, E.L., Cooke, J.E.K., Coltman, D.W., 2020. Linking genotype to phenotype to identify genetic variation relating to host susceptibility in the mountain pine beetle system. *Evolutionary Applications* 13, 48–61. <https://doi.org/10.1111/eva.12773>

## D

- Dauphin, P., 2017. Présence de *Hermetia illucens* (LINNÉ, 1758) dans le sud- ouest de la France (Diptera Stratiomyidae) 3.
- De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C., 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- De Donato, M., Manrique, R., Ramirez, R., Mayer, L., Howell, C., 2005. Mass selection and inbreeding effects on a cultivated strain of *Panaeus* (*Litopenaeus*) *vannamei* in Venezuela. *Aquaculture, Genetics In Aquaculture VIII* 247, 159–167. <https://doi.org/10.1016/j.aquaculture.2005.02.005>
- Delahaye, C., Nicolas, J., 2021. Sequencing DNA with nanopores: Troubles and biases. *PLoS ONE* 16, e0257521. <https://doi.org/10.1371/journal.pone.0257521>
- Delsuc, F., Douzery, E.J.P., 2004. Les méthodes probabilistes en phylogénie moléculaire: (1) Les modèles d'évolution des séquences et le maximum de vraisemblance. *Biosystema* 22, 59–74.
- Demetriou, J., Kalaentzis, K., Kazilas, C., Kunz, G., Mostovski, M.B., Koutsoukos, E., 2022. An “alien” species on the loose: new records and updated distribution of the black soldier fly *Hermetia illucens* in the Western Palearctic 6.
- Desjardins, P., Conklin, D., 2010. NanoDrop microvolume quantitation of nucleic acids. *J Vis Exp* 2565. <https://doi.org/10.3791/2565>
- Di Somma, A., Moretta, A., Cané, C., Scieuzo, C., Salvia, R., Falabella, P., Duilio, A., 2021. Structural and Functional Characterization of a Novel Recombinant Antimicrobial Peptide from *Hermetia illucens*. *Curr Issues Mol Biol* 44, 1–13. <https://doi.org/10.3390/cimb44010001>
- Diener, S., Zurbrügg, C., Gutiérrez, F.R., Nguyen, D.H., Morel, A., Koottatep, T., Tockner, K., 2011. BLACK SOLDIER FLY LARVAE FOR ORGANIC WASTE TREATMENT – PROSPECTS AND CONSTRAINTS 8.
- Diener, S., Zurbrügg, C., Tockner, K., 2015. Bioaccumulation of heavy metals in the black soldier fly, *Hermetia illucens* and effects on its life cycle. *Journal of Insects as Food and Feed* 1, 261–270. <https://doi.org/10.3920/JIFF2015.0030>
- DNA Sequencing Costs: Data [WWW Document], 2021. . Genome.gov. URL <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data> (accessed 8.16.22).

Driscoll, C.A., Macdonald, D.W., O'Brien, S.J., 2009. From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences* 106, 9971–9978. <https://doi.org/10.1073/pnas.0901586106>

## E

Eggleton, P., 2020. The State of the World's Insects. *Annu. Rev. Environ. Resour.* 45, 61–82. <https://doi.org/10.1146/annurev-environ-012420-050035>

El-Kassaby, Y.A., Ritland, K., 1996. Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodivers Conserv* 5, 795–813. <https://doi.org/10.1007/BF00051787>

Erickson, M.C., Islam, M., Sheppard, C., Liao, J., Doyle, M.P., 2004a. Reduction of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Enteritidis in chicken manure by larvae of the black soldier fly. *J Food Prot* 67, 685–690. <https://doi.org/10.4315/0362-028x-67.4.685>

Erlich, Y., Mitra, P.P., delaBastide, M., McCombie, W.R., Hannon, G.J., 2008. Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nat Methods* 5, 679–682. <https://doi.org/10.1038/nmeth.1230>

European Commission, 2021. List of Invasive Alien Species of Union concern - Environment - European Commission [WWW Document]. URL [https://ec.europa.eu/environment/nature/invasivealien/list/index\\_en.htm](https://ec.europa.eu/environment/nature/invasivealien/list/index_en.htm) (accessed 8.10.22).

Ewald, N., Vidakovic, A., Langeland, M., Kiessling, A., Sampels, S., Lalander, C., 2020a. Fatty acid composition of black soldier fly larvae (*Hermetia illucens*) – Possibilities and limitations for modification through diet. *Waste Management* 102, 40–47. <https://doi.org/10.1016/j.wasman.2019.10.014>

## F

Facchini, E., Shrestha, K., van den Boer, E., Junes, P., Sader, G., Peeters, K., Schmitt, E., 2022. Long-Term Artificial Selection for Increased Larval Body Weight of *Hermetia illucens* in Industrial Settings. *Front. Genet.* 13, 865490. <https://doi.org/10.3389/fgene.2022.865490>

Frankham, R., 2005. Genetics and extinction. *Biological Conservation* 126, 131–140. <https://doi.org/10.1016/j.biocon.2005.05.002>

Frankham, R., 2003. Genetics and conservation biology. *Comptes Rendus Biologies* 326, 22–29. [https://doi.org/10.1016/S1631-0691\(03\)00023-4](https://doi.org/10.1016/S1631-0691(03)00023-4)

Franks, K., Kooienga, E., Sanders, M., Pendarvis, K., Yang, F., Tomberlin, J. k., Jordan, H. r., 2021. The effect of *Rhodococcus rhodochrous* supplementation on black soldier fly (Diptera: Stratiomyidae) development, nutrition, and waste conversion. *Journal of Insects as Food and Feed* 7, 397–408. <https://doi.org/10.3920/JIFF2020.0033>

Furman, P., Young, D., Paulcatts, E., 1959. *Hermetia illucens* (Linnaeus) as a Factor in the Natural Control of *Musca domestica* Linnaeus. *JOURNAL OF ECONOMIC ENTOMOLOGY* 5.

## G

Geneious, n.d. Geneious | Bioinformatics Software for Sequence Data Analysis [WWW Document]. Geneious. URL <https://www.geneious.com/> (accessed 8.10.22).

Generalovic, T.N., McCarthy, S.A., Warren, I.A., Wood, J.M.D., Torrance, J., Sims, Y., Quail, M., Howe, K., Pipan, M., Durbin, R., Jiggins, C.D., 2021a. A high-quality, chromosome-level genome assembly of the Black Soldier Fly (*Hermetia illucens* L.). *G3 (Bethesda)* 11, jkab085. <https://doi.org/10.1093/g3journal/jkab085>

- Gerard, P.J., Burton, G.P., 1983. Growth and development of Australian soldier fly, *Inopus rubriceps* (Diptera: Stratiomyidae), in New Zealand. *New Zealand Journal of Zoology* 10, 75–81. <https://doi.org/10.1080/03014223.1983.10423892>
- Giannetto, A., Oliva, S., Ceccon Lanes, C.F., de Araújo Pedron, F., Savastano, D., Baviera, C., Parrino, V., Lo Paro, G., Spanò, N.C., Cappello, T., Maisano, M., Mauceri, A., Fasulo, S., 2019. *Hermetia illucens* (Diptera: Stratiomyidae) larvae and prepupae: biomass production, fatty acid profile and expression of key genes involved in lipid metabolism. *Journal of Biotechnology*. <https://doi.org/10.1016/j.jbiotec.2019.10.015>
- Gogvadze, E., Buzdin, A., 2009. Retroelements and their impact on genome evolution and functioning. *Cell Mol Life Sci* 66, 3727–3742. <https://doi.org/10.1007/s00018-009-0107-2>
- Gold, M., Tomberlin, J.K., Diener, S., Zurbrügg, C., Mathys, A., 2018a. Decomposition of biowaste macronutrients, microbes, and chemicals in black soldier fly larval treatment: A review. *Waste Manag* 82, 302–318. <https://doi.org/10.1016/j.wasman.2018.10.022>
- González, J., Lenkov, K., Lipatov, M., Macpherson, J.M., Petrov, D.A., 2008. High Rate of Recent Transposable Element-Induced Adaptation in *Drosophila melanogaster*. *PLoS Biol* 6, e251. <https://doi.org/10.1371/journal.pbio.0060251>
- Goodier, J.L., Kazazian, H.H., 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135, 23–35. <https://doi.org/10.1016/j.cell.2008.09.022>
- Gorrens, E., De Smet, J., Vandeweyer, D., Bossaert, S., Crauwels, S., Lievens, B., Van Campenhout, L., 2022. The bacterial communities of black soldier fly larvae (*Hermetia illucens*) during consecutive, industrial rearing cycles. *Journal of Insects as Food and Feed* 1–16. <https://doi.org/10.3920/IIFF2021.0150>
- Goubert, C., Modolo, L., Vieira, C., ValienteMoro, C., Mavingui, P., Boulesteix, M., 2015. De Novo Assembly and Annotation of the Asian Tiger Mosquito (*Aedes albopictus*) Repeatome with dnaPipeTE from Raw Genomic Reads and Comparative Analysis with the Yellow Fever Mosquito (*Aedes aegypti*). *Genome Biology and Evolution* 7, 1192–1205. <https://doi.org/10.1093/gbe/evw050>
- Grimaldi, D.A., Engel, M.S., 2005. *Evolution of the insects*. Cambridge University Press, Cambridge [U.K.]; New York.
- Gu, P., Liu, W., Yao, Y., Ni, Q., Zhang, M., Li, D., Xu, H., 2016. Evidence of adaptive evolution of alpine pheasants to high-altitude environment from mitogenomic perspective. *Mitochondrial DNA Part A* 27, 455–462. <https://doi.org/10.3109/19401736.2014.900667>
- Guarnieri, A., Triunfo, M., Scieuzo, C., Ianniciello, D., Tafi, E., Hahn, T., Zibek, S., Salvia, R., De Bonis, A., Falabella, P., 2022. Antimicrobial properties of chitosan from different developmental stages of the bioconverter insect *Hermetia illucens*. *Sci Rep* 12, 8084. <https://doi.org/10.1038/s41598-022-12150-3>
- Guilliet, J., Baudouin, G., Pollet, N., Filée, J., 2022. What complete mitochondrial genomes tell us about the evolutionary history of the black soldier fly, *Hermetia illucens*. *BMC Ecol Evol* 22, 72. <https://doi.org/10.1186/s12862-022-02025-6>
- Guiné, R.P.F., Correia, P., Coelho, C., Costa, C.A., 2021. The role of edible insects to mitigate challenges for sustainability. *Open Agriculture* 6, 24–36. <https://doi.org/10.1515/opag-2020-0206>
- Gunawan, A., 2022. The effect of live maggot (*Hermetia illucens*) supplementation on organoleptic quality and yolk color of duck egg. *IOP Conf. Ser.: Earth Environ. Sci.* 1001, 012035. <https://doi.org/10.1088/1755-1315/1001/1/012035>

## H

- Hale, O.M., 1973. Dried *Hermetia illucens* larvae (Diptera: Stratiomyidae) as a feed additive for poultry. undefined.
- Hall, D.W., 2008. Popularity of Insects, in: Capinera, J.L. (Ed.), *Encyclopedia of Entomology*. Springer Netherlands, Dordrecht, pp. 2999–3006. [https://doi.org/10.1007/978-1-4020-6359-6\\_3070](https://doi.org/10.1007/978-1-4020-6359-6_3070)
- Hall, R.D., Gerhardt, R.R., 2002. TAXONOMY 129 MORPHOLOGY 130 LIFE HISTORY 132 BEHAVIOR AND ECOLOGY 133 FAMILIES OF MINOR MEDICAL OR VETERINARY INTEREST 134 PUBLIC HEALTH IMPORTANCE 141 VETERINARY IMPORTANCE 142 PREVENTION AND CONTROL 142 REFERENCES AND FURTHER READING 143 19.
- Hansen, K.D., Brenner, S.E., Dudoit, S., 2010. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res* 38, e131. <https://doi.org/10.1093/nar/gkq224>
- Hari, R., Parthasarathy, S., 2019. Next Generation Sequencing Data Analysis, in: Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C. (Eds.), *Encyclopedia of Bioinformatics and Computational Biology*. Academic Press, Oxford, pp. 157–163. <https://doi.org/10.1016/B978-0-12-809633-8.20093-9>
- Hauser, M., Woodley, M., 2015. The historical spread of the Black Soldier Fly, *Hermetia illucens* (L.) (Diptera, Stratiomyidae, Hermetiinae), and its establishment in Canada. *Journal of the Kansas Entomological Society* 146, 51–54.
- Hedges, S.B., Marin, J., Suleski, M., Paymer, M., Kumar, S., 2015. Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* 32, 835–845. <https://doi.org/10.1093/molbev/msv037>
- Hernandez, R.D., Uricchio, L.H., Hartman, K., Ye, C., Dahl, A., Zaitlen, N., 2019. Ultrarare variants drive substantial cis heritability of human gene expression. *Nat Genet* 51, 1349–1355. <https://doi.org/10.1038/s41588-019-0487-7>
- Heuel, M., Sandrock, C., Leiber, F., Mathys, A., Gold, M., Zurbrüegg, C., Gangnat, I.D.M., Kreuzer, M., Terranova, M., 2022. Black soldier fly larvae meal and fat as a replacement for soybeans in organic broiler diets: effects on performance, body N retention, carcass and meat quality. *British Poultry Science* 1–12. <https://doi.org/10.1080/00071668.2022.2053067>
- Heuel, M., Sandrock, C., Leiber, F., Mathys, A., Gold, M., Zurbrüegg, C., Gangnat, I.D.M., Kreuzer, M., Terranova, M., 2021. Black soldier fly larvae meal and fat can completely replace soybean cake and oil in diets for laying hens. *Poult Sci* 100, 101034. <https://doi.org/10.1016/j.psj.2021.101034>
- Hoc, B., Genva, M., Fauconnier, M.-L., Lognay, G., Francis, F., Caparros Megido, R., 2020. About lipid metabolism in *Hermetia illucens* (L. 1758): on the origin of fatty acids in prepupae. *Sci Rep* 10, 11916. <https://doi.org/10.1038/s41598-020-68784-8>
- Hof, A.E. van't, Campagne, P., Rigden, D.J., Yung, C.J., Lingley, J., Quail, M.A., Hall, N., Darby, A.C., Saccheri, I.J., 2016. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* 534, 102–105. <https://doi.org/10.1038/nature17951>
- Hoffmann, L., 2021. Genetic diversity and mating systems in a mass-reared black soldier fly (*Hermetia illucens*) population (Thesis). Stellenbosch : Stellenbosch University.
- Hoffmann, L., Hull, K.L., Bierman, A., Badenhorst, R., Bester-van der Merwe, A.E., Rhode, C., 2021a. Patterns of Genetic Diversity and Mating Systems in a Mass-Reared Black Soldier Fly Colony. *Insects* 12, 480. <https://doi.org/10.3390/insects12060480>
- Hoorweg, D., Bhada-Tata, P., 2012. *What a Waste : A Global Review of Solid Waste Management*. World Bank, Washington, DC.
- Huis, A. van, 2013. *Edible insects: future prospects for food and feed security*, FAO forestry paper. Food and Agriculture Organization of the United Nations, Rome.

## I

- Ikegawa, S., 2012. A Short History of the Genome-Wide Association Study: Where We Were and Where We Are Going. *Genomics Inform* 10, 220–225. <https://doi.org/10.5808/GI.2012.10.4.220>
- IndexMundi - Country Facts [WWW Document], n.d. URL <https://www.indexmundi.com/> (accessed 9.6.22).
- IPIFF, 2019. Promoting Insects for Human Consumption & Animal Feed – IPIFF [WWW Document]. International Platform of Insects for Food and Feed, Brussels. URL <https://ipiff.org/> (accessed 7.29.22).
- Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., Paszkowski, J., 2011. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472, 115–119. <https://doi.org/10.1038/nature09861>

## J

- James, M.T., 1936. The Stratiomyidae of Colorado and Utah. *Journal of the Kansas Entomological Society* 9, 33–36.
- Jiang, L., Schlesinger, F., Davis, C.A., Zhang, Y., Li, R., Salit, M., Gingeras, T.R., Oliver, B., 2011. Synthetic spike-in standards for RNA-seq experiments. *Genome Res* 21, 1543–1551. <https://doi.org/10.1101/gr.121095.111>
- Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S.R., McCouch, S.R., Wessler, S.R., 2003. An active DNA transposon family in rice. *Nature* 421, 163–167. <https://doi.org/10.1038/nature01214>
- Jones, B.M., Tomberlin, J.K., 2021. Effects of adult body size on mating success of the black soldier fly, *Hermetia illucens* (L.) (Diptera: Stratiomyidae). *Journal of Insects as Food and Feed* 7, 5–20. <https://doi.org/10.3920/JIFF2020.0001>
- Jones, B.M., Tomberlin, J.K., 2020. Validation of Acrylic Paint as a Marking Technique for Examining Mating Success of the Black Soldier Fly (Diptera: Stratiomyidae). *Journal of Economic Entomology* 113, 2128–2133. <https://doi.org/10.1093/jee/toaa129>
- Jordan, H.R., Tomberlin, J.K., 2021. Microbial influence on reproduction, conversion, and growth of mass produced insects. *Current Opinion in Insect Science, NeuroScience*. Special Section on Insects as food and feed 48, 57–63. <https://doi.org/10.1016/j.cois.2021.10.001>
- Jurka, J., Kapitonov, V.V., Kohany, O., Jurka, M.V., 2007. Repetitive sequences in complex genomes: structure and evolution. *Annu Rev Genomics Hum Genet* 8, 241–259. <https://doi.org/10.1146/annurev.genom.8.080706.092416>

## K

- Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E., Schulman, A.H., 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences* 97, 6603–6607. <https://doi.org/10.1073/pnas.110587497>
- Kaneshiro, K.Y., Boake, C.R.B., 1987. Sexual selection and speciation: Issues raised by Hawaiian *Drosophila*. *Trends in Ecology & Evolution* 2, 207–212. [https://doi.org/10.1016/0169-5347\(87\)90022-X](https://doi.org/10.1016/0169-5347(87)90022-X)
- Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., Soukal, P., Hroudová, M., Doležal, P., Stairs, C.W., Roger, A.J., Eliáš, M., Dacks, J.B., Vlček, Č., Hampl, V., 2016. A Eukaryote without a Mitochondrial Organelle. *Current Biology* 26, 1274–1284. <https://doi.org/10.1016/j.cub.2016.03.053>



- Katakura, H., Saitoh, S., Aoki, M., 1996. Sexual isolation between three forms of flightless *Chrysolina* leaf beetles (Coleoptera: Chrysomelidae) parapatrically distributed in the vicinity of Sapporo, Hokkaido, northern Japan. *Genes & Genetic Systems* 71, 139–144. <https://doi.org/10.1266/ggs.71.139>
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kawahara, A.Y., Breinholt, J.W., 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proc Biol Sci* 281, 20140970. <https://doi.org/10.1098/rspb.2014.0970>
- Kaya, C., Generalovic, T.N., Ståhls, G., Hauser, M., Samayoa, A.C., Nunes-Silva, C.G., Roxburgh, H., Wohlfahrt, J., Ewusie, E.A., Kenis, M., Hanboonsong, Y., Orozco, J., Carrejo, N., Nakamura, S., Gasco, L., Rojo, S., Tanga, C.M., Meier, R., Rhode, C., Picard, C.J., Jiggins, C.D., Leiber, F., Tomberlin, J.K., Hasselmann, M., Blanckenhorn, W.U., Kapun, M., Sandrock, C., 2021. Global population genetic structure and demographic trajectories of the black soldier fly, *Hermetia illucens*. *BMC Biology* 19, 94. <https://doi.org/10.1186/s12915-021-01029-w>
- Kaza, S., Yao, L.C., Bhada-Tata, P., Van Woerden, F., 2018. What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. World Bank, Washington, DC. <https://doi.org/10.1596/978-1-4648-1329-0>
- Kelley, J.L., Peyton, J.T., Fiston-Lavier, A.-S., Teets, N.M., Yee, M.-C., Johnston, J.S., Bustamante, C.D., Lee, R.E., Denlinger, D.L., 2014. Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nat Commun* 5, 4611. <https://doi.org/10.1038/ncomms5611>
- Kenis, M., Wermelinger, B., Grégoire, J.-C., 2004. Research on Parasitoids and Predators of Scolytidae – A Review, in: *Bark and Wood Boring Insects in Living Trees in Europe, A Synthesis*. pp. 237–290. [https://doi.org/10.1007/978-1-4020-2241-8\\_11](https://doi.org/10.1007/978-1-4020-2241-8_11)
- Khamis, F.M., Ombura, F.L.O., Akutse, K.S., Subramanian, S., Mohamed, S.A., Fiaboe, K.K.M., Saijuntha, W., Van Loon, J.J.A., Dicke, M., Dubois, T., Ekesi, S., Tanga, C.M., 2020. Insights in the Global Genetics and Gut Microbiome of Black Soldier Fly, *Hermetia illucens*: Implications for Animal Feed Safety Control. *Front. Microbiol.* 11. <https://doi.org/10.3389/fmicb.2020.01538>
- Kidwell, M.G., 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115, 49–63. <https://doi.org/10.1023/A:1016072014259>
- Kim, W., Bae, S., Park, H., Park, K., Lee, S., Choi, Y., Han, S., Koh, Y., 2010. The Larval Age and Mouth Morphology of the Black Soldier Fly, *Hermetia illucens* (Diptera: Stratiomyidae) 3.
- Kim, W., Bae, S., Park, K., Lee, S., Choi, Y., Han, S., Koh, Y., 2011. Biochemical characterization of digestive enzymes in the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae). *Journal of Asia-Pacific Entomology* 14, 11–14. <https://doi.org/10.1016/j.aspen.2010.11.003>
- Kim, Y.B., Oh, J.H., McIver, L.J., Rashkovetsky, E., Michalak, K., Garner, H.R., Kang, L., Nevo, E., Korol, A.B., Michalak, P., 2014. Divergence of *Drosophila melanogaster* repeatomes in response to a sharp microclimate contrast in Evolution Canyon, Israel. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10630–10635. <https://doi.org/10.1073/pnas.1410372111>
- Kimura, M., 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press.
- Kishawy, A.T.Y., Mohammed, H.A., Zagloul, A.W., Attia, M.S., Hassan, F.A.M., Roushdy, E.M., Ismail, T.A., Ibrahim, D., 2022. Partial defatted black soldier larvae meal as a promising strategy to replace fish meal protein in diet for Nile tilapia (*Oreochromis niloticus*): Performance, expression of protein and fat transporters, and cytokines related genes and economic efficiency. *Aquaculture* 555, 738195. <https://doi.org/10.1016/j.aquaculture.2022.738195>
- Klammsteiner, T., Walter, A., Bogataj, T., Heussler, C.D., Stres, B., Steiner, F.M., Schlick-Steiner, B.C., Arthofer, W., Insam, H., 2020. The Core Gut Microbiome of Black Soldier Fly (*Hermetia illucens*) Larvae Raised on Low-Bioburden Diets. *Front. Microbiol.* 11, 993. <https://doi.org/10.3389/fmicb.2020.00993>

- Klein, nicolas, 2022. La sélection génétique est une menace pour le bien-être animal : l'exemple de la race bovine Blanc bleu belge. Welfarm. URL <https://welfarm.fr/vaches-blanc-bleu-belge-selection-genetique/> (accessed 9.21.22).
- Kongyeli Badii, B., 2020. Phylogeny and Functional Morphology of Diptera (Flies), in: Sarwar, M. (Ed.), Life Cycle and Development of Diptera. IntechOpen. <https://doi.org/10.5772/intechopen.90421>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35, 1547–1549. <https://doi.org/10.1093/molbev/msy096>

## L

- Lalander, C., Diener, S., Magri, M.E., Zurbrügg, C., Lindström, A., Vinnerås, B., 2013. Faecal sludge management with the larvae of the black soldier fly (*Hermetia illucens*) — From a hygiene aspect. Science of The Total Environment 458–460, 312–318. <https://doi.org/10.1016/j.scitotenv.2013.04.033>
- Lalander, C., Vinnerås, B., 2022. Actions needed before insects can contribute to a real closed-loop circular economy in the EU. Journal of Insects as Food and Feed 8, 337–342. <https://doi.org/10.3920/JIFF2022.x003>
- Larson, B.M.H., Kevan, P.G., Inouye, D.W., 2001. Flies and flowers: taxonomic diversity of anthophiles and pollinators. Can Entomol 133, 439–465. <https://doi.org/10.4039/Ent133439-4>
- Latreille, P.A., Buffon, G.L.L., Sonnini, C.S., Sève, J.E. de, 1802. Histoire naturelle, générale et particulière des crustacés et des insectes. F. Dufart, Paris. <https://doi.org/10.5962/bhl.title.15764>
- Leclercq, M., 1969. Dispersion and transport of noxious insects: *Hermetia illucens* L.(Diptera Stratiomyidae). Bulletin des recherches agronomiques de Gembloux.
- Lee, H.L., Cheong, W.H., 1982. Larva of *Hermetia* species from a dead human female. Southeast Asian J Trop Med Public Health 13, 289–290.
- Lee, Y.C.G., Langley, C.H., 2010. Transposable elements in natural populations of *Drosophila melanogaster*. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 1219–1228. <https://doi.org/10.1098/rstb.2009.0318>
- Lee, Y.H., Kim, S.C., Nam, K.D., Kim, T.H., Jung, B.O., Park, Y.-I., Synytsya, A., Park, J.K., 2022. Chitosan isolated from black soldier flies *Hermetia illucens*: Structure and enzymatic hydrolysis. Process Biochemistry 118, 171–181. <https://doi.org/10.1016/j.procbio.2022.04.020>
- Lei, X.J., Kim, T.H., Park, J.H., Kim, I.H., 2019. Evaluation of Supplementation of Defatted Black Soldier Fly (*Hermetia illucens*) Larvae Meal in Beagle Dogs. Annals of Animal Science 19, 767–777. <https://doi.org/10.2478/aoas-2019-0021>
- Leigh, J.W., Bryant, D., 2015. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6, 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Leinonen, R., Sugawara, H., Shumway, M., International Nucleotide Sequence Database Collaboration, 2011. The sequence read archive. Nucleic Acids Res 39, D19–21. <https://doi.org/10.1093/nar/gkq1019>
- Leni, G., Soetemans, L., Jacobs, J., Depraetere, S., Gianotten, N., Bastiaens, L., Caligiani, A., Sforza, S., 2020. Protein hydrolysates from *Alphitobius diaperinus* and *Hermetia illucens* larvae treated with commercial proteases. Journal of Insects as Food and Feed 6, 393–404. <https://doi.org/10.3920/JIFF2019.0037>
- Leong, S.Y., Kutty, S.R.M., Malakahmad, A., Tan, C.K., 2016. Feasibility study of biodiesel production using lipids of *Hermetia illucens* larva fed with organic waste. Waste Manag 47, 84–90. <https://doi.org/10.1016/j.wasman.2015.03.030>



- Les Echos, 2017. Elevage d'insectes : les promesses d'une nouvelle filière industrielle [WWW Document]. Les Echos. URL <https://www.lesechos.fr/2017/07/elevage-dinsectes-les-promesses-dune-nouvelle-filiere-industrielle-175269> (accessed 7.29.22).
- Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* 49, W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- Li, Q., Zheng, L., Cai, H., Garza, E., Yu, Z., Zhou, S., 2011. From organic waste to biodiesel: Black soldier fly, *Hermetia illucens*, makes it feasible. *Fuel* 90, 1545–1548. <https://doi.org/10.1016/j.fuel.2010.11.016>
- Li, X., Qin, C., Fang, Z., Sun, X., Shi, H., Wang, Q., Zhao, H., 2022. Replacing Dietary Fish Meal with Defatted Black Soldier Fly (*Hermetia illucens*) Larvae Meal Affected Growth, Digestive Physiology and Muscle Quality of Tongue Sole (*Cynoglossus semilaevis*). *Front. Physiol.* 13, 855957. <https://doi.org/10.3389/fphys.2022.855957>
- Li, X.-D., Jiang, G.-F., Yan, L.-Y., Li, R., Mu, Y., Deng, W.-A., 2018. Positive Selection Drove the Adaptation of Mitochondrial Genes to the Demands of Flight and High-Altitude Environments in Grasshoppers. *Frontiers in Genetics* 9, 605. <https://doi.org/10.3389/fgene.2018.00605>
- Li, Z., Chen, Y., Mu, D., Yuan, J., Shi, Y., Zhang, H., Gan, J., Li, N., Hu, X., Liu, B., Yang, B., Fan, W., 2012. Comparison of the two major classes of assembly algorithms: overlap-layout-consensus and de-bruijn-graph. *Briefings in Functional Genomics* 11, 25–37. <https://doi.org/10.1093/bfgp/elr035>
- Lievens, S., Poma, G., De Smet, J., Van Campenhout, L., Covaci, A., Van Der Borght, M., 2021. Chemical safety of black soldier fly larvae (*Hermetia illucens*), knowledge gaps and recommendations for future research: a critical review. *Journal of Insects as Food and Feed* 7, 383–396. <https://doi.org/10.3920/JIFF2020.0081>
- Lindner, E., 1936. Die amerikanische *Hermetia illucens* L. im Mittelmeergebiet: (Stratiomyidae, Dipt.). Akadem. Verlagsges.
- Liu, C., Wang, C., Yao, H., 2019. Comprehensive Resource Utilization of Waste Using the Black Soldier Fly (*Hermetia illucens* (L.)) (Diptera: Stratiomyidae). *Animals* 9, 349. <https://doi.org/10.3390/ani9060349>
- Liu, Q., Tomberlin, J.K., Brady, J.A., Sanford, M.R., Yu, Z., 2008a. Black soldier fly (Diptera: Stratiomyidae) larvae reduce *Escherichia coli* in dairy manure. *Environ Entomol* 37, 1525–1530. <https://doi.org/10.1603/0046-225x-37.6.1525>
- Liu, Y., Cai, C., Huang, D., 2017. First soldier fly from the mid-Cretaceous Burmese amber (Diptera, Stratiomyidae). *Cretaceous Research* 70, 142–146. <https://doi.org/10.1016/j.cretres.2016.10.007>
- Lodish, M.H., Berk, M.A., Matsudaira, P., Bretscher, M.A., Krieger, M.M., Scott, M.M.P., Zipursky, S.L., Darnell, J., 2005. *Biologie moléculaire de la cellule*. De Boeck Supérieur.
- Lopez, P., Casane, D., Philippe, H., 2002. Phylogénie et évolution moléculaires - Bio-informatique (5). *Med Sci (Paris)* 18, 1146–1154. <https://doi.org/10.1051/medsci/200218111146>
- Lord, W. D., Goff, M.L., Adkins, T.R., Haskell, N.H., 1994. The Black Soldier Fly *Hermetia illucens* (Diptera: Stratiomyidae) As a Potential Measure of Human Postmortem Interval: Observations and Case Histories. *JFS* 39, 215–222. <https://doi.org/10.1520/JFS135871>
- Lu, J., Guo, Y., Muhmood, A., Zeng, B., Qiu, Y., Wang, P., Ren, L., 2022. Probing the antioxidant activity of functional proteins and bioactive peptides in *Hermetia illucens* larvae fed with food wastes. *Sci Rep* 12, 2799. <https://doi.org/10.1038/s41598-022-06668-9>

Luo, Y., Yang, X., Gao, Y., 2013. Mitochondrial DNA response to high altitude: A new perspective on high-altitude adaptation. *Mitochondrial DNA* 24, 313–319. <https://doi.org/10.3109/19401736.2012.760558>

Lupi, D., Savoldelli, S., Leonardi, M.G., Jucker, C., 2019. Feeding in the adult of *Hermetia illucens* (Diptera Stratiomyidae): reality or fiction? *J Entomol Acarol Res* 51. <https://doi.org/10.4081/jear.2019.8046>

## M

Mackillop, S., Keitel, C., Latty, T., 2022. Beyond *Hermetia illucens*: An investigation of the garden soldier fly *Exaireta spinigera* as a potential bioconverter of food waste. *Journal of Applied Entomology* 146, 168–174. <https://doi.org/10.1111/jen.12936>

Maddison, D.R., Swofford, D.L., Maddison, W.P., 1997. NEXUS: an extensible file format for systematic information. *Syst Biol* 46, 590–621. <https://doi.org/10.1093/sysbio/46.4.590>

MAGNE, M.-A., NOZIÈRES-PETIT, M.-O., COURNUT, S., OLLION, É., PUILLET, L., RENAUDEAU, D., FORTUN-LAMOTHE, L., 2019. Gérer la diversité animale dans les systèmes d'élevage : laquelle, comment et pour quels bénéfices ? *INRAE Productions Animales* 32, 263–280. <https://doi.org/10.20870/productions-animales.2019.32.2.2496>

Maquart, P.-O., Denis, R., Willems, J., 2020. First record of the Black Soldier Fly, *Hermetia illucens*, in the Western regions of France (Vendée, Loire-Atlantique, Ille-et-Vilaine) with notes on its worldwide repartition (Diptera, Stratiomyidae). *Bulletin de la Société entomologique de France* 125, 13–18. [https://doi.org/10.32475/bsef\\_2104](https://doi.org/10.32475/bsef_2104)

Martínez-Sánchez, A., Magaña, C., Saloña, M., Rojo, S., 2011. First record of *Hermetia illucens* (Diptera: Stratiomyidae) on human corpses in Iberian Peninsula. *Forensic Science International* 206, e76–e78. <https://doi.org/10.1016/j.forsciint.2010.10.021>

Martín-Forés, I., Martín-López, B., Montes, C., 2013. Anthropomorphic Factors Influencing Spanish Conservation Policies of Vertebrates. *International Journal of Biodiversity* 2013, e142670. <https://doi.org/10.1155/2013/142670>

Mason, F., Rozkošný, R., Hauser, M., 2009. A review of the soldier flies (Diptera: Stratiomyidae) of Sardinia. *Zootaxa* 2318, 507–530. <https://doi.org/10.11646/zootaxa.2318.1.20>

May, B.M., 1961. The occurrence in New Zealand and the life-history of the soldier fly *Hermetia illucens* (L.) (Diptera: Stratiomyidae). *New Zealand Journal of Science Wellington* 4, 55–65.

McDonald, J.F., Matzke, M.A., Matzke, A.J., 2005. Host defenses to transposable elements and the evolution of genomic imprinting. *Cytogenet Genome Res* 110, 242–249. <https://doi.org/10.1159/000084958>

McLaughlin, P., Lemaitre, R., 1997. Carcinization in the Anomura - fact or fiction? *Contributions to Zoology* 45.

Meneguz, M., Schiavone, A., Gai, F., Dama, A., Lussiana, C., Renna, M., Gasco, L., 2018. Effect of rearing substrate on growth performance, waste reduction efficiency and chemical composition of black soldier fly (*Hermetia illucens*) larvae: Rearing substrate effects on performance and nutritional composition of black soldier fly. *J. Sci. Food Agric.* 98, 5776–5784. <https://doi.org/10.1002/jsfa.9127>

Meng, G., Li, Y., Yang, C., Liu, S., 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research* 47, e63. <https://doi.org/10.1093/nar/gkz173>

Merdith, A.S., Williams, S.E., Collins, A.S., Tetley, M.G., Mulder, J.A., Blades, M.L., Young, A., Armistead, S.E., Cannon, J., Zahirovic, S., Müller, R.D., 2021. Extending full-plate tectonic models into deep time: Linking the Neoproterozoic and the Phanerozoic. *Earth-Science Reviews* 214, 103477. <https://doi.org/10.1016/j.earscirev.2020.103477>

- Mertenat, A., Diener, S., Zurbrügg, C., 2019. Black Soldier Fly biowaste treatment – Assessment of global warming potential. *Waste Management* 84, 173–181. <https://doi.org/10.1016/j.wasman.2018.11.040>
- meticulous research, 2022. Edible Insects Market by Size, Share, Forecasts, & Trends Analysis | Meticulous Research® [WWW Document]. URL <https://www.meticulousresearch.com/product/edible-insects-market-5156#description> (accessed 7.29.22).
- Mikołajczak, Z., Rawski, M., Mazurkiewicz, J., Kierończyk, B., Kołodziejski, P., Pruszyńska-Oszmátek, E., Józefiak, D., 2022. The first insight into black soldier fly meal in brown trout nutrition as an environmentally sustainable fish meal replacement. *animal* 16, 100516. <https://doi.org/10.1016/j.animal.2022.100516>
- Mohamed, H., Marusich, E., Afanasev, Y., Leonov, S., 2022. Bacterial Outer Membrane Permeability Increase Underlies the Bactericidal Effect of Fatty Acids From *Hermetia illucens* (Black Soldier Fly) Larvae Fat Against Hypermucoviscous Isolates of *Klebsiella pneumoniae*. *Front Microbiol* 13, 844811. <https://doi.org/10.3389/fmicb.2022.844811>
- Montagna, M., Tong, K.J., Magoga, G., Strada, L., Tintori, A., Ho, S.Y.W., Lo, N., 2019. Recalibration of the insect evolutionary time scale using Monte San Giorgio fossils suggests survival of key lineages through the End-Permian Extinction. *Proc. R. Soc. B* 286, 20191854. <https://doi.org/10.1098/rspb.2019.1854>
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How Many Species Are There on Earth and in the Ocean? *PLoS Biol* 9, e1001127. <https://doi.org/10.1371/journal.pbio.1001127>
- Morneau, D., 2021. Pan-genomes: moving beyond the reference. *Nature Research*. <https://doi.org/10.1038/d42859-020-00115-3>

## N

- Nakatsuji, T., Kao, M.C., Fang, J.-Y., Zouboulis, C.C., Zhang, L., Gallo, R.L., Huang, C.-M., 2009. Antimicrobial Property of Lauric Acid Against *Propionibacterium Acnes*: Its Therapeutic Potential for Inflammatory *Acne Vulgaris*. *Journal of Investigative Dermatology* 129, 2480–2488. <https://doi.org/10.1038/jid.2009.93>
- Nellåker, C., Keane, T.M., Yalcin, B., Wong, K., Agam, A., Belgard, T.G., Flint, J., Adams, D.J., Frankel, W.N., Ponting, C.P., 2012. The genomic landscape shaped by selection on transposable elements across 18 mouse strains. *Genome Biol* 13, R45. <https://doi.org/10.1186/gb-2012-13-6-r45>
- Neumann, C., Velten, S., Liebert, F., 2018. N Balance Studies Emphasize the Superior Protein Quality of Pig Diets at High Inclusion Level of Algae Meal (*Spirulina platensis*) or Insect Meal (*Hermetia illucens*) when Adequate Amino Acid Supplementation Is Ensured. *Animals (Basel)* 8. <https://doi.org/10.3390/ani8100172>
- Newton, G.L., Booram, C.V., Barker, R.W., Hale, O.M., 1977. Dried *Hermetia Illucens* Larvae Meal as a Supplement for Swine. *Journal of Animal Science* 44, 395–400. <https://doi.org/10.2527/jas1977.443395x>
- Newton, G.L., Sheppard, D.C., Burtle, G.J., 2005. Research Briefs: Black Soldier Fly Prepupae - A Compelling Alternative to Fish Meal and Fish Oil 3.
- Ng, P.C., Kirkness, E.F., 2010. Whole Genome Sequencing, in: Barnes, M.R., Breen, G. (Eds.), *Genetic Variation: Methods and Protocols, Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp. 215–226. [https://doi.org/10.1007/978-1-60327-367-1\\_12](https://doi.org/10.1007/978-1-60327-367-1_12)
- Nixon, K.C., 2001. Phylogeny, in: Levin, S.A. (Ed.), *Encyclopedia of Biodiversity (Second Edition)*. Academic Press, Waltham, pp. 16–23. <https://doi.org/10.1016/B978-0-12-384719-5.00108-8>
- Noerhatini, P., Al Hakim, A.D., Setiawati, Y., Azhari, S.K., Munaf, D.R., Azka, R.M., 2022. Business Model Based on Coffee Husk Bioconversion by Black Soldier Fly (*Hermetia Illucens*) and Agroforestry System in Bandung Regency, West Java: Presented at the International Conference on Economics, Management and Accounting (ICEMAC 2021), Sukabumi, Indonesia. <https://doi.org/10.2991/aebmr.k.220204.031>

Notter, D.R., 1999. The importance of genetic diversity in livestock populations of the future. *Journal of Animal Science* 77, 61–69. <https://doi.org/10.2527/1999.77161x>

Nyakeri, E.M., Ogola, H.J., Ayieko, M.A., Amimo, F.A., 2017. An open system for farming black soldier fly larvae as a source of proteins for smallscale poultry and fish production. *Journal of Insects as Food and Feed* 3, 51–56. <https://doi.org/10.3920/JIFF2016.0030>

## O

O’Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., de Queiroz, A., Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M.O., Aguilera, O., Aubry, M.-P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J., Jackson, J.B.C., 2016. Formation of the Isthmus of Panama. *Sci Adv* 2, e1600883. <https://doi.org/10.1126/sciadv.1600883>

OGM : le cadre réglementaire [WWW Document], 2018. . Ministère de l’Agriculture et de la Souveraineté alimentaire. URL <https://agriculture.gouv.fr/ogm-le-cadre-reglementaire> (accessed 8.17.22).

Oldroyd, B.P., 2012. Domestication of honey bees was associated with expansion of genetic diversity. *Molecular Ecology* 21, 4409–4411. <https://doi.org/10.1111/j.1365-294X.2012.05641.x>

Oliveira, F., Doelle, K., List, R., O’Reilly, J.R., 2015. Assessment of Diptera: Stratiomyidae, genus *Hermetia illucens* (L., 1758) using electron microscopy. *Journal of Entomology and Zoology Studies* 6.

Oliveira, F., Doelle, K., Smith, R., 2016. External Morphology of *Hermetia illucens* Stratiomyidae: Diptera (L.1758) Based on Electron Microscopy. *ARRB* 9, 1–10. <https://doi.org/10.9734/ARRB/2016/22973>

Ooninx, D.G.A.B., de Boer, I.J.M., 2012. Environmental Impact of the Production of Mealworms as a Protein Source for Humans – A Life Cycle Assessment. *PLoS ONE* 7, e51145. <https://doi.org/10.1371/journal.pone.0051145>

Ooninx, D.G.A.B., Volk, N., Diehl, J.J.E., van Loon, J.J.A., Belušič, G., 2016. Photoreceptor spectral sensitivity of the compound eyes of black soldier fly (*Hermetia illucens*) informing the design of LED-based illumination to enhance indoor reproduction. *Journal of Insect Physiology* 95, 133–139. <https://doi.org/10.1016/j.jinsphys.2016.10.006>

Outram, A.K., Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., Evershed, R.P., 2009. The Earliest Horse Harnessing and Milking. *Science* 323, 1332–1335. <https://doi.org/10.1126/science.1168594>

## P

Palma, L., Ceballos, S.J., Johnson, P.C., Niemeier, D., Pitesky, M., VanderGheynst, J.S., 2018. Cultivation of black soldier fly larvae on almond byproducts: impacts of aeration and moisture on larvae growth and composition. *J. Sci. Food Agric.* 98, 5893–5900. <https://doi.org/10.1002/jsfa.9252>

Paradis, E., 2018. Analysis of haplotype networks: The randomized minimum spanning tree method. *Methods in Ecology and Evolution* 9, 1308–1317. <https://doi.org/10.1111/2041-210X.12969>

Park, J.-Y., Jung, S., Na, Y.-G., Jeon, C.-H., Cheon, H.-Y., Yun, E.-Y., Lee, S., Kwon, E.E., Kim, J.-K., 2021. Biodiesel production from the black soldier fly larvae grown on food waste and its fuel property characterization as a potential transportation fuel. *Environmental Engineering Research* 27, 200704–0. <https://doi.org/10.4491/eer.2020.704>

Park, S., Choi, H., Choi, J., Jeong, G., 2017. Population Structure of the Exotic Black Soldier Fly, *Hermetia illucens* (Diptera: Stratiomyidae) in Korea. *Korean J. Environ. Ecol.* 31, 520–528. <https://doi.org/10.13047/KJEE.2017.31.6.520>

- Park, S.-I., Chang, B.S., Yoe, S.M., 2014. Detection of antimicrobial substances from larvae of the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae): Antimicrobial activity of *H. illucens*. *Entomological Research* 44, 58–64. <https://doi.org/10.1111/1748-5967.12050>
- Paterson, S., Wilson, K., Pemberton, J., 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries* L.). *Proceedings of the National Academy of Sciences* 95, 3714–3719. <https://doi.org/10.1073/pnas.95.7.3714>
- Pei, Y., Zhao, S., Chen, X., Zhang, J., Ni, H., Sun, M., Lin, H., Liu, X., Chen, H., Yang, S., 2022. *Bacillus velezensis* EEAM 10B Strengthens Nutrient Metabolic Process in Black Soldier Fly Larvae (*Hermetia illucens*) via Changing Gut Microbiome and Metabolic Pathways. *Front Nutr* 9, 880488. <https://doi.org/10.3389/fnut.2022.880488>
- Peng, Q., Tang, L., Tan, S., Li, Z., Wang, J., Zou, F., 2012. Mitogenomic analysis of the genus *Pseudois*: evidence of adaptive evolution of morphological variation in the ATP synthase genes. *Mitochondrion* 12, 500–505. <https://doi.org/10.1016/j.mito.2012.07.107>
- Petersen, M., Armisén, D., Gibbs, R.A., Hering, L., Khila, A., Mayer, G., Richards, S., Niehuis, O., Misof, B., 2019. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evol Biol* 19, 11. <https://doi.org/10.1186/s12862-018-1324-9>
- Pezzi, M., Scapoli, C., Bharti, M., Faucheux, M.J., Chicca, M., Leis, M., Marchetti, M.G., Mamolini, E., Salvia, R., Falabella, P., Bonacci, T., 2021. Fine Structure of Maxillary Palps in Adults of *Hermetia illucens* (Diptera: Stratiomyidae). *Journal of Medical Entomology* 58, 658–665. <https://doi.org/10.1093/jme/tjaa251>
- PHOCAS, F., BELLOC, C., BIDANEL, J., DELABY, L., DOURMAD, J.-Y., DUMONT, B., EZANNO, P., FORTUN-LAMOTHE, L., FOUCRAS, G., FRAPPAT, B., GONZALEZ-GARCIA, E., HAZARD, D., LARZUL, C., LUBAC, S., MIGNON-GRASTEAU, S., MORENO-ROMIEUX, C., TIXIER-BOICHARD, M., BROCHARD, M., 2017. Quels programmes d'amélioration génétique des animaux pour des systèmes d'élevage agro-écologiques? *INRAE Productions Animales* 30, 31–46. <https://doi.org/10.20870/productions-animales.2017.30.1.2232>
- Picker, M., Griffiths, C., Weaving, A., 2002. *Field Guide to Insects of South Africa*. Penguin Random House South Africa.
- Pujol-Luz, J.R., Francez, P.A. da C., Ururahy-Rodrigues, A., Constantino, R., 2008. The Black Soldier-fly, *Hermetia illucens* (Diptera, Stratiomyidae), Used to Estimate the Postmortem Interval in a Case in Amapá State, Brazil. *J Forensic Sci* 53, 476–478. <https://doi.org/10.1111/j.1556-4029.2008.00659.x>

## Q

- Qi, Y., Xu, J., Tian, X., Bai, Y., Gu, X., 2017a. The complete mitochondrial genome of *Hermetia illucens* (Diptera: Stratiomyidae). *Mitochondrial DNA Part B* 2, 189–190. <https://doi.org/10.1080/23802359.2017.1307708>

## R

- RCC, n.d. Réseau Compost Citoyen – Le réseau des acteurs de la gestion de proximité des biodéchets. URL <https://reseaucompost.org/> (accessed 8.10.22).
- Rech, G.E., Bogaerts-Márquez, M., Barrón, M.G., Merenciano, M., Villanueva-Cañas, J.L., Horváth, V., Fiston-Lavier, A.-S., Luyten, I., Venkataram, S., Quesneville, H., Petrov, D.A., González, J., 2019. Stress response, behavior, and development are shaped by transposable element-induced mutations in *Drosophila*. *PLOS Genetics* 15, e1007900. <https://doi.org/10.1371/journal.pgen.1007900>
- Reswita, R., Noli, Z.A., Rahayu, R., 2022. Effect of Giving Frass *Hermetia Illucen* L. on Soil Physical Chemical Properties, Chlorophyll Content and Yield of Upland Rice (*Oryza Sativa* L.) on Ultisol Soil. *Eduvest - Journal of Universal Studies* 2, 335–346. <https://doi.org/10.36418/edv.v2i2.320>

- Rhode, C., Badenhorst, R., Hull, K.L., Greenwood, M.P., Bester-van der Merwe, A.E., Andere, A.A., Picard, C.J., Richards, C., 2020. Genetic and phenotypic consequences of early domestication in black soldier flies (*Hermetia illucens*). *Animal Genetics* 51, 752–762. <https://doi.org/10.1111/age.12961>
- Ricci, M., Peona, V., Guichard, E., Taccioli, C., Boattini, A., 2018. Transposable Elements Activity is Positively Related to Rate of Speciation in Mammals. *J Mol Evol* 86, 303–310. <https://doi.org/10.1007/s00239-018-9847-7>
- Richoux, P., 2009. Sur la présence d'*Hermetia illucens*(Linnaeus, 1758) (Diptère Stratiomyidae) dans la région lyonnaise/Presence of *Hermetia illucens* (Linnaeus, 1758) (Diptera Stratiomyidae) in the region of Lyons. *Bulletin mensuel de la Société linnéenne de Lyon* 78, 137–138. <https://doi.org/10.3406/linly.2009.13722>
- Rota-Stabelli, O., Daley, A.C., Pisani, D., 2013. Molecular Timetrees Reveal a Cambrian Colonization of Land and a New Scenario for Ecdysozoan Evolution. *Current Biology* 23, 392–398. <https://doi.org/10.1016/j.cub.2013.01.026>
- Rowe, A., 2020. Insects raised for food and feed — global scale, practices, and policy - EA Forum [WWW Document]. URL <https://forum.effectivealtruism.org/posts/ruFmR5oBgqLgTc2b/insects-raised-for-food-and-feed-global-scale-practices-and> (accessed 8.29.22).
- Rozkosný, R., 1983. A Biosystematic Study of the European Stratiomyidae (Diptera): Volume 2 - Clitellariinae, Hermediinae, Pachygasterinae and Bibliography, Series Entomologica. Springer Netherlands.
- Russo, C.A.M., Mello, B., Frazão, A., Voloch, C.M., 2013. Phylogenetic analysis and a time tree for a large drosophilid data set (Diptera: Drosophilidae). *Zoological Journal of the Linnean Society* 169, 765–775. <https://doi.org/10.1111/zoj.12062>

## S

- Saint-Leandre, B., Capy, P., Hua-Van, A., Filée, J., 2020. piRNA and Transposon Dynamics in *Drosophila*: A Female Story. *Genome Biology and Evolution* 12, 931–947. <https://doi.org/10.1093/gbe/evaa094>
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4, 406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
- Sandrock, C., Leupi, S., Wohlfahrt, J., Kaya, C., Heuel, M., Terranova, M., Blanckenhorn, W.U., Windisch, W., Kreuzer, M., Leiber, F., 2022. Genotype-by-Diet Interactions for Larval Performance and Body Composition Traits in the Black Soldier Fly, *Hermetia illucens*. *Insects* 13, 424. <https://doi.org/10.3390/insects13050424>
- Schiavone, A., Dabbou, S., Petracci, M., Zampiga, M., Sirri, F., Biasato, I., Gai, F., Gasco, L., 2019. Black soldier fly defatted meal as a dietary protein source for broiler chickens: effects on carcass traits, breast meat quality and safety. *Animal* 13, 2397–2405. <https://doi.org/10.1017/S1751731119000685>
- Schmitt, E., de Vries, W., 2020. Potential benefits of using *Hermetia illucens* frass as a soil amendment on food production and for environmental impact reduction. *Current Opinion in Green and Sustainable Chemistry* 25, 100335. <https://doi.org/10.1016/j.cogsc.2020.03.005>
- Schreven, S.J.J., de Vries, H., Hermes, G.D.A., Zeni, G., Smidt, H., Dicke, M., van Loon, J.J.A., 2022. Black Soldier Fly Larvae Influence Internal and Substrate Bacterial Community Composition Depending on Substrate Type and Larval Density. *Appl Environ Microbiol* 88, e0008422. <https://doi.org/10.1128/aem.00084-22>
- SDES, 2020. Production de déchets et recyclage [WWW Document]. Bilan environnemental de la France. URL <https://www.statistiques.developpement-durable.gouv.fr/edition-numerique/bilan-environnemental/16-production-de-dechets-et-recyclage.php>



- Seyedalmoosavi, M.M., Mielenz, M., Veldkamp, T., Daş, G., Metges, C.C., 2022a. Growth efficiency, intestinal biology, and nutrient utilization and requirements of black soldier fly (*Hermetia illucens*) larvae compared to monogastric livestock species: a review. *Journal of Animal Science and Biotechnology* 13, 31. <https://doi.org/10.1186/s40104-022-00682-7>
- Shah, S. r. a., Çetingül, I. s., 2022. Anti-microbial and digestibility trials of black soldier fly (*Hermetia illucens*) in rats, rabbits, and reptiles as a model for humans – a review. *Journal of Insects as Food and Feed* 1–8. <https://doi.org/10.3920/JIFF2021.0173>
- Sharma, K., Monobrullah, M., Mohanasundaram, A., Ramani, R., 2016. BENEFICIAL INSECT FARMING (Benefits and Livelihood Generation).
- Sheppard, C., Larry Newton, G., Thompson, S.A., Savage, S., 1994. A value added manure management system using the black soldier fly. *Bioresource Technology* 50, 275–279. [https://doi.org/10.1016/0960-8524\(94\)90102-3](https://doi.org/10.1016/0960-8524(94)90102-3)
- Sheppard, D.C., Tomberlin, J.K., Joyce, J.A., Kiser, B.C., Sumner, S.M., 2002. Rearing Methods for the Black Soldier Fly (Diptera: Stratiomyidae). *JOURNAL OF MEDICAL ENTOMOLOGY* 39, 4.
- Siddiqui, S.A., Ristow, B., Rahayu, T., Putra, N.S., Widya Yuwono, N., Nisa', K., Mategeko, B., Smetana, S., Saki, M., Nawaz, A., Nagdalian, A., 2022. Black soldier fly larvae (BSFL) and their affinity for organic waste processing. *Waste Management* 140, 1–13. <https://doi.org/10.1016/j.wasman.2021.12.044>
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Smetana, S., Schmitt, E., Mathys, A., 2019. Sustainable use of *Hermetia illucens* insect biomass for feed and food: Attributional and consequential life cycle assessment. *Resources, Conservation and Recycling* 144, 285–296. <https://doi.org/10.1016/j.resconrec.2019.01.042>
- Sprangers, T., Ottoboni, M., Klootwijk, C., Ovyne, A., Deboosere, S., De Meulenaer, B., Michiels, J., Eeckhout, M., De Clercq, P., De Smet, S., 2017. Nutritional composition of black soldier fly (*Hermetia illucens*) prepupae reared on different organic waste substrates. *J Sci Food Agric* 97, 2594–2600. <https://doi.org/10.1002/jsfa.8081>
- Ståhls, G., Meier, R., Sandrock, C., Hauser, M., Šašić Zorić, L., Laiho, E., Aracil, A., Doderović, J., Badenhorst, R., Unadirekkul, P., Mohd Adom, N.A.B., Wein, L., Richards, C., Tomberlin, J.K., Rojo, S., Veselić, S., Parviainen, T., 2020. The puzzling mitochondrial phylogeography of the black soldier fly (*Hermetia illucens*), the commercially most important insect protein species. *BMC Evol Biol* 20, 60. <https://doi.org/10.1186/s12862-020-01627-2>
- Stork, N.E., 2018. How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth? *Annu. Rev. Entomol.* 63, 31–45. <https://doi.org/10.1146/annurev-ento-020117-043348>
- Strümpher, W.P., Sole, C.L., Villet, M.H., Scholtz, C.H., 2016. Allopatric speciation in the flightless *Phoberus capensis* (Coleoptera: Trogidae) group, with description of two new species. *Insect Systematics & Evolution* 47, 149–179. <https://doi.org/10.1163/1876312X-47022138>
- Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4. <https://doi.org/10.1093/ve/vey016>
- Surendra, K.C., Tomberlin, J.K., van Huis, A., Cammack, J.A., Heckmann, L.-H.L., Khanal, S.K., 2020. Rethinking organic wastes bioconversion: Evaluating the potential of the black soldier fly (*Hermetia illucens* (L.)) (Diptera: Stratiomyidae) (BSF). *Waste Manag* 117, 58–80. <https://doi.org/10.1016/j.wasman.2020.07.050>
- The IUCN Red List of Threatened Species [WWW Document], n.d. . IUCN Red List of Threatened Species. URL <https://www.iucnredlist.org/resources/summary-statistics> (accessed 7.27.22).

Tichit, M.M., Puillet, L., Sabatier, R.R., Teillard d'Eyry, F.F., 2011. Multicriteria performance and sustainability in livestock farming systems: Functional diversity matters. *Livestock Science* 139, 161–171. <https://doi.org/10.1016/j.livsci.2011.03.006>

## T

Tomberlin, J. k., van Huis, A., 2020. Black soldier fly from pest to ‘crown jewel’ of the insects as feed industry: an historical perspective. *Journal of Insects as Food and Feed* 6, 1–4. <https://doi.org/10.3920/JIFF2020.0003>

Tomberlin, J.K., 2001. Biological, behavioral, and toxicological studies on the black soldier fly (Diptera: stratiomyidae). University of Georgia.

Tomberlin, J.K., Sheppard, D.C., 2002. Factors Influencing Mating and Oviposition of Black Soldier Flies (Diptera: Stratiomyidae) in a Colony. *Journal of Entomological Science* 37, 345–352. <https://doi.org/10.18474/0749-8004-37.4.345>

Tomberlin, J.K., Sheppard, D.C., Joyce, J.A., 2002. Susceptibility of Black Soldier Fly (Diptera: Stratiomyidae) Larvae and Adults to Four Insecticides. *ec* 95, 598–602. <https://doi.org/10.1603/0022-0493-95.3.598>

Turchetto, M., Lafisca, S., Costantini, G., 2001. Postmortem interval (PMI) determined by study sarcophagous biocenoses: three cases from the province of Venice (Italy). *Forensic Science International* 120, 28–31. [https://doi.org/10.1016/S0379-0738\(01\)00412-1](https://doi.org/10.1016/S0379-0738(01)00412-1)

Turner, N., 1960. The Effect of Inbreeding and Crossbreeding on Numbers of Insects. *Annals of the Entomological Society of America* 53, 686–688. <https://doi.org/10.1093/aesa/53.5.686>

## U

United Nations, 2021. UNDESA - Global Population Growth and Sustainable Development [WWW Document]. URL [https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa\\_pd\\_2022\\_global\\_population\\_growth.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2022_global_population_growth.pdf) (accessed 7.29.22).

Urquiza-Haas, E.G., Kotrschal, K., 2015. The mind behind anthropomorphic thinking: attribution of mental states to other species. *Animal Behaviour* 109, 167–176. <https://doi.org/10.1016/j.anbehav.2015.08.011>

## V

van Huis, A., 2013. Edible insects: future prospects for food and feed security, FAO forestry paper. Food and Agriculture Organization of the United Nations, Rome.

Verheyen, G.R., Ooms, T., Vogels, L., Vreysen, S., Bovy, A., Miert, S.V., Meersman, F., Kempen, C., 2018. Insects as an Alternative Source for the Production of Fats for Cosmetics. *JOURNAL OF COSMETIC SCIENCE* 17.

Vicoso, B., Bachtrog, D., 2015. Numerous transitions of sex chromosomes in Diptera. *PLoS Biol* 13, e1002078. <https://doi.org/10.1371/journal.pbio.1002078>

Vieira, C., Biémont, C., 2004. Transposable element dynamics in two sibling species: *Drosophila melanogaster* and *Drosophila simulans*. *Genetica* 120, 115–123. <https://doi.org/10.1023/b:gene.0000017635.34955.b5>

Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., Schatz, M.C., 2017. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* 33, 2202–2204. <https://doi.org/10.1093/bioinformatics/btx153>



## W

- Wang, H., Rehman, K. ur, Liu, X., Yang, Q., Zheng, L., Li, W., Cai, M., Li, Q., Zhang, J., Yu, Z., 2017. Insect biorefinery: a green approach for conversion of crop residues into biodiesel and protein. *Biotechnol Biofuels* 10, 304. <https://doi.org/10.1186/s13068-017-0986-7>
- Wang, Xianhui, Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., Li, B., Cui, F., Wei, J., Ma, C., Wang, Y., He, J., Luo, Y., Wang, Z., Guo, X., Guo, W., Wang, Xuesong, Zhang, Y., Yang, M., Hao, S., Chen, B., Ma, Z., Yu, D., Xiong, Z., Zhu, Y., Fan, D., Han, L., Wang, B., Chen, Y., Wang, Junwen, Yang, L., Zhao, W., Feng, Y., Chen, G., Lian, J., Li, Q., Huang, Z., Yao, X., Lv, N., Zhang, G., Li, Y., Wang, Jian, Wang, Jun, Zhu, B., Kang, L., 2014. The locust genome provides insight into swarm formation and long-distance flight. *Nat Commun* 5, 2957. <https://doi.org/10.1038/ncomms3957>
- Wang, Y.-S., Shelomi, M., 2017. Review of Black Soldier Fly (*Hermetia illucens*) as Animal Feed and Human Food. *Foods* 6. <https://doi.org/10.3390/foods6100091>
- Waterston, R., Ainscough, R., Anderson, K., Berks, M., Blair, D., Connell, M., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Green, P., Halloran, N., Hawkins, T., Hillier, L., Huynh, C., Johnston, L., Kershaw, J., Kirsten, J., Kozono, Y., Laister, N., Latreille, P., Mortimore, B., Panussis, D., Percy, C., Rifkin, L., Roopra, A., Shownkeen, R., Smaldon, N., Smith, A., Smith, M., Staden, R., Sulston, J., Thierry-Mieg, J., Vaudin, M., Vaughan, K., Weinstock, L., Wilson, R., Wohldmann, P., 1993. The Genome of the Nematode *Caenorhabditis elegans*. *Cold Spring Harb Symp Quant Biol* 58, 367–376. <https://doi.org/10.1101/SQB.1993.058.01.043>
- Weiner, J., 2003. Ecology – the science of agriculture in the 21st century. *The Journal of Agricultural Science* 141, 371–377. <https://doi.org/10.1017/S0021859603003605>
- West, B., Zhou, B.-X., 1989. Did chickens go north? New evidence for domestication. *World's Poultry Science Journal* 45, 205–218. <https://doi.org/10.1079/WPS19890012>
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., Schulman, A.H., 2007. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8, 973–982. <https://doi.org/10.1038/nrg2165>
- Winkel, T.D., Wahlen, S., Jensen, T., 2016. Food waste: Book of Abstracts, Nordic Conference on Consumer Research 126–127.
- Winston, M.E., Kronauer, D.J.C., Moreau, C.S., 2017. Early and dynamic colonization of Central America drives speciation in Neotropical army ants. *Mol Ecol* 26, 859–870. <https://doi.org/10.1111/mec.13846>
- Woodley, N., 2001. A World Catalog of the Stratiomyidae (Insecta: Diptera). *Myia* 11, 1–475.

## Y

- Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, Rong, Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, Ruiqiang, Qiao, J., Tang, F., 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 20, 1131–1139. <https://doi.org/10.1038/nsmb.2660>
- Yeates, D., Wiegmann, B., 2017. Phylogeny of Diptera.
- Yeates, D.K., Wiegmann, B.M., 1999. CONGRUENCE AND CONTROVERSY: Toward a Higher-Level Phylogeny of Diptera. *Annual Review of Entomology* 44, 397–428. <https://doi.org/10.1146/annurev.ento.44.1.397>

## Z

- Zhan, S., Fang, G., Cai, M., Kou, Z., Xu, Jun, Cao, Y., Bai, L., Zhang, Y., Jiang, Y., Luo, X., Xu, Jian, Xu, X., Zheng, L., Yu, Z., Yang, H., Zhang, Z., Wang, S., Tomberlin, J.K., Zhang, J., Huang, Y., 2020. Genomic landscape and genetic manipulation of the black soldier fly *Hermetia illucens*, a natural waste recycler. *Cell Res* 30, 50–60. <https://doi.org/10.1038/s41422-019-0252-6>
- Zhang, Q.-L., Zhang, L., Zhao, T.-X., Wang, J., Zhu, Q.-H., Chen, J.-Y., Yuan, M.-L., 2017. Gene sequence variations and expression patterns of mitochondrial genes are associated with the adaptive evolution of two *Gynaephora* species (Lepidoptera: Lymantriinae) living in different high-elevation environments. *Gene* 610, 148–155. <https://doi.org/10.1016/j.gene.2017.02.014>
- Zhang, S., Shu, J., Wang, Y., Liu, Y., Peng, H., Zhang, W., Wang, H., 2019. The complete mitochondrial genomes of two sibling species of camellia weevils (Coleoptera: Curculionidae) and patterns of Curculionini speciation. *Sci Rep* 9, 3412. <https://doi.org/10.1038/s41598-019-39895-8>
- Zhang, Xingxiao, Zhang, Junzhe, Jiang, L., Yu, X., Zhu, H., Zhang, Jianlong, Feng, Z., Zhang, Xiang, Chen, G., Zhang, Z., 2021. Black Soldier Fly (*Hermetia illucens*) Larvae Significantly Change the Microbial Community in Chicken Manure. *Curr Microbiol* 78, 303–315. <https://doi.org/10.1007/s00284-020-02276-w>
- Zhang, Z.-Q., 2013. Animal biodiversity: An update of classification and diversity in 2013\*. *Zootaxa* 3703, 5. <https://doi.org/10.11646/zootaxa.3703.1.3>
- Zhineng, Y., Ying, M., Bingjie, T., Rouxian, Z., Qiang, Z., 2021. Intestinal microbiota and functional characteristics of black soldier fly larvae (*Hermetia illucens*). *Ann Microbiol* 71, 13. <https://doi.org/10.1186/s13213-021-01626-8>
- Zhou, F., Tomberlin, J.K., Zheng, L., Yu, Z., Zhang, J., 2013. Developmental and waste reduction plasticity of three black soldier fly strains (Diptera: Stratiomyidae) raised on different livestock manures. *J Med Entomol* 50, 1224–1230. <https://doi.org/10.1603/me13021>
- Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., Yorke, J.A., 2013. The MaSuRCA genome assembler. *Bioinformatics* 29, 2669–2677. <https://doi.org/10.1093/bioinformatics/btt476>
- Zuckerlandl, E., Pauling, L., 1965. Molecules as documents of evolutionary history. *Journal of Theoretical Biology* 8, 357–366. [https://doi.org/10.1016/0022-5193\(65\)90083-4](https://doi.org/10.1016/0022-5193(65)90083-4)

# Liste des annexes

## Chapitre 1

220 : Profile GenomeScope en utilisant des kmer de taille 31 du génome Cycle Farms.

221 : Script assemblage hybride.

## Chapitre 2

222 : List of all individuals used during the study. Origin and accession number

227 : CO1 gene sequencing protocol. Primers used and Tm.

228 : Phylogenetic tree of the CO1 gene.

229 : Annotated mitochondrial genome of *Hermetia illucens*

230 : Paired blasts of the mitochondrial genomes used in the study

231 : Phylogeny based on the complete mitochondrial genome of *Hermetia illucens* with the ATP8 gene sequence removed.

232 : Separate phylogenetic and time trees.

## Chapitre 3

233 : Code Python de tri des données DnaPipeTE

234 : Code R pour la représentation des famille de TE

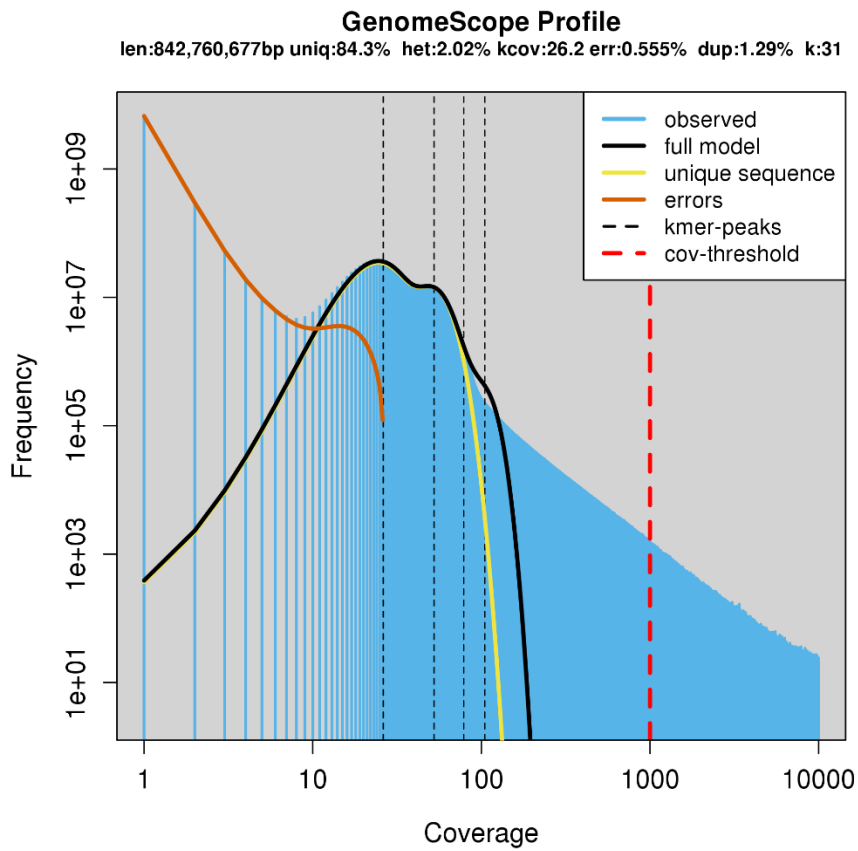
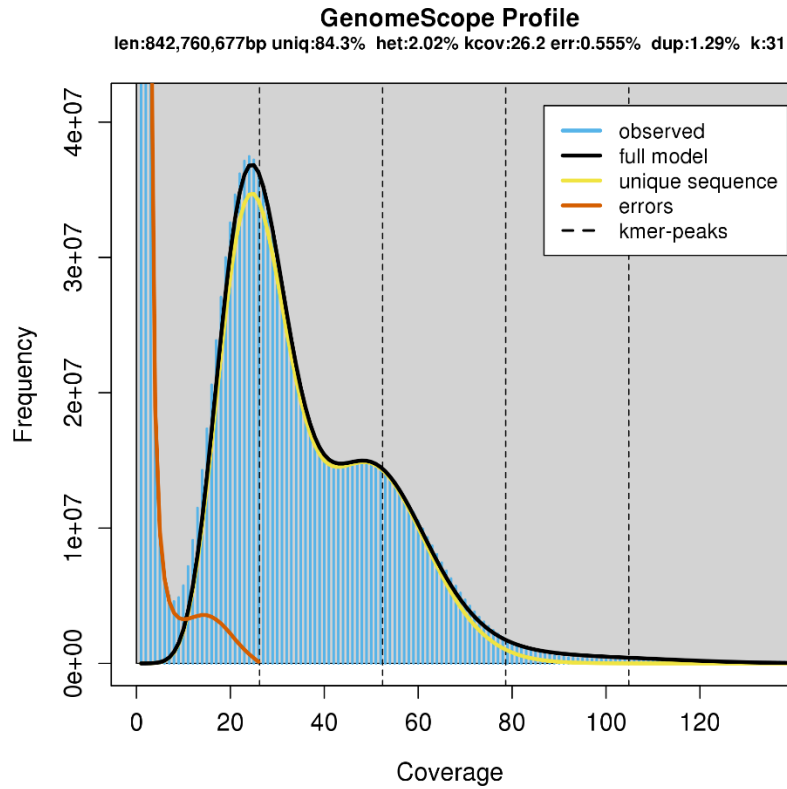
## Discussion

237 : Profile GenomeScope en utilisant des kmer de taille 21 des génomes des souches Ghanéenne et Montpelliéraine.

## En supplément

238 : Zones d'échantillonnage

239 : Article



Annexe 1.01 Profile GenomeScope en utilisant des kmer de taille 31 du génome Cycle Farms.

## Annexe 1.02 : Script assemblage hybride

```
#on s'occupe en premier des données
DATA
#donnée illumina PE pour paire end, un nom à deux lettre qu'on choisit
#la taille moyenne des fragments l'écart type (si inconnu prendre environ 15%)
#le chemin vers les R1 et R2 fastq
PE= OM 150 22
/home/guilliet/hermetia_data/illumina/AOM17/fastq_after_trimming/J4_AOM17_S2_all_R1_001_cutadapt.fastq
/home/guilliet/hermetia_data/illumina/AOM17/fastq_after_trimming/J4_AOM17_S2_all_R2_001_cutadapt.fastq
PE= FM 150 22
/home/guilliet/hermetia_data/illumina/ADF04/fastq_after_trimming/J4_ADF04_S1_all_R1_001_cutadapt.fastq
/home/guilliet/hermetia_data/illumina/ADF04/fastq_after_trimming/J4_ADF04_S1_all_R2_001_cutadapt.fastq

#pour les long reads on prend tout ensemble d'un coup en fasta
NANOPORE= /home/guilliet/hermetia_data/Minion/Hilu01_02_94_206.fasta
END

PARAMETERS
EXTEND_JUMP_READS=0
GRAPH_KMER_SIZE = auto
USE_LINKING_MATES = 0
USE_GRID=0
GRID_QUEUE=all.q
GRID_BATCH_SIZE=300000000
LHE_COVERAGE=30
LIMIT_JUMP_COVERAGE=300
CA_PARAMETERS = cgwErrorRate=0.15
KMER_COUNT_THRESHOLD=1
CLOSE_GAPS=1
NUM_THREADS=48
JF_SIZE = 10000000000
SOAP_ASSEMBLY=0
FLYE_ASSEMBLY=0

END
```

Database	Country	accession number	Name	haplotype	Continent	Lab	Captive	Wild	Mined	GPS	List of all individuals used during the study. Origin and accession number.
Bold	Australia	GCOT2822	LR92227 18 Hermetia COI 5P Australia Ql		Oceania		0	0	0	1	
Bold	Australia	LR92227	LR92227 Australia Queensland Polly Creek		Oceania		0	0	0	1	
Bold	Australia	DIOT1070	DIOT1070 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT1134	DIOT1134 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT1199	DIOT1199 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT1690	DIOT1690 09 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GBMNC1277	GBMNC1277 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1279	GBMNC1279 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1281	GBMNC1281 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1292	GBMNC1292 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1294	GBMNC1294 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1296	GBMNC1296 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1314	GBMNC1314 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1355	GBMNC1355 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1366	GBMNC1366 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1367	GBMNC1367 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GMAOB475	GMAOB475 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB476	GMAOB476 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB477	GMAOB477 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB478	GMAOB478 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB479	GMAOB479 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB480	GMAOB480 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB481	GMAOB481 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB482	GMAOB482 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB483	GMAOB483 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB484	GMAOB484 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB485	GMAOB485 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB486	GMAOB486 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOB487	GMAOB487 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK868	GMAOK868 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK869	GMAOK869 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK874	GMAOK874 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK875	GMAOK875 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK880	GMAOK880 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK882	GMAOK882 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK883	GMAOK883 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK884	GMAOK884 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK885	GMAOK885 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK886	GMAOK886 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK887	GMAOK887 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK888	GMAOK888 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK889	GMAOK889 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK890	GMAOK890 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK891	GMAOK891 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK892	GMAOK892 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK893	GMAOK893 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK895	GMAOK895 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK896	GMAOK896 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK897	GMAOK897 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK899	GMAOK899 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAOK901	GMAOK901 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAQL936	GMAQL936 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAQL938	GMAQL938 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMAQL942	GMAQL942 17 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMQB248	GMQB248 18 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMQB249	GMQB249 18 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMQB250	GMQB250 18 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GMQK014	GMQK014 18 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	NSWHP4133	NSWHP4133 19 Hermetia Australia New Sou		Oceania		0	0	0	1	
Bold	Australia	NSWHP4134	NSWHP4134 19 Hermetia Australia New Sou		Oceania		0	0	0	1	
Bold	Australia	GBMNC1445	GBMNC1445 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1446	GBMNC1446 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1447	GBMNC1447 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1448	GBMNC1448 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1449	GBMNC1449 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1450	GBMNC1450 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	DIOT1119	DIOT1119 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT200	DIOT200 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT282	DIOT282 08 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	DIOT698	DIOT698 10 Hermetia Australia Queensland		Oceania		0	0	0	1	
Bold	Australia	GBMNC1278	GBMNC1278 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1280	GBMNC1280 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1293	GBMNC1293 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1295	GBMNC1295 20 Hermetia Australia		Oceania		0	0	0	1	
Bold	Australia	GBMNC1358	GBMNC1358 20 Hermetia Australia		Oceania		0	0	0	1	
NCBI	Australia	LR792225	LR792225 Australia Queensland Polly Creek		Oceania		0	0	0	1	
NCBI	Australia	LR792253	LR792253 Australia West Australia Perth		Oceania		0	0	0	1	
NCBI	Australia	LR792255	LR792255 Australia New South Wales Beaur		Oceania		0	0	0	1	
NCBI	Australia	LR792257	LR792257 Australia New South Wales Wym		Oceania		0	0	0	1	
NCBI	Australia	LR812706	LR812706 Australia West Australia Perth		Oceania		0	0	0	1	
NCBI	Australia	LR812708	LR812708 Australia New South Wales Beaur		Oceania		0	0	0	1	
NCBI	Australia	LR812710	LR812710 Australia New South Wales Wym		Oceania		0	0	0	1	
NCBI	Australia	LR922254	LR922254 Australia West Australia Perth		Oceania		0	0	0	1	
NCBI	Australia	LR922256	LR922256 Australia New South Wales Beaur		Oceania		0	0	0	1	
NCBI	Australia	LR812707	LR812707 Australia West Australia Perth		Oceania		0	0	0	1	
NCBI	Australia	LR812709	LR812709 Australia New South Wales Beaur		Oceania		0	0	0	1	
Bold	Bangladesh	GMBCF5433	GMBCF5433 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5435	GMBCF5435 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5436	GMBCF5436 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5437	GMBCF5437 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5438	GMBCF5438 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5439	GMBCF5439 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5441	GMBCF5441 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5443	GMBCF5443 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5446	GMBCF5446 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5447	GMBCF5447 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF5452	GMBCF5452 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCH3831	GMBCH3831 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCI4775	GMBCI4775 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCN4781	GMBCN4781 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCN750	GMBCN750 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCN835	GMBCN835 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCN836	GMBCN836 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCN2486	GMBCN2486 20 Hermetia Bangladesh		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF51593	GMBCF51593 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Bangladesh	GMBCF3748	GMBCF3748 15 Hermetia Bangladesh Chitag		Asia		0	0	0	1	
Bold	Benin	GBMNC1389	GBMNC1389 20 Hermetia Benin		IG		0	0	0	1	
Bold	Bhutan	GBMNC1341	GBMNC1341 20 Hermetia Bhutan		Asia		0	0	0	1	
NCBI	Bhutan	LR778160	LR778160 Bhutan		Asia		0	0	0	1	
NCBI	Bhutan	LR778159	LR778159 Bhutan		Asia		0	0	0	1	
Bold	Bolivia	GBMNC1335	GBMNC1335 20 Hermetia Bolivia		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778202	LR778202 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778199	LR778199 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778197	LR778197 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778198	LR778198 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778201	LR778201 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778203	LR778203 Bolivia La Paz Pilon Lajas Indig		SouthAmerica		0	0	0	1	
NCBI	Bolivia	LR778194	LR7781								

<b>Bold</b>	France	GBMNC1371	GBMNC1371_20_Hermelia_France_Avignon	C	Europe	0	0	0	1	
<b>Bold</b>	France	GBMNC1372	GBMNC1372_20_Hermelia_France_Seie	C	Europe	0	0	0	1	
<b>Bold</b>	France	GBMNC1373	GBMNC1373_20_Hermelia_France_NextAlim	C	Europe	0	0	0	1	
	France	This study	37b_LaRochele	A	1	0	0	1		
	France	This study	16b_StMarcellinEnForez	A	1	0	1	0		
	France	This study	16a_StMarcellinEnForez	A	1	0	1	0		
	France	This study	20b_Lyon	A	1	0	1	0		
	France	This study	20c_Lyon	A	1	0	1	0		
	France	This study	21b_Montpellier	A	1	0	1	0		
	France	This study	21c_Montpellier	A	1	0	1	0		
	France	This study	21d_Montpellier	A	1	0	1	0		
	France	This study	21e_Montpellier	A	1	0	1	0		
	France	This study	21f_Montpellier	A	1	0	1	0		
	France	This study	21g_Montpellier	A	1	0	1	0		
	France	This study	21h_Montpellier	A	1	0	1	0		
	France	This study	28b_Lormont	A	1	0	1	0		
	France	This study	28c_Lormont	A	1	0	1	0		
	France	This study	29b_Lormont	A	1	0	1	0		
	France	This study	30a_LaForce	A	1	0	1	0		
	France	This study	30b_LaForce	A	1	0	1	0		
	France	This study	31a_SaintGeorges	A	1	0	1	0		
	France	This study	34c_Vincennes	A	1	0	1	0		
	France	This study	38c_LaRochele	A	1	0	1	0		
	France	This study	39b_Poitiers	A	1	0	1	0		
	France	This study	40a_Poitiers	A	1	0	1	0		
	France	This study	40b_Poitiers	A	1	0	1	0		
	France	This study	40c_Poitiers	A	1	0	1	0		
	France	This study	41b_Poitiers	A	1	0	1	0		
	France	This study	42b_Poitiers	A	1	0	1	0		
	France	This study	46b_VallieresSurCisse	A	1	0	1	0		
	France	This study	47b_Blois	A	1	0	1	0		
	France	This study	47c_Blois	A	1	0	1	0		
	France	This study	48b_Blois	A	1	0	1	0		
	France	This study	50b_Blois	A	1	0	1	0		
	France	This study	52c_Blois	A	1	0	1	0		
	France	This study	53a_Corbais	A	1	0	1	0		
	France	This study	54a_Dardilly	A	1	0	1	0		
	France	This study	58c_Paris	A	1	0	1	0		
	France	This study	7b_Pereguex	A	1	0	1	0		
	France	This study	7b_Pereguex	A	1	0	1	0		
	France	This study	a11_CO1_Angbou	A	1	0	1	0	55 F - Angoulême	
	France	This study	a13_CO1_Treslecleux	A	1	0	1	0	42 F - Treslecleux	
	France	This study	a14_CO1_Dieuleff	A	1	0	1	0	52 F - Dieuleff	
	France	This study	a29_CO1_1a_Pereguex	A	1	0	1	0	41 F - Pereguex	
	France	This study	a30_CO1_19a_Champagne_Au_Mont	A	1	0	1	0	28 F - Champagne au Mont D'Or	
	France	This study	a31_CO1_20a_Lyon	A	1	0	1	0	27 F - Lyon	
	France	This study	a32_CO1_21a_Montpellier	A	1	0	1	0	40 F - Montpellier	
	France	This study	a35_CO1_29a_Lormont	A	1	0	1	0	45 F - Lormont	
	France	This study	a36_CO1_29a_Lormont	A	1	0	1	0	39 F - Lormont	
	France	This study	a37_CO1_30a_LaForce	A	1	0	1	0	38 F - La Force	
	France	This study	a38_CO1_SiGasprest.B	A	1	0	1	0	44 F - St Georges Les Baillargeaux	
	France	This study	a41_CO1_35a_Paris	A	1	0	1	0	37 F - Paris	
	France	This study	a43_CO1_39a_LaRochele	A	1	0	1	0	51 F - La Rochele	
	France	This study	a44_CO1_39a_Poitiers	A	1	0	1	0	36 F - Poitiers	
	France	This study	a45_CO1_40a_Poitiers	A	1	0	1	0	43 F - Poitiers	
	France	This study	a46_CO1_41a_Poitiers	A	1	0	1	0	35 F - Poitiers	
	France	This study	a47_CO1_42a_Poitiers	A	1	0	1	0	34 F - Poitiers	
	France	This study	a49_CO1_46a_VallieresSurCisse	A	1	0	1	0	33 F - Vallieres Sur Cisse	
	France	This study	a50_CO1_47a_Blois	A	1	0	1	0	32 F - Blois	
	France	This study	a51_CO1_48a_Blois	A	1	0	1	0	30 F - Blois	
	France	This study	a53_CO1_50a_Blois	A	1	0	1	0	31 F - Blois	
	France	This study	a54_CO1_15a_SaintMarcellinEnForez	A	1	0	1	0	26 F - St Marcellin En Forez	
	France	This study	a55_CO1_27a_Lormont	A	1	0	1	0	29 F - Lormont	
	France	This study	a56_CO1_38b_LaRochele	A	1	0	1	0	25 F - La Rochele	
	France	This study	a57_CO1_58a_Paris	A	1	0	1	0	24 F - Paris	
	France	This study	a58_CO1_27a_LaRochele	A	1	0	1	0	22 F - La Rochele	
	France	This study	JAFD810_Dieuleff	A	1	0	1	0		
	France	This study	JAFD810_Treslecleux	A	1	0	1	0		
	France	This study	Ave	B	1	0	1	0		
	France	This study	a10_CO1_Aix	C	1	1	0	0	16 F - Avignon	43.917545 4.864731
	France	This study	a12_CO1_Angbou	C	1	1	0	0	11 F - Angoulême	45.6454 0.176845
	France	This study	ADFO4_illumina_CO1	C	1	1	0	0	8 F - Beaufort En Vallée	
	France	This study	ADM17_illumina_CO1	C	1	1	0	0	10 F - Beaufort En Vallée	
	France	This study	JAFD94a_CF	C	1	1	0	0		
	France	This study	JAMND128Av_Ava	C	1	1	0	0		
	France	This study	JAMND128Av	C	1	1	0	0		
	France	This study	Ava	C	1	1	0	0		
	France	This study	Avc	C	1	1	0	0		
	France	This study	Afr	C	1	1	0	0		
	France	This study	Avt	C	1	1	0	0		
	France	This study	35b_Paris	F	1	0	1	0		
	France	This study	35d_Paris	F	1	0	1	0		
	France	This study	58b_Paris	F	1	0	1	0		
	France	This study	49b_Blois	G	1	0	1	0		
	France	This study	a52_CO1_49a_Blois	G	1	0	1	0	47 F - Blois	
	France	This study	34b_Vincennes	I	1	0	1	0		
	France	This study	a39_CO1_33a_NogentSurMarne	I	1	0	1	0	48 F - Nogent Sur Marne	
	France	This study	a40_CO1_34a_Vincennes	I	1	0	1	0	54 F - Vincennes	
	France	This study	43a_Colomiers	J	1	0	1	0		
	France	This study	a48_CO1_43a_Colomiers	J	1	0	1	0	56 F - Colomiers	
<b>NCBI</b>	French/Guiana	LR792234	LR792234_French_Guiana_Roura_Relais_Patay	C	5 SouthAmerica	0	0	0	1	
<b>NCBI</b>	French/Guiana	This study	JAMND128Av	H	SouthAmerica	1	1	0	0	
<b>Bold</b>	French/Polynesia	SYC1218	SYC1218_14_Hermelia_French_Polynesia_A	A	Oceania	0	0	0	1	
<b>Bold</b>	French/Polynesia	SYC3851	SYC3851_14_Hermelia_French_Polynesia_A	A	Oceania	1	0	0	1	
<b>NCBI</b>	French/Polynesia	KX053393	KX053393_French_Polynesia_moorea_island_fm	A	Oceania	0	0	0	1	
<b>NCBI</b>	French/Polynesia	KX053394	KX053394_French_Polynesia_moorea_island_fm	A	Oceania	0	0	0	1	
<b>Bold</b>	Gabon	GMGG0353	GMGG0353_17_Hermelia_Gabon_Moven_Ogou	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1433	GBMNC1433_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1439	GBMNC1439_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1440	GBMNC1440_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1441	GBMNC1441_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1442	GBMNC1442_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1443	GBMNC1443_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>Bold</b>	Ghana	GBMNC1444	GBMNC1444_20_Hermelia_Ghana	G	Africa	0	0	0	1	
<b>NCBI</b>	Ghana	LR778158	LR778158_Ghana	G	Africa	0	0	0	1	
	Ghana	This study	a1_CO1_Gha	C	Africa	1	1	0	0	17 Ghana
	Ghana	This study	a19_CO1_Gha2	C	Africa	1	1	0	0	21 Ghana
	Ghana	This study	a2_CO1_Gha	C	Africa	1	1	0	0	9 Ghana
	Ghana	This study	a20_CO1_Gha2	C	Africa	1	1	0	0	20 Ghana
	Ghana	This study	a21_CO1_Gha2	C	Africa	1	1	0	0	12 Ghana
	Ghana	This study	a3_CO1_Gha	C	Africa	1	1	0	0	13 Ghana
	Ghana	This study	JAFD1540Gha	C	Africa	1	1	0	0	
	Ghana	This study	a15_CO1_Gha	H	SouthAmerica	1	1	0	0	50 French Guiana
	Ghana	This study	a17_CO1_Goyana	H	SouthAmerica	1	1	0	0	49 French Guiana
	Ghana	This study		H	SouthAmerica	1	1	0	0	53 French Guiana
<b>Bold</b>	India	GBDP27981	GBDP27981_19_Hermelia_India	A	40 Asia	0	0	0	1	
<b>Bold</b>	India	AGIR258	AGIR258_17_Hermelia_India	A	Asia	0	0	0	1	
<b>Bold</b>	India	GBDP27980	GBDP27980_19_Hermelia_India	A	Asia	0	0	0	1	
<b>Bold</b>	Indonesia	GBMNC1369	GBMNC1369_20_Hermelia_Indonesia	C	Asia	0	0	0	1	
<b>Bold</b>	Italy	GBMNC1276	GBMNC1276_20_Hermelia_Italy	F	Europe	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1291	GBMNC1291_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1428	GBMNC1428_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1429	GBMNC1429_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1430	GBMNC1430_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1431	GBMNC1431_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1432	GBMNC1432_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>Bold</b>	Kenya	GBMNC1433	GBMNC1433_20_Hermelia_Kenya	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	HMS81262	HMS81262_USA	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	LR778157	LR778157_Kenya	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	LR792258	LR792258_Kenya_Nairobi_Nairobi	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	LR792259	LR792259_Kenya_Nairobi_Nairobi	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	LR812711	LR812711_Kenya_Nairobi_Nairobi	C	Africa	0	0	0	1	
<b>NCBI</b>	Kenya	LR812712	LR812712_Kenya_Nairobi_Nairobi	C	Africa	0	0	0	1	
	Kenya	This study	JAMND128KenBot	C	Africa	1	1	0	0	
	Kenya	This study	JLD278KenKen	C	Africa	1	1	0	0	
	Kenya	This study	JLD278bKenMabu	C	Africa	1	1	0	0	
	Kenya	This study	a25_CO1_Keni	C	Africa	1	1	0	0	6 Kenya
	Kenya	This study	a26_CO1_Keni	C	Africa	1	1	0	0	15 Kenya
	Kenya	This study	a27_CO1_Keni	C	Africa	1	1	0	0	5 Kenya
	Kenya	This study	a28_CO1_KenB	C	Africa	1	1	0	0	14 Kenya
	Kenya	This study	a7_CO1_KenB	C	Africa	1	1	0	0	19 Kenya
<b>Bold</b>	Madagascar	GBMNC1346	GBMNC1346_20_Hermelia_Madagascar	C	Africa	0	0	0	1	
<b>Bold</b>	Madagascar	GBMNC1347	GBMNC1347_20_Hermelia_Madagascar	A	Africa	0	0	0	1	
<b>NCBI</b>	Madagascar	LR778154	LR778154_Madagascar	A	Africa	0	0	0	1	
<b>NCBI</b>	Madagascar	LR778155	LR778155_Madagascar	A	Africa	0	0	0	1	
<b>Bold</b>	Malaysia	GBMNC1364	GBMNC1364_20_Hermelia_Malaysia	A	34 Asia	0	0	0	1	
<b>Bold</b>	Malaysia	GBMNC1315	GBMNC1315_20_Hermelia_Malaysia	A	Asia	0	0	0	1	
<b>Bold</b>	Malaysia	MBPT075	MBPT075_15_Hermelia_Malaysia_Selangor	A	Asia	0	0	0	1	
<b>Bold</b>	Malaysia	GBMNC1348	GBMNC1348_20_Hermelia_Malaysia	E	Asia	0	0	0	1	
<b>NCBI</b>	Malaysia	LR792223	LR792223_Malaysia_Borneo_Sabah_Penampas	A	Asia	0	0	0	1	
<b>NCBI</b>	Malaysia	MT90920	MT90920_Malaysia	E	Asia	0	0	0	1	
<b>NCBI</b>	Mexico	LR778208	LR778208_Mexico_unknown_unknown	D	SouthAmerica	0	0	0	1	
	Mexico	This study	a15_CO1_Mex	D	SouthAmerica	1	1	0	0	57 Mexico
	Mexico	This study	JAFD810mex	D	SouthAmerica	1	1	0	0	

Bold	Netherlands	GBMNC1435	GBMNC1435 20 Hermetia Netherlands	C	Europe	0	0	0	1
Bold	Netherlands	GBMNC1436	GBMNC1436 20 Hermetia Netherlands	C	Europe	0	0	0	1
Bold	Netherlands	GBMNC1437	GBMNC1437 20 Hermetia Netherlands	C	Europe	0	0	0	1
Bold	Netherlands	GBMNC1438	GBMNC1438 20 Hermetia Netherlands	C	Europe	0	0	0	1
Bold	Nigeria	GBMNC1491	GBMNC1491 17 Hermetia Nigeria	C	Africa	0	0	0	1
Bold	Nigeria	GBMNC1411	GBMNC1411 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Nigeria	GBMNC1412	GBMNC1412 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Nigeria	GBMNC1413	GBMNC1413 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Nigeria	GBMNC1414	GBMNC1414 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Nigeria	GBMNC1415	GBMNC1415 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Nigeria	GBMNC1417	GBMNC1417 20 Hermetia Nigeria	IG	Africa	0	0	0	1
Bold	Pakistan	MADIP442	MADIP442 10 Hermetia Pakistan Azad Kashmir	E	Asia	0	0	0	1
Bold	Pakistan	MADIP447	MADIP447 11 Hermetia Pakistan	E	Asia	0	0	0	1
Bold	Pakistan	MADIP448	MADIP448 11 Hermetia Pakistan	E	Asia	0	0	0	1
Bold	Pakistan	MADIP449	MADIP449 11 Hermetia Pakistan	E	Asia	0	0	0	1
NCBI	Panama	LR79233	LR79233 Panama Chiriqui Boquete		28 SouthAmerica	0	0	0	1
Bold	PapuaNewGuinea	PNGFF143	PNGFF143 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
Bold	PapuaNewGuinea	PNGFF146	PNGFF146 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
Bold	PapuaNewGuinea	PNGFF147	PNGFF147 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
Bold	PapuaNewGuinea	PNGFF148	PNGFF148 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
Bold	PapuaNewGuinea	PNGFF150	PNGFF150 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
Bold	PapuaNewGuinea	PNGFF153	PNGFF153 11 Hermetia Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308280	JN308280 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308281	JN308281 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308282	JN308282 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308283	JN308283 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308284	JN308284 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308285	JN308285 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308286	JN308286 Papua New Guinea	A	Oceania	0	0	0	1
NCBI	PapuaNewGuinea	JN308287	JN308287 Papua New Guinea	A	Oceania	0	0	0	1
Bold	Paraguay	GBMNC1300	GBMNC1300 20 Hermetia Paraguay		3 SouthAmerica	0	0	0	1
NCBI	Paraguay	LR792239	LR792239 Paraguay Amambay Chacarru		3 SouthAmerica	0	0	0	1
NCBI	Paraguay	LR792240	LR792240 Paraguay Amambay Chacarru		12 SouthAmerica	0	0	0	1
NCBI	Paraguay	LR792241	LR792241 Paraguay Amambay Chacarru		12 SouthAmerica	0	0	0	1
NCBI	Peru	LR792231	LR792231 Peru Loreto San Juan de Pampom		7 SouthAmerica	0	0	0	1
NCBI	Peru	LR792232	LR792232 Peru Loreto San Juan de Pampom		10 SouthAmerica	0	0	0	1
NCBI	Peru	LR778212	LR778212 Peru El Ingenio unknown		11 SouthAmerica	0	0	0	1
NCBI	Peru	LR778213	LR778213 Peru El Ingenio unknown		16 SouthAmerica	0	0	0	1
NCBI	Peru	LR792230	LR792230 Peru Madre de Dios Rio		18 SouthAmerica	0	0	0	1
Bold	Poland	GBMNC1354	GBMNC1354 20 Hermetia Poland		F Europe	0	0	0	1
Bold	Portugal	IBDPO12	IBDPO12 19 Hermetia Portugal Porto		C Europe	0	0	0	1
Bold	Portugal	IBDPO32	IBDPO32 19 Hermetia Portugal		C Europe	0	0	0	1
Bold	Portugal	IBDPO106	IBDPO106 19 Hermetia Portugal Braganca		C Europe	0	0	0	1
Bold	Portugal	IBDPO19	IBDPO19 Hermetia Portugal Braga		C Europe	0	0	0	1
NCBI	Portugal	MN868749	MN868749 Portugal Porto Via Nova de Gaia		C Europe	0	0	0	1
NCBI	Portugal	MN868766	MN868766 Portugal Braganca Alfandega da F		C Europe	0	0	0	1
NCBI	Portugal	MN868787	MN868787 Portugal		C Europe	0	0	0	1
NCBI	Portugal	MN868751	MN868751 Portugal Braga Barcelos Gilmonde		IG Europe	0	0	0	1
Bold	Russia	GBMNC1389	GBMNC1389 20 Hermetia Rounion		C Europe	0	0	0	1
NCBI	Russia	KY817115	KY817115 Russia		C Europe	0	0	0	1
NCBI	Samoa	LR792226	LR792226 USA Oceania American Samoa		E Oceania	0	0	0	1
Bold	Singapore	GBMNC1365	GBMNC1365 20 Hermetia Singapore		A Asia	0	0	0	1
Bold	Singapore	GBMNC1370	GBMNC1370 20 Hermetia Singapore		E Asia	0	0	0	1
Bold	Singapore	GBMNC1381	GBMNC1381 20 Hermetia Singapore		A Asia	0	0	0	1
Bold	Singapore	GBMNC1350	GBMNC1350 20 Hermetia Singapore		E Asia	0	0	0	1
Bold	Singapore	GBMNC1357	GBMNC1357 20 Hermetia Singapore		E Asia	0	0	0	1
Bold	Singapore	GBMNC1376	GBMNC1376 20 Hermetia Singapore		E Asia	0	0	0	1
Bold	Singapore	GBMNC1380	GBMNC1380 20 Hermetia Singapore		E Asia	0	0	0	1
Bold	SouthAfrica	GBMNC1268	GBMNC1268 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1271	GBMNC1271 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1283	GBMNC1283 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1285	GBMNC1285 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1333	GBMNC1333 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1421	GBMNC1421 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1422	GBMNC1422 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1423	GBMNC1423 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1424	GBMNC1424 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1425	GBMNC1425 20 Hermetia South Africa		C Africa	0	0	0	1
Bold	SouthAfrica	GBMNC1426	GBMNC1426 20 Hermetia South Africa		C Africa	0	0	0	1
NCBI	SouthAfrica	LR792263	LR792263 South Africa Western Cape Cape		C Africa	0	0	0	1
NCBI	SouthAfrica	LR792264	LR792264 South Africa Western Cape Cape		C Africa	0	0	0	1
NCBI	SouthAfrica	LR792265	LR792265 South Africa KwaZulu Natal Durbar		C Africa	0	0	0	1
NCBI	SouthAfrica	LR792266	LR792266 South Africa KwaZulu Natal Durbar		C Africa	0	0	0	1
NCBI	SouthAfrica	LR812716	LR812716 South Africa Western Cape Cape		C Africa	0	0	0	1
NCBI	SouthAfrica	LR812717	LR812717 South Africa Western Cape Cape		C Africa	0	0	0	1
NCBI	SouthAfrica	LR812718	LR812718 South Africa KwaZulu Natal Durbar		C Africa	0	0	0	1
NCBI	SouthAfrica	LR812719	LR812719 South Africa KwaZulu Natal Durbar		C Africa	0	0	0	1
NCBI	SouthKorea	HO541236	HO541236 Korea		32 Asia	0	0	0	1
NCBI	SouthKorea	HO541300	HO541300 Korea		36 Asia	0	0	0	1
NCBI	SouthKorea	HO541253	HO541253 Korea		37 Asia	0	0	0	1
NCBI	SouthKorea	FJ794358	FJ794358 South Korea Buyeo		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794361	FJ794361 South Korea Buyeo		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794363	FJ794363 South Korea Buyeo		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794364	FJ794364 South Korea Buyeo		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794370	FJ794370 South Korea Yangyang		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794370	FJ794370 South Korea Yangyang		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794373	FJ794373 South Korea Yangyang		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794374	FJ794374 South Korea Yangyang		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794389	FJ794389 South Korea Jeju Island		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794392	FJ794392 South Korea Jeju Island		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794394	FJ794394 South Korea Jeju Island		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794403	FJ794403 South Korea Suwon		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794404	FJ794404 South Korea Suwon		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794405	FJ794405 South Korea Suwon		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794406	FJ794406 South Korea Suwon		A Asia	0	0	0	1
NCBI	SouthKorea	HO541186	HO541186 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541194	HO541194 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541207	HO541207 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541208	HO541208 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541209	HO541209 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541224	HO541224 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541228	HO541228 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541233	HO541233 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541234	HO541234 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541246	HO541246 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541249	HO541249 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541272	HO541272 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541280	HO541280 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541281	HO541281 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541286	HO541286 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541287	HO541287 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541289	HO541289 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541289	HO541289 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541291	HO541291 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541293	HO541293 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541294	HO541294 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541305	HO541305 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541311	HO541311 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541315	HO541315 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	HO541316	HO541316 Korea		A Asia	0	0	0	1
NCBI	SouthKorea	FJ794367	FJ794367 South Korea Yangyang		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794369	FJ794369 South Korea Yangyang		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794372	FJ794372 South Korea Yangyang		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794375	FJ794375 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794377	FJ794377 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794378	FJ794378 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794379	FJ794379 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794380	FJ794380 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794381	FJ794381 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794383	FJ794383 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794384	FJ794384 South Korea Bonghwa		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794385	FJ794385 South Korea Jeju Island		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794387	FJ794387 South Korea Jeju Island		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794393	FJ794393 South Korea Jeju Island		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794397	FJ794397 South Korea Suwon		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794398	FJ794398 South Korea Suwon		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794399	FJ794399 South Korea Suwon		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794407	FJ794407 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794408	FJ794408 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794409	FJ794409 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794410	FJ794410 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794411	FJ794411 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794412	FJ794412 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794413	FJ794413 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794414	FJ794414 South Korea Gurye		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794415	FJ794415 South Korea Youngchun		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794416	FJ794416 South Korea Youngchun		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794417	FJ794417 South Korea Youngchun		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794418	FJ794418 South Korea Youngchun		B Asia	0	0	0	1
NCBI	SouthKorea	FJ794419	FJ794419 South Korea Young						





Bold	Uganda	GBMNC1397	GBMNC1397_20_Hermetia_Uganda	F	Africa	0	0	0	1
Bold	Uganda	GBMNC1398	GBMNC1398_20_Hermetia_Uganda	F	Africa	0	0	0	1
Bold	Uganda	GBMNC1399	GBMNC1399_20_Hermetia_Uganda	F	Africa	0	0	0	1
Bold	Uganda	GBMNC1400	GBMNC1400_20_Hermetia_Uganda	F	Africa	0	0	0	1
Bold	Uganda	GBMNC1401	GBMNC1401_20_Hermetia_Uganda	F	Africa	0	0	0	1
Bold	Uganda	GBMNC1402	GBMNC1402_20_Hermetia_Uganda	F	Africa	0	0	0	1
NCBI	United Kingdom	PRJEB31575	Cambindgs_UK_CO1	C	Europe	0	0	0	1
Bold	USA	GBMNC1311	GBMNC1311_20_Hermetia_United_States		24 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1351	GBMNC1351_20_Hermetia_United_States		38 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1377	GBMNC1377_20_Hermetia_United_States		38 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1387	GBMNC1387_20_Hermetia_United_States		38 NorthAmerica	0	0	0	1
Bold	USA	NBNS789	NBNS789_16_Hermetia_United_States_Californ		38 NorthAmerica	0	0	0	1
Bold	USA	NBNS865	NBNS865_17_Hermetia_United_States_Californ		38 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1352	GBMNC1352_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1374	GBMNC1374_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1382	GBMNC1382_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1383	GBMNC1383_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1384	GBMNC1384_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1385	GBMNC1385_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1386	GBMNC1386_20_Hermetia_United_States		41 NorthAmerica	0	0	0	1
Bold	USA	USOJP376	USOJP376_09_Hermetia_United_States_Texas		41 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1375	GBMNC1375_20_Hermetia_United_States		42 NorthAmerica	0	0	0	1
Bold	USA	GBMNC1403	GBMNC1403_20_Hermetia_United_States	A	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1404	GBMNC1404_20_Hermetia_United_States	A	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1405	GBMNC1405_20_Hermetia_United_States	A	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1406	GBMNC1406_20_Hermetia_United_States	A	NorthAmerica	0	0	0	1
Bold	USA	BBDIT097	BBDIT097_11_Hermetia_United_States_Florida	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1282	GBMNC1282_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1376	GBMNC1376_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1391	GBMNC1391_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1392	GBMNC1392_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1393	GBMNC1393_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1394	GBMNC1394_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1395	GBMNC1395_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1396	GBMNC1396_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1407	GBMNC1407_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1408	GBMNC1408_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1409	GBMNC1409_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1410	GBMNC1410_20_Hermetia_United_States	C	NorthAmerica	0	0	0	1
Bold	USA	QUDP145	QUDP145_09_Hermetia_United_States_Oklabor	C	NorthAmerica	0	0	0	1
Bold	USA	GBMNC1313	GBMNC1313_20_Hermetia_United_States	E	NorthAmerica	0	0	0	1
NCBI	USA	LR792228	LR792228_USA_Florida_Panama_City		24 NorthAmerica	0	0	0	1
NCBI	USA	HM433911	HM433911_USA		41 NorthAmerica	0	0	0	1
NCBI	USA	LR792229	LR792229_USA_North_Carolina_Durham		41 NorthAmerica	0	0	0	1
NCBI	USA	PRJNA268391	Berkley_female_CO1	C	NorthAmerica	0	0	0	1
NCBI	USA	PRJNA268391	Berkley_male_CO1	C	NorthAmerica	0	0	0	1
NCBI	USA	LR792267	LR792267_USA_Texas_College_Station	C	NorthAmerica	0	0	0	1
NCBI	USA	LR792220	LR792220_USA_Texas_College_Station	C	NorthAmerica	0	0	0	1
NCBI	Venezuela	LR778210	LR778210_Venezuela_Aragua_state_unknown	C	25 SouthAmerica	0	0	0	1
NCBI	Venezuela	LR778211	LR778211_Venezuela_San_Antonio_de_Los_Alt	C	26 SouthAmerica	0	0	0	1
NCBI	Venezuela	LR778209	LR778209_Venezuela_Yaracuy_state_San_Felip	C	27 SouthAmerica	0	0	0	1
NCBI	Viet	LR792224	LR792224_Viet_Nam_Thua_Thien_Hue_Bach_M	A	Asia	0	0	0	1
Bold	Vietnam	GBMNC1353	GBMNC1353_20_Hermetia_Vietnam	C	Asia	0	0	0	1
Bold	Zambia	GBMNC1345	GBMNC1345_20_Hermetia_Zambia	A	Africa	0	0	0	1
NCBI	Zambia	LR778156	LR778156_Zambia	A	Africa	0	0	0	1

114 23 91 563

Name on figure	Access number			
1 China	NC035232			
2 United Kingdom	PRJEB37575			
3 United States	SAMN01909157			
4 United States	SAMN01909156			
5 Kenya	ERS11394303			
6 Kenya	ERS11394301			
7 Taiwan	ERS11394300			
Vallee				
9 Ghana	ERS11394282			
Vallee				
11 F - Angoulême	ERS11394290			
12 Ghana	ERS11394299			
13 Ghana	ERS11394283			
14 Kenya	ERS11394284			
15 Kenya	ERS11394302			
16 F - Avignon	ERS11394288			
17 Ghana	ERS11394281			
18 Taiwan	ERS11394286			
19 Kenya	ERS11394285			
20 Ghana	ERS11394288			
21 Ghana	ERS11394297			
22 Taiwan	ERS11394287			
23 F - La Rochelle	ERS11394333			
24 F - Paris	ERS11394332			
25 F - La Rochelle	ERS11394331			
Forez	ERS11394329			
27 F - Lyon	ERS11394307			
Mont D'Or	ERS11394306			
29 F - Lormont	ERS11394330			
30 F - Blois	ERS11394326			
31 F - Blois	ERS11394328			
32 F - Blois	ERS11394325			
Cisse	ERS11394324			
34 F - Poitiers	ERS11394322			
35 F - Poitiers	ERS11394321			
36 F - Poitiers	ERS11394319			
37 F - Paris	ERS11394317			
38 F - La Force	ERS11394313			
39 F - Lormont	ERS11394312			
40 F - Montpeillers	ERS11394308			
41 F - Périgueux	ERS11394305			
42 F - Trésleux	ERS11394291			
43 F - Poitiers	ERS11394320			
Baillargeaux	ERS11394314			
45 F - Lormont	ERS11394311			
46 Taiwan	ERS11394284			
47 F - Blois	ERS11394327			
Marne	ERS11394315			
49 French Guiana	ERS11394295			
50 French Guiana	ERS11394294			
51 F - La Rochelle	ERS11394318			
52 F - Dieulouart	ERS11394292			
53 French Guiana	ERS11394296			
54 F - Vincennes	ERS11394316			
55 F - Angoulême	ERS11394289			
56 F - Colomiers	ERS11394323			
57 Mexico	ERS11394293			
Bayonne	ERS11394334			
Bayonne	ERS11394310			
Bayonne	ERS11394309			

### CO1 gene sequencing protocol. Primers used and Tm.

ADN (10 – 50 ng)	2µl
H2O (qsp 25µl)	12,37 µl
Oligo Forward à 5µM	2,5 µl
Oligo Reverse à 5µM	2,5 µl
NEB 5X standard buffer One Taq	5 µl
dNTPs à 10mM each	0,5 µl
NEB OneTaq Hot-Start 5U/µl	0,13 µl
Total	25 µl

Tampon 5X	3,5 µl
Big Dye	1 µl
Primer à 5µM	0,6 µl
Produit de PCR purifié	1 à 2 µl
H2O QSP 20 µl	12,9 à 13,9 µl

Tableau S2

Sequence (5' to 3'): TTTCAACAAATCATAAAGATATTGG

Type: Primer

Length: 25

%GC: 24.0

Hairpin Tm: 35.4

Self Dimer Tm: None

Tm: 52.7

created by: primer3

Sequence (5' to 3'): TAAACTTCTGGGTGTCCGAAGAATCA

Type: Primer

Length: 26

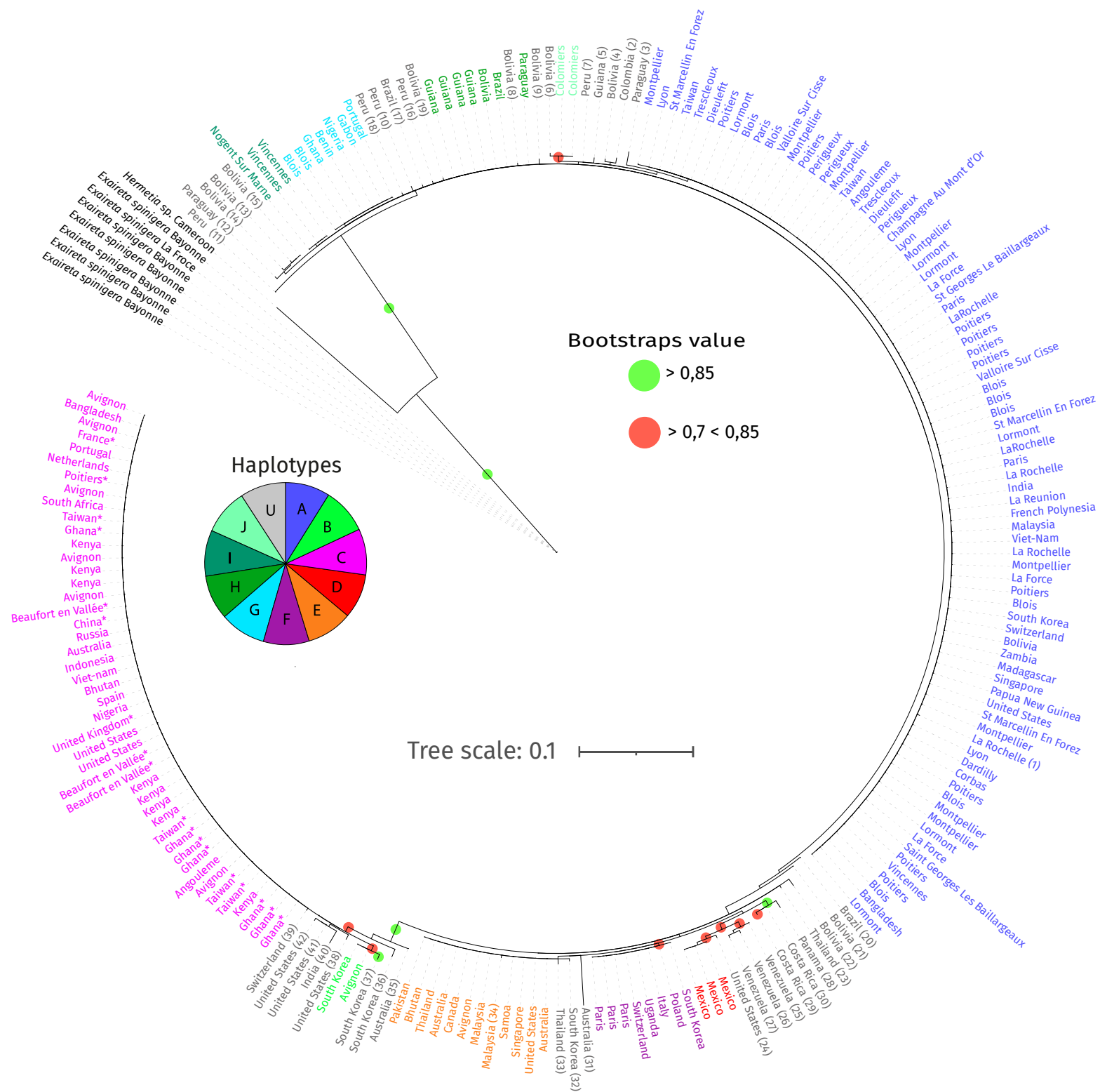
%GC: 42.3

Hairpin Tm: 41.6

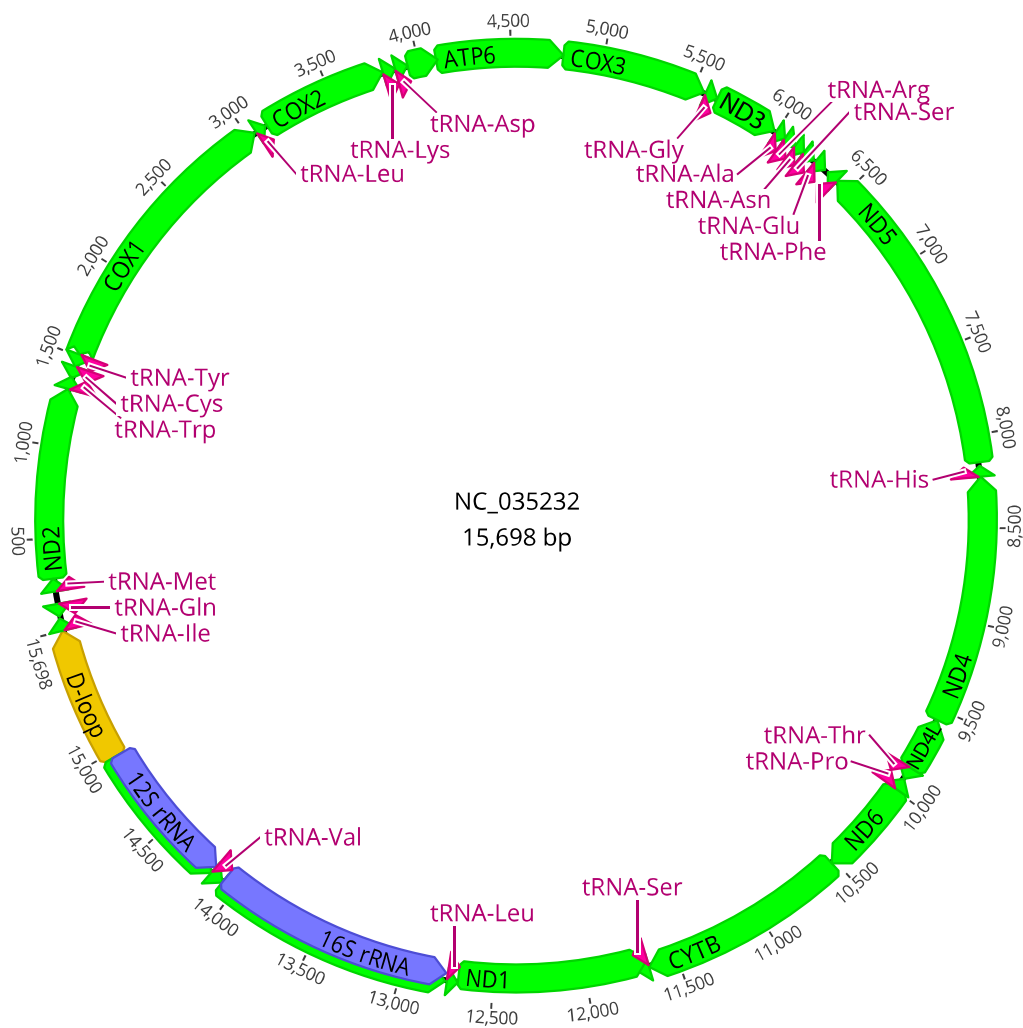
Self Dimer Tm: 17.2

Tm: 62.1

created by: primer3



**Phylogenetic tree of the CO1 gene.**  
 This tree was made by keeping only one sequence per haplotype except for France.  
 The visualization is done using ITOL. An asterisk is placed for individuals from breeding.  
 The statistical support was estimated with 500 Bootstraps with MegaX software using the Tamura 92 G+I model.



### Annotated mitochondrial genome of *Hermetia illucens*

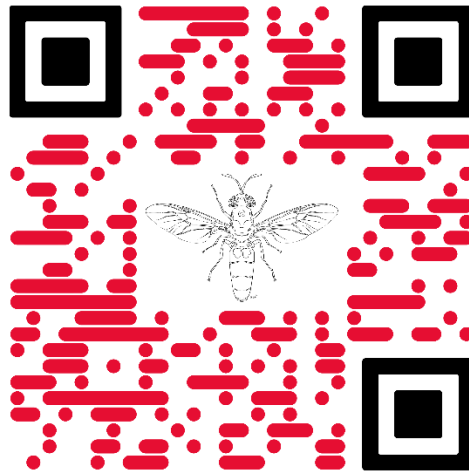
The annotation of the genome was carried out using the MITOS tool.

Visualisation was carried out using the Geneious software.

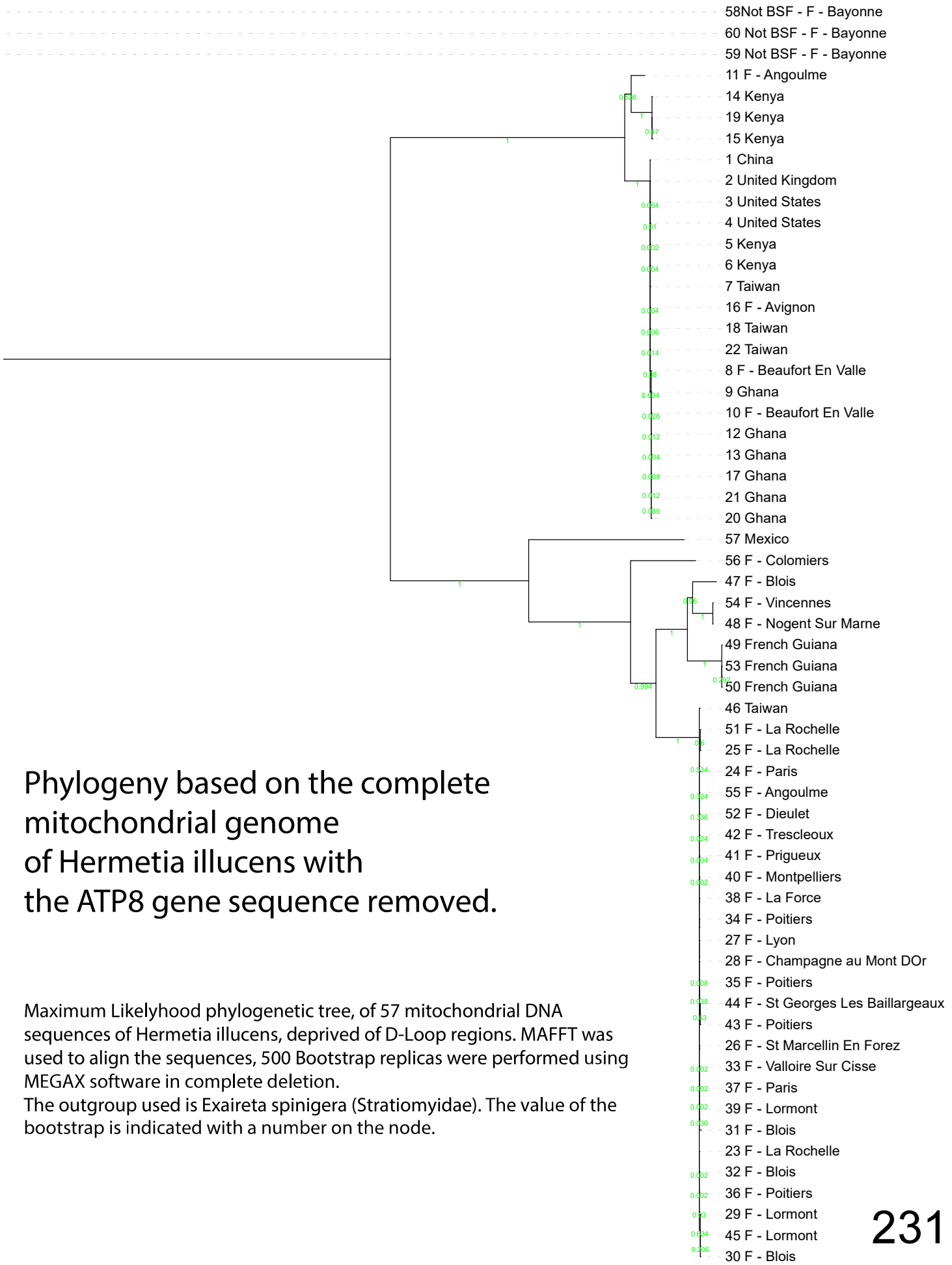
S05 : Paired blasts of the mitochondrial genomes used in the study (Shorter alignment lengths result from comparisons with the outgroup sequences that split the alignments in two fragments.)

Le tableau étant de trop grande taille pour apparaître ici, il peut être téléchargé en suivant le liens suivant : [https://figshare.com/articles/dataset/Supplementary\\_zip/19589332](https://figshare.com/articles/dataset/Supplementary_zip/19589332)

Ou en scannant le code :



Tree scale: 0.01 

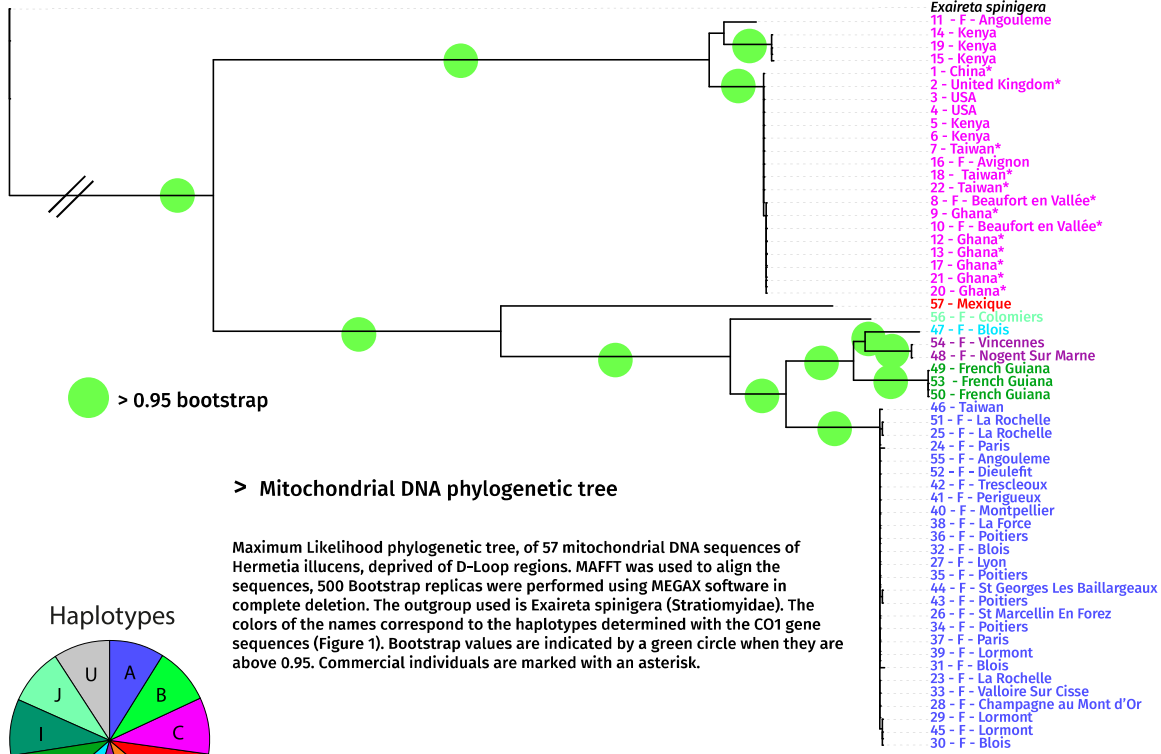


Phylogeny based on the complete mitochondrial genome of *Hermetia illucens* with the ATP8 gene sequence removed.

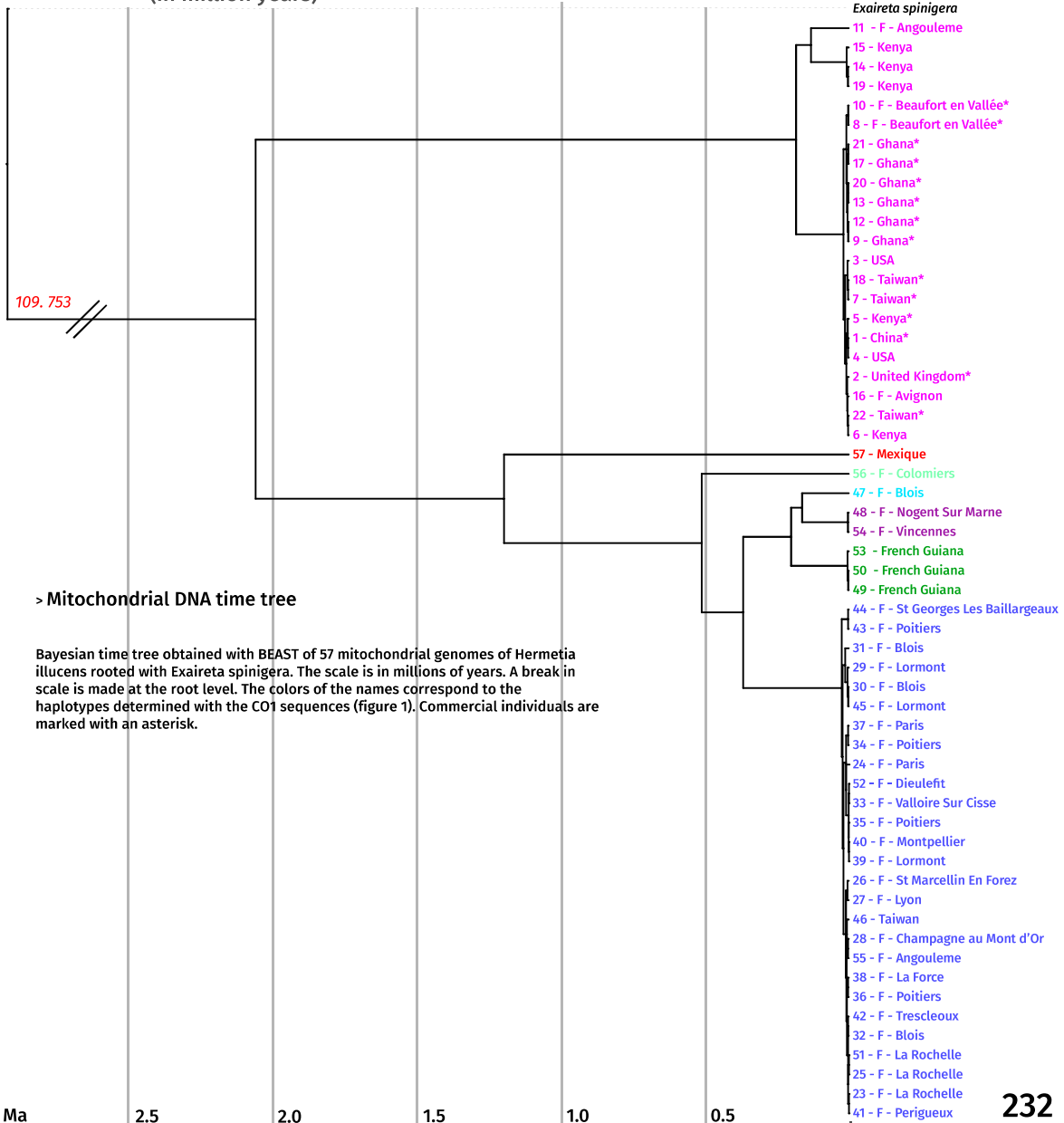
Maximum Likelihood phylogenetic tree, of 57 mitochondrial DNA sequences of *Hermetia illucens*, deprived of D-Loop regions. MAFFT was used to align the sequences, 500 Bootstrap replicas were performed using MEGAX software in complete deletion.

The outgroup used is *Exaireta spinigera* (Stratiomyidae). The value of the bootstrap is indicated with a number on the node.

Tree scale: 0.01



Tree scale : 1  
(in million years)





```

1 # -*- coding: utf-8 -*-
2 """
3 Created on Mon Jan 24 12:28:18 2022
4
5 @author: HP
6 """
7
8 import os #permet de changer de path
9 import pandas as pd #la librairie panda pour l'analyse de datas
10 os.chdir(r"C:\Users\HP\Documents\Thèse\DATAS\dnaPipe_test_reads")
11 #on change pour être dans le bon dossier
12 path=os.getcwd()
13 #on definit la variable path pour y revenir
14
15
16 #on crée une liste de dataframe qui s'ajouteront au fur et à mesure dedans
17 li=[]
18
19 for folder in os.listdir('./'): #on parcourt tous les éléments du dossier
20     if os.path.isdir(folder): #s'il s'agit bien d'un dossier alors
21
22         os.chdir("%s/%s" %(path, folder)) #on se déplace dans ce dossier
23
24         print(os.getcwd())#on print le path pour être sûr
25
26         #noms des colonnes
27         colname = ['counts reads', 'aligned bases', 'dnaPipe contig name', 'RM hit
28         lenght', 'RM annotation', 'RM classification', 'hit lenght / dnaPipeTE contig
29         length']
30
31         #je charge le csv
32         df = pd.read_csv('reads_per_component_and_annotation', delim_whitespace=True,
33         header = None, names=colname)
34
35         #je vais drop les lignes qui ne m'interesse pas
36         df = df.drop(['counts reads', 'dnaPipe contig name', 'RM hit lenght', 'hit
37         lenght / dnaPipeTE contig length', 'RM classification'], axis=1)
38
39         #on drop le lignes qui contiennent des NaN
40         df = df.dropna(axis=0)
41
42         #je renomme la conne aligned bases
43         df = df.rename(columns={"aligned bases": "aligned_bases_%s" %(folder)})
44         df = df.rename(columns={"RM annotation": "RM_annotation"})
45
46         #on garde les lignes dont le nombre de bases aligné_es est >= 1500
47         exec('df = df[df.aligned_bases_%s >= 1500]' %(folder))
48
49         #on regroupe ceux dont le noms en RM_annotation est identique en sommant les
50         aligned bases
51         df = df.groupby('RM_annotation')['aligned_bases_%s' %(folder)].sum()
52         exec('df_%s = df' %(folder))
53
54         exec('li.append(df_%s)' %(folder))
55
56         #retour au path
57         os.chdir(path)
58
59 #on concatène intelligemment tous les dataframes produits
60 frame = pd.concat(li, axis=1)
61
62 #on remplace les NaN par des 0
63 frame = frame.fillna(0)
64 frame.to_csv('all_Reads.csv', index=True) #on enregistre le résultat

```

```
1 library(reshape2)
2 library(ggplot2)
3 library(hrbrthemes)
4 library(pheatmap)
5 library(RColorBrewer)
6 library(viridis)
7 library(dplyr)
8 library(ComplexHeatmap)
9
10
11 setwd("C:/Users/HP/Documents/Thèse/DATAS/dnaPipe/")
12 data <- read.csv("DNA_all_reads_clean_with_accession_number_and_annotation.csv",header
13 = TRUE, row.names=1, sep = ";", stringsAsFactors=FALSE)
14
15 #on transforme les donnée en matrice en les rangeant par la classification de TE
16 mat_data <- as.matrix(data)
17
18 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a12'] <- '01 11 - F -
19 Angouleme'
20 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a28'] <- '02 14 - Kenya'
21 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a7'] <- '03 19 - Kenya'
22 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a26'] <- '04 15 - Kenya'
23 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_Uk_Camb'] <- '06 2 - United
24 Kingdom'
25 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_Berkley_male'] <- '07 3 - USA'
26 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_Berkley_Female'] <- '08 4 -
27 USA'
28 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a27'] <- '09 5 - Kenya'
29 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a25'] <- '10 6 - Kenya'
30 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a22'] <- '11 7 - Taiwan'
31 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a10'] <- '12 16 - F - AVignon'
32 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a8'] <- '13 18 - Taiwan'
33 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a9'] <- '14 22 - Taiwan'
34 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_ADF04'] <- '15 8 - F -
35 Beaufort en Vallée'
36 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a2'] <- '16 9 - Ghana'
37 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_AOM17'] <- '17 10 - F -
38 Beaufort en Vallée'
39 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a21'] <- '18 12 - Ghana'
40 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a3'] <- '19 13 - Ghana'
41 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a1'] <- '20 17 - Ghana'
42 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a19'] <- '21 21 - Ghana'
43 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a20'] <- '22 20 - Ghana'
44 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a15'] <- '23 57 - Mexico'
45 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a48'] <- '24 56 - F -
46 Colomiers'
47 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a52'] <- '25 47 - F - Blois'
48 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a40'] <- '26 54 - F -
49 Vincennes'
50 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a39'] <- '27 48 - F - Nogent
51 Sur Marne'
52 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a17'] <- '28 49 - French
53 Guiana'
54 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a18'] <- '29 53 - French
55 Guiana'
56 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a16'] <- '30 50 - French
57 Guiana'
58 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a4'] <- '31 46 - Taiwan'
59 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a43'] <- '32 51 - F - La
60 Rochelle'
61 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a56'] <- '33 25 - F - La
62 Rochelle'
63 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a57'] <- '34 24 - F - Paris'
64 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a11'] <- '35 55 - F -
65 Angouleme'
66 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a14'] <- '36 52 - F -
67 Dieulefit'
68 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a13'] <- '37 42 - F -
69 Trecleoux'
70 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a29'] <- '38 41 - F -
71 Perigueux'
```

```

54 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a32'] <- '39 40 - F -
Montpellier'
55 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a37'] <- '40 38 - F - La
Force'
56 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a44'] <- '41 36 - F -
Poitiers'
57 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a50'] <- '42 32 - F -Blois'
58 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a31'] <- '43 27 - F - Lyons'
59 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a46'] <- '44 35 - F -Poitiers'
60 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a38'] <- '45 44 - F - St
Georges Les Baillargeaux'
61 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a45'] <- '46 43 - F -
Poitiers'
62 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a54'] <- '47 26 - F - St
Marcellin en Forez'
63 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a47'] <- '48 34 - F -
Poitiers'
64 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a41'] <- '49 37 - F - Paris'
65 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a36'] <- '50 39 - F - Lormont'
66 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a53'] <- '51 31 - F - Blois'
67 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a58'] <- '52 23 - F - La
Rochelle'
68 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a49'] <- '53 33 - F -
Valloire Sur Cisse'
69 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a30'] <- '54 28 - F -
Champagne Au Mont D Or'
70 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a55'] <- '55 29 - F - Lormont'
71 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a35'] <- '56 45 - F - Lormont'
72 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a51'] <- '57 30 - F - Blois'
73 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a33'] <- '00 Exaireta
spinigera 1'
74 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a59'] <- '00 Exaireta
spinigera 2'
75 colnames(mat_data)[colnames(mat_data) == 'aligned_bases_a34'] <- '00 Exaireta
spinigera 3'
76
77
78 #trier dans l'ordre
79 mat_data <- mat_data[,order(colnames(mat_data))]
80
81
82
83 mat_data_inv <- t(mat_data)
84 mat_data_inv <- mat_data_inv[,order(colnames(mat_data_inv))]
85
86
87
88 makeColorRampPalette <- function(colors, cutoff.fraction, num.colors.in.palette)
89 {
90   stopifnot(length(colors) == 3)
91   ramp1 <- colorRampPalette(colors[1])(num.colors.in.palette * cutoff.fraction)
92   ramp2 <- colorRampPalette(colors[2:3])(num.colors.in.palette * (1 - cutoff.fraction
93   ))
94   return(c(ramp1, ramp2))
95 }
96 #on donne le cutoff le 0.001 et nombre de couleur 5000 ici pour mettre le 0 en blanc
97 =>
98 cutoff.distance <- 0.01
99 cols <- makeColorRampPalette(c("white", # distances 0 to cutoff.distance colored from
white to red
100                                     "green", "red"), # distances 3 to max(distmat) colored
101                                     from green to red
102                                     cutoff.distance / max(mat_data_inv),
103                                     1000)
104 annotation_row = data.frame(
105   Haplotypes = factor(rep(c("0Outgroup", "1C", "2D", "3J", "4G", "5F", "6H", "7A"), c(
106     3, 21, 1, 1, 1, 2, 3, 27)))

```

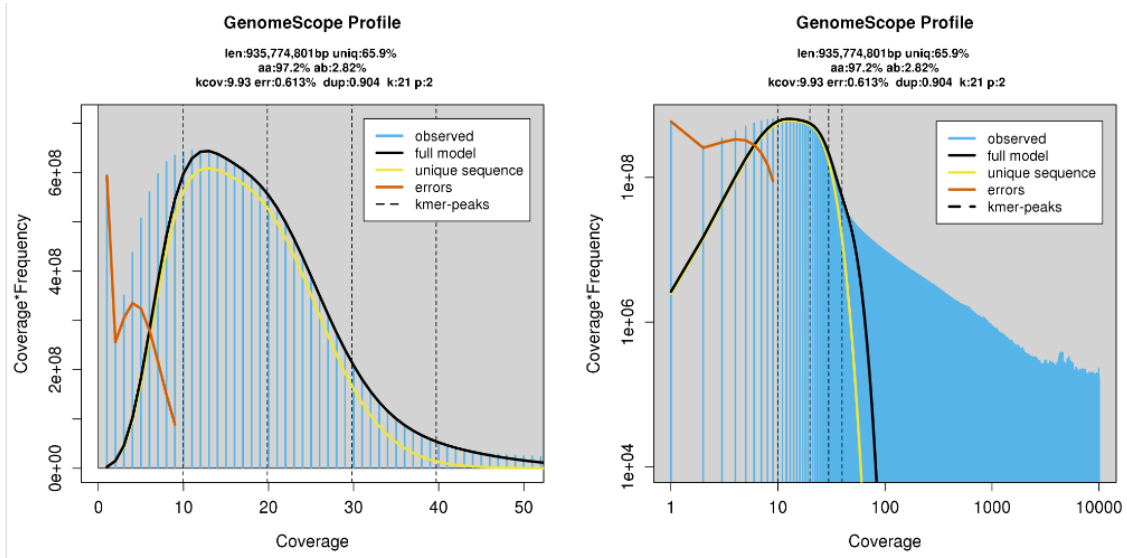
```

107 rownames(annotation_row) <- colnames(mat_data)
108
109
110
111 annotation_col = data.frame(
112   TE = factor(rep(c("Academ", "CMC", "Crypton", "Dada", "Ginger", "hAT", "IS3EU",
113     "Kolobox", "Maverick", "Merlin", "MULE", "P", "PIF", "PiggyBac", "Sola", "TcMar", "Zisupton",
114     "ZUndefined"), c(2, 22, 2, 3, 1, 30, 2, 4, 1, 3, 3, 1, 4, 3, 1, 103, 1, 9)))
115 )
116 rownames(annotation_col) <- colnames(mat_data_inv)
117
118 pheatmap(
119   mat = mat_data_inv,
120   display_numbers = F,
121   cluster_rows = F,
122   cluster_cols = F,
123   color = cols,
124   annotation_row = annotation_row,
125   annotation_col = annotation_col,
126   row_split = annotation_row$Haplotypes,
127   column_split = annotation_col$TE,
128   row_gap = unit(3, "mm"),
129   column_gap = unit(3, "mm"),
130   border = TRUE,
131 )
132

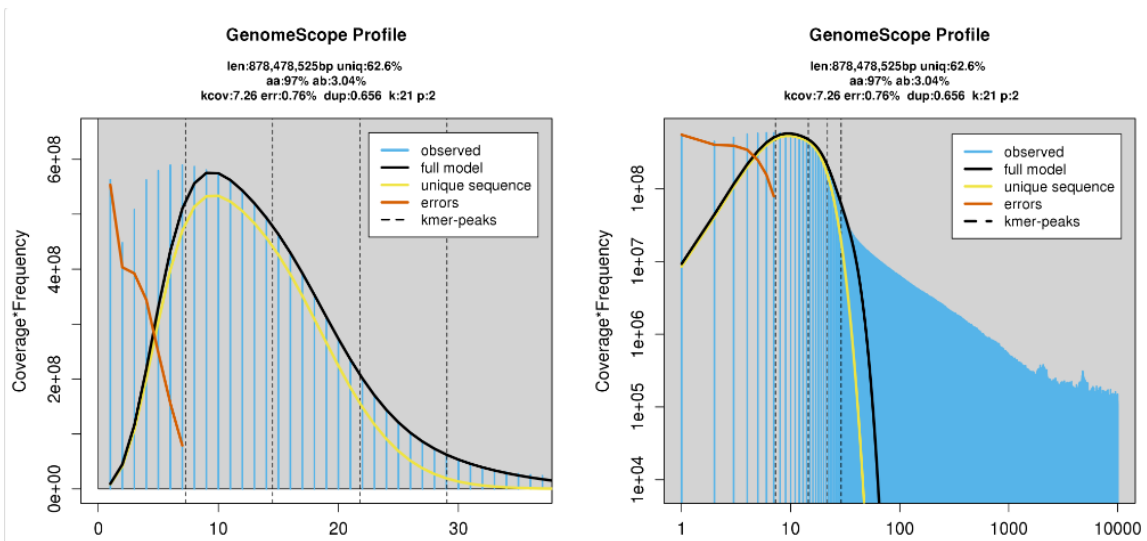
```

Annexe 4.01 Profile GenomeScope en utilisant des kmer de taille 21 des génomes des souches Ghanéenne et Montpelliéraine.

Ghana (haplotype C) :



Montpellier (haplotype A) :



#### Annexe 4.02 Echantillonnage en France

Durant l'été 2020, à l'aide du réseau compost citoyen, nous avons réalisé un échantillonnage de BSF sur le territoire de la France métropolitaine.

Nous avons prospecté dans 25 zones différentes et récupéré des animaux provenant de 18 zones. L'une d'entre elle, à Bayonne, contenait des larves qui n'étaient pas des mouches soldat noires mais la mouche soldat bleue, *Exaireta spinigera*.

La carte ci-après indique les zones de prospections. Les points noirs sont celles où l'on a retrouvé des BSF, les points bleus sont ceux où aucune mouche n'a été trouvée et le point violet est celui où il y avait des mouches soldat bleues.





RESEARCH

Open Access



# What complete mitochondrial genomes tell us about the evolutionary history of the black soldier fly, *Hermetia illucens*

J. Guilliet<sup>1,2,3\*</sup>, G. Baudouin<sup>2</sup>, N. Pollet<sup>1,3</sup> and J. Filée<sup>1,3\*</sup>

## Abstract

**Background:** The Black Soldier Fly (BSF) *Hermetia illucens* is a cosmopolitan fly massively used by industrial companies to reduce biowaste and produce protein and fat for poultry and aquaculture feed. However, the natural history and the genetic diversity of the BSF are poorly known. Here, we present a comprehensive phylogeny and time tree based on a large dataset of complete mitochondrial genomes better to understand the evolution and timing of the BSF.

**Results:** In this study, we analyzed 677 CO1 sequences derived from samples found all over the five continents, leading us to discover 52 haplotypes, including ten major haplotypes. This worldwide cryptic genetic and genomic diversity is mirrored at a local scale in France, in which we found five major haplotypes sometimes in sympatry. Phylogenetic analyses of 60 complete mitochondrial genomes robustly resolved the phylogeny of the major BSF haplotypes. We estimate the separation events of the different haplotypes at more than 2 million years for the oldest branches characterizing the ancestral split between present North American lineages and the other highly diverse south-central American clades, possibly the following radiation beyond the isthmus of Panama northwards. Our data confirm that this North American lineage ultimately gave birth to almost all commercial BSF stocks that participated in the worldwide BSF dissemination through farm escapements.

**Conclusions:** Our data resolve the phylogenetic relationships between the major lineages and give insights into the BSF's short and long-term evolution. Our results indicate that commercial BSF stock's genetic and genomic diversity is very low. These results call for a better understanding of the genomic diversity of the BSF to unravel possible specific adaptations of the different lineages for industrial needs and to initiate the selection process.

**Keywords:** *Hermetia illucens*, Phylogeny, Mitochondrial genome, Phylogeography

## Background

The black soldier fly (BSF) *Hermetia illucens* (Linnaeus, 1758) is a Diptera of the Stratiomyidae family. The native area of the BSF is considered to be Central America and

the northern part of South America, then spread to the entire continent over the last thousands of years [1, 2]. These geographical areas are now considered their natural area [1, 3]. The BSF is now a cosmopolitan fly found in tropical, subtropical, and tempered regions, from the 40th parallel north to the 45th parallel south [3, 4] where it is considered an exotic, non-invasive species, according to the current regulations [5–7]. This fly likely spread through shipping routes [8]. While studies based on naturalistic observations show that it was first discovered in Malta in 1926 [9], it has been suggested that the earliest

\*Correspondence: joseph.guilliet@universite-paris-saclay.fr; jonathan.filee@universite-paris-saclay.fr

<sup>3</sup> IDEEV - Laboratoire Evolution, Génomes Comportement, Ecologie CNRS Université Paris Sud UMR 9191, IRD UMR 247, 12 rue 128, 91198 Gif sur Yvette, France  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

record in Europe was in Italy about 500 years ago [10]. The BSF appeared in the early 1950s in the South-East of France [11] before spreading through the South-West [12] along the Rhône valley [13] and finally along the Atlantic coast [14]. Currently, the BSF can be found on the whole European territory, with few exceptions. It is reasonable to assume that the rise of international traffic and the industrial use of the BSF may have led to a steady flow of introductions in France and worldwide [15, 16].

The BSF was initially used in forensic entomology to date the post-mortem time [17]. Industrial companies heavily use it to reduce biowastes [18, 19], to produce protein [19–21], and fat [19, 20, 22]. These products are used to make animal feed [23] or biodiesel [24, 25].

BSF industries face similar issues as other sectors: knowing how to manage their species and genetic aspects to get the best out of it. To this end, it is crucial to know the genetic diversity of BSF existing internationally, nationally, and locally. Many articles exist about the nutritional aspect of the BSF [22, 26–29] or its microbiota [30–32], but few have focused on the genetic diversity of the BSF. We have some evidence of BSF genetic diversity from previous studies. Ten haplotypes based on the CO1 gene were found in South Korea with 245 individuals [33]. Still, on the CO1 gene and a worldwide sampling, 56 haplotypes were found with a divergence rate up to 4.9%, and it has been shown that individuals belonging to divergent lineages are still interfertile [34]. Finally, based on 15 microsatellite markers, 16 genetic clusters were found with hot spots in South America [2]. Other works focused on more fundamental aspects of BSF and assembled a reference mitochondrial genome and a chromosomal-scale nuclear genome [35–37].

Within the framework of the BSF industry, phylogeography can allow on the one side to account for genetic diversity at different scales (global, national, and local) and, on the other side, to have keys to the geographical origin of the species present in a territory or used by companies. This study may suggest local adaptations or some phenotypic plasticity due to habituation to different biomes.

Our study quantified genetic diversity at different scales: at the global level by studying the variation within the CO1 gene and more locally at the French level. Our report is based on data obtained by sequencing 114 new individuals for the CO1 gene focusing specifically on France (90 French individuals were analyzed) and extracting 563 CO1 sequences from databases. We confirm the hidden genetic diversity within *Hermetia illucens* at all scales of observation, allowing us to find nine different haplotypes. We sequenced and assembled 56 mitochondrial genomes from four different continents and 40 different locations. In addition, we collected sequencing data from databases allowing us to assemble three additional mitochondrial genomes. We also collected a mitochondrial genome already assembled from databases, resulting in 60 mitochondrial genomes. These data allowed us to resolve the phylogenetic relationships between the major haplotypes robustly. In addition, we also provide here the date of the separation events of the different haplotypes found. This separation turns out to be more than 2 million years old for the most distant haplotypes.

## Materials and methods

### Sampling

We decided to conduct a global sampling followed by a specific focus on France. A detailed table of samples and associated metadata is provided in (Additional file 1: Table S1). The origin of all CO1 sequences used in our study is summarised in Table 1, and a summary of sample numbers by continent is provided in Table 2.

Regarding the samples, one adult individual was found in Mexico in the wild in 2016. Individuals coming from Asia (Taiwan), Africa (Ghana and Kenya), and South America (French Guiana) were collected during the spring and summer of 2019. They come either from the wild or breeding farms (see Additional file 1). In France, the collection of BSF was done with the help of the citizen compost network [38] during the summer of 2020. We recovered mainly L5 larvae exclusively from shared and personal composts and some nearby adults. Between

**Table 1** Origin of CO1 sequences and mitochondrial genomes

Origin	Number of individuals (CO1 sequences)	Number of individuals (mitochondrial genome)
Collection from the wild	91	43
Commercial origin	23	13
Mined from databases (GenBank and BOLD)	560	1
Assembled from SRA (raw sequences)	3	3
Total	677	60



**Table 2** Geographic origin of CO1 sequences and mitochondrial genomes

Geographic origin	Number of individuals (CO1 sequences)	Number of individuals (mitochondrial genome)
South America	55	4
North America	43	2
Asia	277	4
Oceania	106	0
Africa	78	11
Europe (including France)	118 (90)	39 (38)
Total	677	60

one and ten individuals were collected per area. In some French localities, collections in the wild have been made near BSF breeding areas (about 5 km). We obtained larvae in the areas of Blois, Bordeaux, Toulouse, Montpellier, Lyon, Paris. All these samples were stored in 100% ethanol at  $-20^{\circ}\text{C}$  prior to DNA extraction.

In the NCBI Sequence Read Archive (SRA), we gathered all the CO1 sequences on NCBI and Bold and whole-genome sequencing data and reconstructed the mitochondrial genome. We were surprised in the Bayonne area to have BSF-like larvae, which turned out to be *Exaireta spinigera* larvae. These larvae were analyzed in the same way as the BSF.

#### DNA extraction

We extracted DNA from larvae, pupae, and imago under the same conditions by grinding individuals in liquid nitrogen using a mortar and pestle. The ground samples were then processed according to the protocol of the Macherey–Nagel AXG 100 [39] kit for the isolation of genomic DNA from tissue with some modification: we added a second purification step with 3.5 ml of freshly prepared ethanol and centrifugation at 15,000 rpm at  $4^{\circ}\text{C}$  for 15 min just before the final re-suspension. The DNA pellet was finally re-suspended in 100  $\mu\text{l}$  of Tris 10 mM EDTA 1 mM pH 8.0. The purity was evaluated by spectrophotometry using a Nanodrop [40], and the concentration was quantified using the Qubit kit DNA Broad range [41]. DNA integrity was monitored using routine agarose gel electrophoresis.

#### CO1 sequencing

We designed specific CO1 primers from the reference mitochondrial DNA [35] (Additional file 2: Table S2), then a mix was realized using NEB 5 $\times$  standard buffer One Taq. The mix is then amplified in ONETAQ HOT 45 PCR. We purified the PCR products using the Illustra

Exostar 1-step GE Healthcare kit before Sanger cycle sequencing. Cycle sequencing reactions were performed using the BigDye 3.1 chemistry, purified by precipitation, and sequenced using an ABI 3130 sequencer (Applied Biosystems). Base-calling was performed using the sequencing analysis software version 5.3 (Applied Biosystems). We assembled forward and reverse reads using the Geneious V R11 [42] assembler with the sensitivity set at medium/fast. The regions containing the primers were excised, and we kept only the sequences having more than 75% of high-quality base calls.

#### Multiple sequence alignment and haplotype network reconstruction

We aligned CO1 sequences obtained in the laboratory and those recovered from the databases using the MAFFT [43] aligner included as a plug-in in the Geneious software [42]. We used the algorithm G-INS-i with a scoring matrix of 200 PAM/k=2, a gap open penalty at 1.53, and an offset value at 0.123. We exported the multiple sequence alignment in nexus format [44] for haplotype network reconstruction and visualization using POPART [45]. We modified the output file to add geographical origins to the sequences.

#### Shotgun sequencing and mitochondrial DNA assembly

A total of 56 BSF individuals were used for whole-genome shotgun sequencing using an Illumina method by Novogene UK [46–50]. The technique used is pair-end with a medium coverage rate of  $4.68\times$  (SD 0.50). The resulting sequences were checked for quality and counted for read parity. We take advantage of the fact that mtDNA is highly repeated in insect DNA. In insects, mtDNA represents on average 0.42% of the total DNA in the genome sequence project [51]. As stipulated, we used 5000 kb of data/sample (4–5 $\times$  coverage) for the mitochondrial assembly, which on average contains  $(5000 \times 0.42)/100 = 21,000$  kb of mtDNA. As the mitochondrial genome is about 15 kb, we obtain coverage of  $1400\times$  for the mitochondrial genome, which is mainly sufficient to obtain a complete genome in one contig. The sequences were assembled de novo with the MEGAHIT software [52] (based options); the mitochondrial genomes were thus recovered from the assemblies by blast [53] using the reference BSF mitochondrial genome sequence (NC\_035232). The mitochondrial DNAs were annotated using Mitos online software [54], and the D-Loop zone was removed. We added four complete mitochondrial sequences from databases or reconstructed them via reading sequences available on the SRA

[55] using the same assembly procedure and then aligned them under the same conditions.

**CO1 and whole mitochondrial genome phylogeny**

We use the MEGAX [56] software to determine the best model and compute an ML Tree with 500 Bootstraps using a complete deletion parameter. We visualized the data with the online tool ITOL [57], and we rooted the tree with the Stratiomyidae *Exaireta spinigera*. We chose *Exaireta spinigera* because it was the closest species for which we have a complete mitochondrial genome (the blastn result of the complete mitochondrial genome of *Hermetia illucens* vs. *Exaireta spinigera* gives us 81.28% of similarity). To simplify the interpretation of the phylogeny, we have kept only a single identical sequence per haplotype.

**Time tree**

We used the BEAST software [58] to estimate the divergence time of the different BSF groups. The analysis was performed with an expansion growth model and an uncorrelated lognormal relaxed clock with the proposed

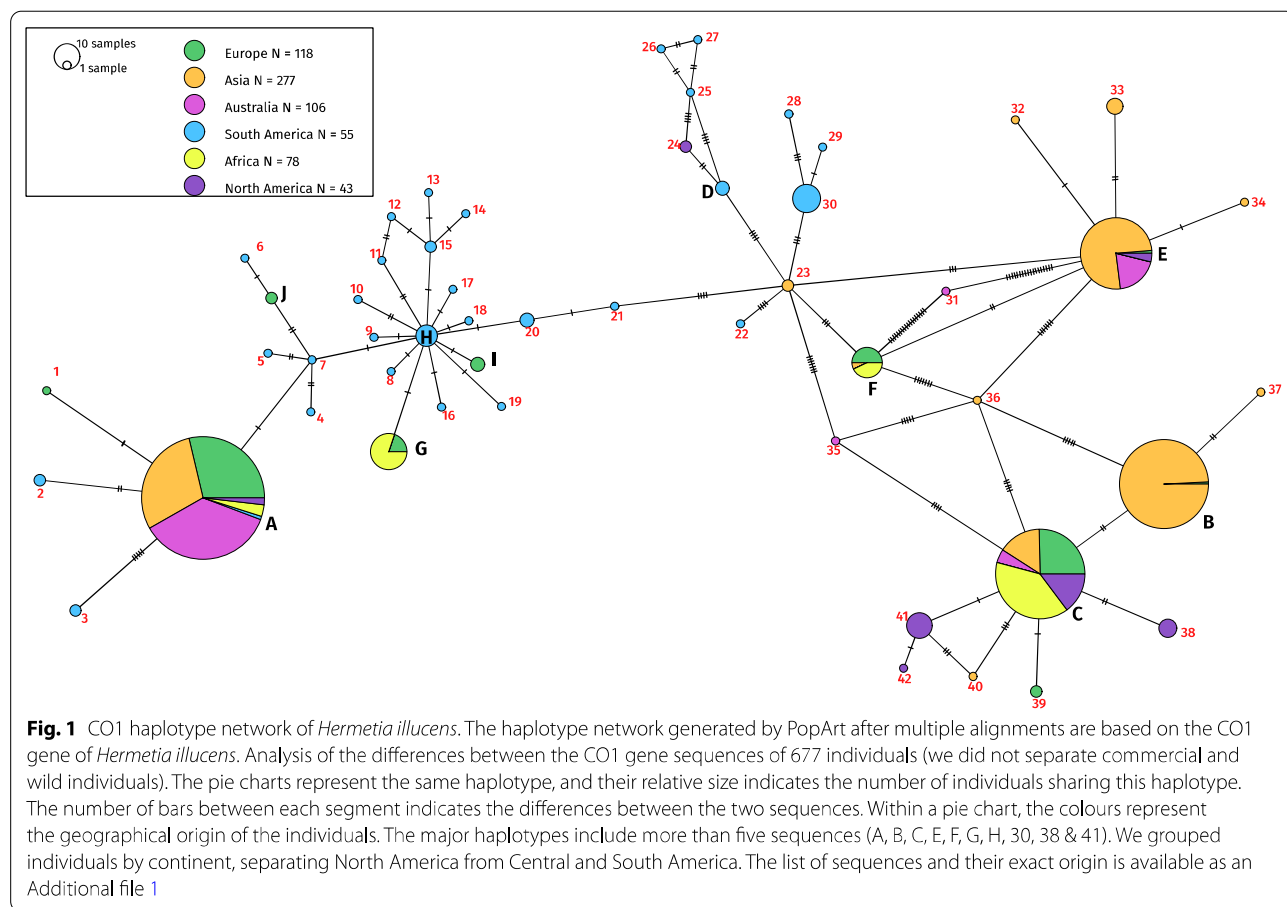
insect molecular clock. Ten million iterations were performed, and then with the TreeAnnotator [59], the obtained trees were clustered with a burn-in of 25% of the trees. The visualization of the final tree was done with the online software ITOL [57].

**Results**

**Haplotype network**

We reconstructed a haplotype network from the analysis of 677 sequences spanning 658 nucleotides of the CO1 mitochondrial gene. Using the minimum spanning method [60], 52 distinct haplotypes were delineated from the analysis of this CO1 fragment. These different haplotypes group up to 230 sequences (Fig. 1).

There are 31 haplotypes containing only one sequence and 11 haplotypes with between 2 and 4 sequences from the same continent. We have named these 42 haplotypes “minor haplotypes.” We observed ten haplotypes between five and 230 sequences, including six multi-continental haplotypes that we will refer to as major haplotypes. We found an important diversity both within and between continents using the CO1 phylogenetic mitochondrial gene. Minor haplotypes represent 8.5% of all sequences



(58 sequences) and 80.8% of haplotypes. Major haplotypes represent 91.5% of the sequences (619 sequences) and 19.2% of the haplotypes.

In Europe, we found ten distinct haplotypes from 118 sequences, nine for Asia from 277 sequences, five for Oceania from 106 sequences, 30 for South America from 55 sequences, four in Africa from 78 sequences, and seven in North America from 43 sequences. Some major haplotypes are represented in a more restricted area. For example, haplotype B is mainly in Asia (Fig. 1). We observed a high haplotype diversity in South America with 28 minor haplotypes that contain between one and four sequences. The haplotype C (we choose C for “commercial”) contains all the sequences derived from farmed strains of the industry working on BSF.

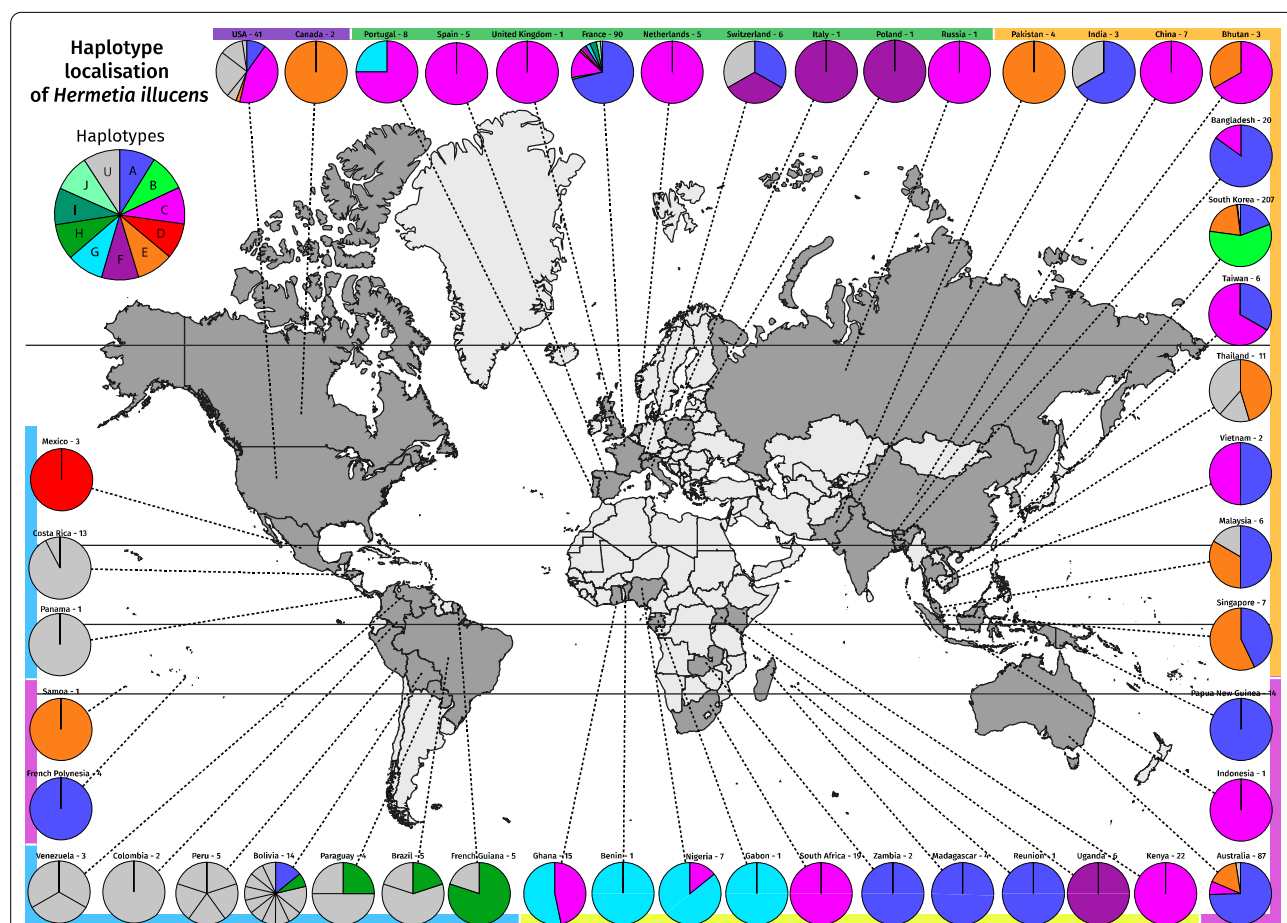
**Global distribution analysis**

We visualized the actual distribution of the different haplotypes found previously on a distribution map (Fig. 2).

Each haplotype from A to J is now represented with a color (Fig. 2); haplotypes numbered in the haplotype network (Fig. 1) are now represented in grey.

We observed great variability in South America, highlighted on the map by the grey pie charts representing mostly minor haplotypes. We observe 22 singletons in South America out of the 52 haplotypes found, more than 40% of the total diversity. On the other hand, we found six out of ten major haplotypes in Europe.

Haplotype A and C were the most represented haplotypes in our sampling: Haplotype A was present on all continents and represented 34% of all sequences, and haplotype C was only absent from South America and represented 18% of all sequences. Some haplotypes



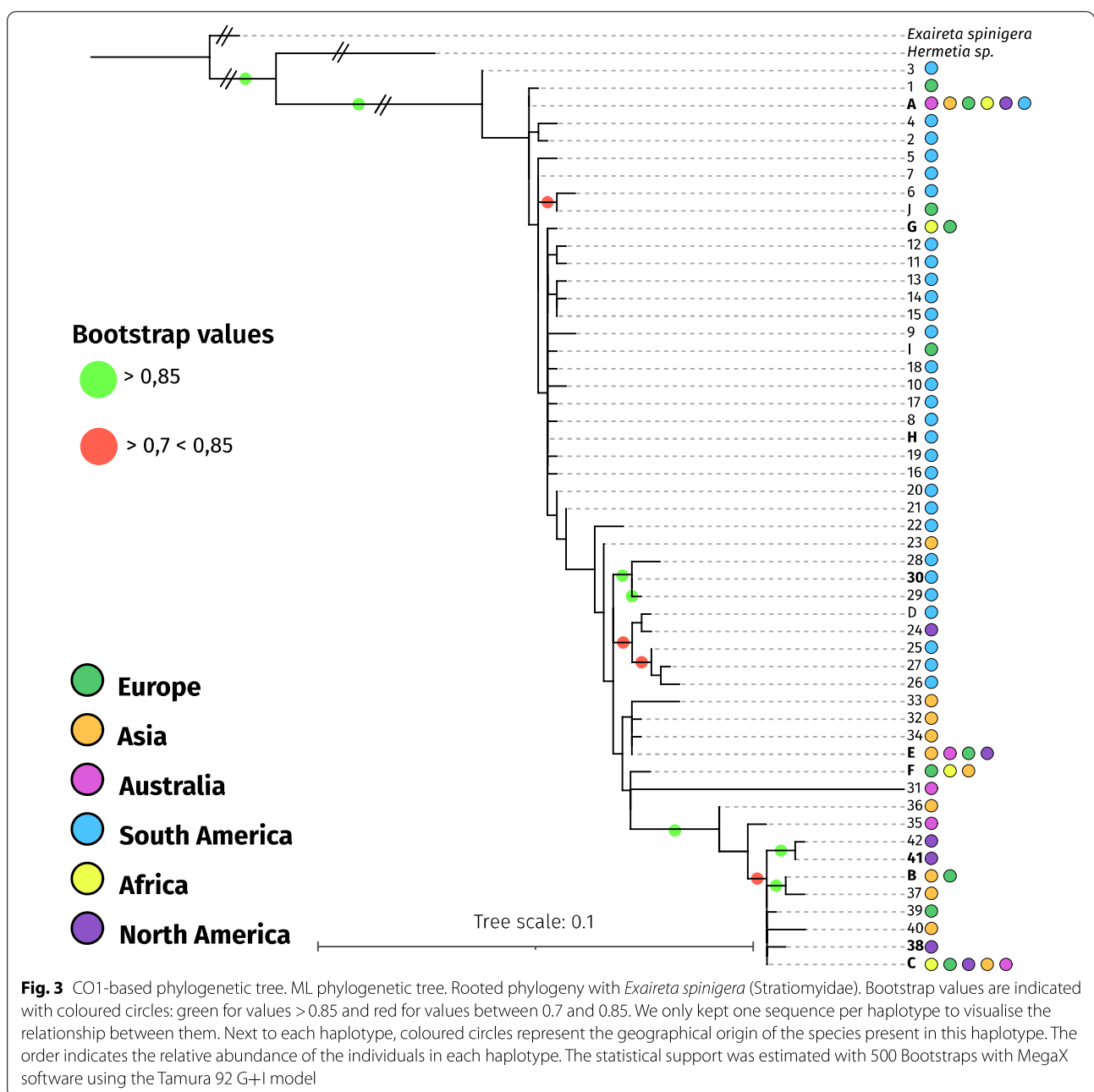
**Fig. 2** World CO1 haplotype distribution. Global distribution of the different *Hermetia illucens* haplotypes based on the CO1 gene. The pie charts represent the proportion of each haplotype within a country. The country's name is followed by the number of individuals analysed. A coloured band representing the continents is next to each pie chart. The colours represent the haplotypes determined in Fig. 1. The haplotypes represented with grey colours, indicated with the letter U, are represented with a number (1–42) in Fig. 1. Separation is made in the pie chart when the haplotypes change for the same country. For example, Venezuela contains three unique haplotypes, meaning that these three individuals are different. A summary table can be found in Additional file 1

seemed to be restricted to some world areas, such as haplotype G, which was present in all the localities in West Africa and surprisingly in a Portuguese sample, or the haplotype H in South America. Haplotype F was present in Europe, Uganda, and South Korea. Haplotype E was more present in Asia, Oceania, and North America. Finally, some haplotypes such as B, I, and J present rare sequences found in more restricted areas such as B in Korea or H and J in France (Fig. 2).

**CO1 phylogeny**

During our CO1 gene analyses, some BSF-like larvae were identified as *Exaireta spinigera* larvae. The sequences were amplified and sequenced with the same primers as for BSF.

By taking only one sequence for each haplotype, we have reconstructed a phylogeny of the CO1 gene sequence. The CO1 phylogeny was rooted with the *Exaireta spinigera* (Fig. 3). A sample from Cameroon was revealed to belong to a closely unidentified species



**Fig. 3** CO1-based phylogenetic tree. ML phylogenetic tree. Rooted phylogeny with *Exaireta spinigera* (Stratiomyidae). Bootstrap values are indicated with coloured circles: green for values > 0.85 and red for values between 0.7 and 0.85. We only kept one sequence per haplotype to visualise the relationship between them. Next to each haplotype, coloured circles represent the geographical origin of the species present in this haplotype. The order indicates the relative abundance of the individuals in each haplotype. The statistical support was estimated with 500 Bootstraps with MegaX software using the Tamura 92 G+I model



of BSF and was termed *Hermetia* sp. Cameroon in our subsequent analysis (Blastn score between *Hermetia* sp. from Cameroon and BSF gives us 85.01% identity with an E Value of  $2e-174$ ).

The CO1 gene phylogeny supports the analyses performed with the haplotype network (Fig. 1), showing remarkable sequence diversity among the CO1 gene. Interestingly, all the commercial strains are found in the C haplotype (a detailed phylogenetic tree is given as a supplementary with an asterisk placed on individuals from farms (Additional file 3: Fig. S3).

A group containing haplotypes 35 to 42 and haplotypes B and C forms a divergent group from the others, with strong statistical support (bootstrap value of 0.99). Some individuals captured in the wild have been found near BSF commercial breeding sites. For example, we have captured C haplotype individuals in Avignon close by companies working with the BSF.

Although the global resolution of the deep branches of the tree is low, it is also possible to better understand the relationship between the haplotypes: haplotypes B and C are close to each other and well separated from other haplotypes. The other haplotypes are not very well resolved with the CO1 sequences. Finally, the samples coming from Latin America are scattered all over the tree, showing their great diversity and supporting the hypothesis of the geographical origin of the BSF.

#### France displays vast BSF genetic diversity at local scales

Among the ten major haplotypes, our sampling effort in France showed that five are present in France, and a unique haplotype is only found in France (Fig. 4) in La Rochelle. Haplotype A remains the dominant haplotype in France (71.1% of all sequences) in seven of the eight localities (excluding Beaufort en Vallée, where the flies come from a company). Regarding the haplotype C, BSF belonging to this group have been found in four locations in the regions of Poitiers, Bordeaux, Beaufort-En-Vallée and Montpellier. Except for Beaufort-En-Vallée, where the flies come from a company, the other BSF come from natural environments. BSF industries exist in all these localities. More surprisingly, different haplotypes coexist in some localities. For example, we found three different haplotypes in the region of Paris and four in the region of Montpellier. Finally, we found a singular haplotype in Colomiers (J) and La Rochelle (1).

#### Complete mitochondrial genome analyses

To better understand BSF phylogeny and estimate the timing of haplotype divergence, we analyzed the mitochondrial genomes of 60 individuals (56 that have been sequenced and assembled, three from which sequences have been recovered and assembled and one recovered),

including 3 *Exaireta spinigera* genomes (used as an outgroup). The 60 reconstructed mitochondrial genomes have the typical structure of insect mtDNA: 13 protein-coding genes, 22 tRNA, two rRNA-coding genes, and non-coding regions and control regions (an example is given as an Additional file 4: Fig. S4). The genetic differences are homogeneously distributed overall mtDNA and are not localized on specific genes (Fig. 5A). A great genetic diversity at the mitochondrial genome level is observed. Two very homogeneous groups, A and C, are well represented, and sequences of the C haplotype display many mutations all along the genomes from the others. The two sequences with the most differences are 55-F-Angouleme and 21-Ghana, with a similarity percentage of 96.373% (all results are in Additional file 5. Shorter alignment lengths result from comparisons with the outgroup sequences that split the alignments into two fragments.) The mitochondrial genomes of the three *Exaireta spinigera* were extremely close to each other (100% by BLAST—Additional file 5), so we kept only one to serve as an outgroup.

We also wanted to know if the mutations are equally (or not) distributed between the different genes. Therefore, we compared the similarity of the 13 mitochondrial genes of the 57 BSFs (Fig. 5B) to those of *Exaireta spinigera*. The results are presented as percent sequence similarity.

We grouped the individuals into three different categories:

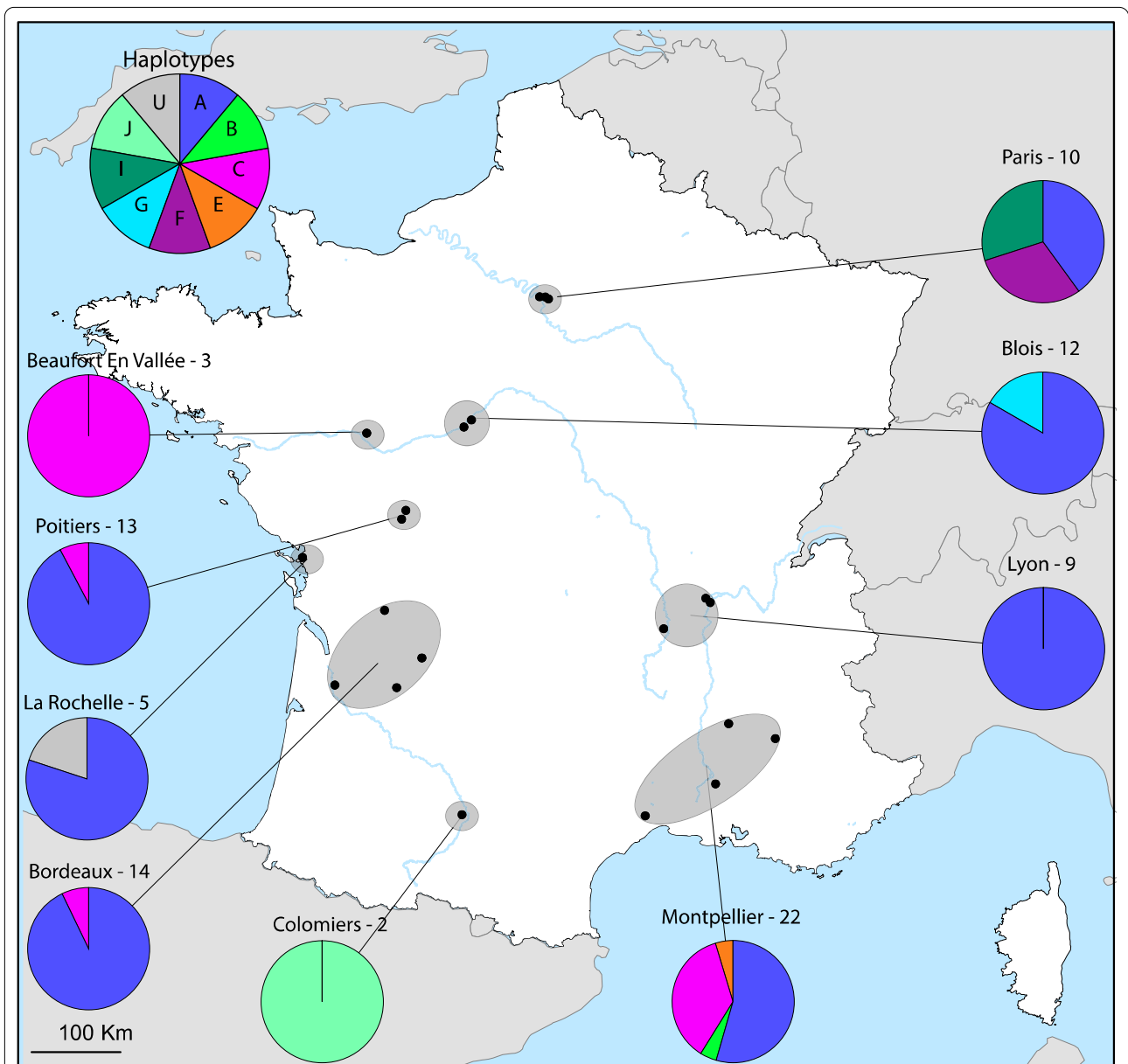
Haplotype A (Fig. 5B1) represents the percentages of differences of the 27 individuals of haplotype A compared with *Exaireta spinigera*. All individuals show the same level of gene difference as those of *Exaireta spinigera*. Only the NAD6 gene presents slight differences.

Haplotype C (Fig. 5B2) represents the percentages of differences of the 22 individuals of haplotype C compared to *Exaireta spinigera*. It shows that most of the mitochondrial genes are identical except for the ATP8 gene, which is different for four individuals (11 Angouleme; 14 Kenya; 19 Kenya, and 15 Kenya). Slight differences are also present in the NAD3 and NAD6 genes.

Haplotype D–F–G–H–I (Fig. 5B3) represents the percentages of differences of the eight other individuals of haplotype C compared with *Exaireta spinigera*. The profiles are slightly different, with a more substantial disparity for the ATP8, NAD1, and NAD5 genes, especially the Mexican one (57). The profiles show the same pattern except for the ATP8 gene of the individual from Mexico (57), which is more conserved.

#### Complete mitochondrial genome phylogeny

The complete mitochondrial genome phylogeny demonstrates the very high genetic divergence among BSF



**Fig. 4** France CO1 haplotype distribution. Distribution of the different haplotypes of *Hermetia illucens* in France. The dots represent the specific area of collection. The pie charts represent the distribution of haplotypes in the corresponding area. The names of the main cities in each area are indicated above, and the number of sequenced individuals. Paris includes Vincennes and Nogent Sur Marnes. Blois includes Valloire Sur Cisse. Lyon includes Saint Marcellin En Forez, Corbas, Champagne Au Mont d'Or and Dardilly. Montpellier includes Avignon, Trescléoux and Dieulefit. Bordeaux includes Lormont La Force and Périgueux. Poitiers includes Saint Georges Lès Baillargeaux

populations found with the CO1 gene analysis (Additional file 7: Fig. S7). The increased sensitivity of the analysis, due to the size of the sequences (13.96 kb, 13 genes), robustly resolves the relationships between different haplotypes identified previously (nodes > 95%). In addition, it allows us to distinguish subgroups within the previously determined haplotypes. Two major divisions emerge within the tree: group C (which includes all commercial

strains) and others. Two subgroups emerge within the C group (11—F—Angouleme; 14 Kenya; 19 Kenya; 15 Kenya and the others).

The other large group contains sequences of A, D, I, G, and H haplotypes. The I G and H haplotypes are related and close to each other. The very cosmopolitan A haplotype is well conserved with limited mitochondrial sequence divergences.

Moreover, although phylogenetically close, the I, G, and H haplotypes are found in very disparate areas (the G haplotype is mainly in Africa, the I in Europe, and the H in Latin America), supporting multiple worldwide introductions.

Strikingly the high level of the conservation of the ATP8 sequence described above in the Mexican sample (haplotype D) does not explain the basal position of the sequence in the phylogeny; indeed, we tested the phylogeny by removing the ATP8 gene, and no change was found (Additional file 6: Fig. S6).

The ATP8 gene alone is not sufficient to explain the appearance of the observed subgroups. However, unlike the CO1 gene, which varies only slightly within this haplotype, the NAD2, NAD3, NAD5, and NAD6 genes are more variable.

The average percentage of identity between haplotypes A and C is about 96.39%. Between haplotypes C and D, 96.54%. Between haplotype C and H, 96.40% and 96.43% between C and I. Between A and D, 97.82% (all results are in Additional file 5: Fig. S5).

#### The divergence time of the different groups

The time tree (Fig. 6) gives us information about the divergence time of the different mitochondrial genomes. The phylogenetic tree presented in Fig. 6 contains both the Bayesian and the ML trees. The two separate trees are presented in Additional file 7. We noted a relatively long divergence time between haplotype C and the others (2.2 million years). The others have diverged more recently (about 1.2 million years). Within the C haplotype, a divergence has been about 300,000 years, while in the other groups, the divergence times are much lower, about 20,000 years for the A haplotype.

The time of divergence with *Exaireta spinigera* has been estimated at 110 million years, confirmed by the literature that gives us 112 million years with the Timetree tool [61].

#### Discussion

*Hermetia illucens* is now considered a species of major interest for the feed industry [62, 63]. Thus, it became urgent to study the genetic diversity within the different populations at different scales. Analyzing the different haplotypes in the same geographical area is a powerful tool for studying the recent history of BSF. Hence, an in-depth study of the genetic variations between the different haplotypes may allow us to resolve better the scenario of the arrival of BSF in a territory.

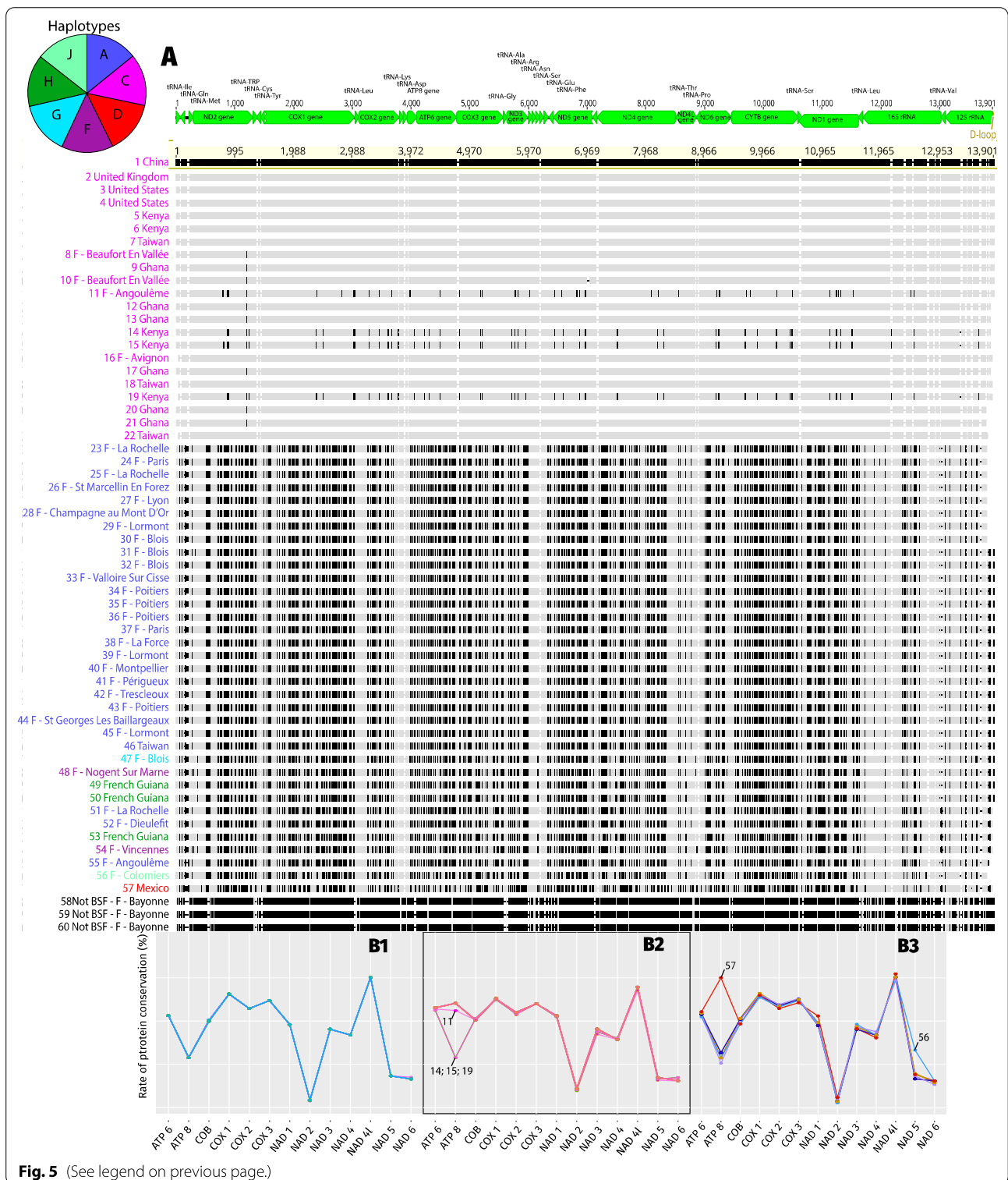
Our study uncovered a high genetic variability of mitogenomes within the species *Hermetia illucens*. We found that haplotype C is the most divergent. The dating of the division between the C haplotype and the others goes back more than 2 million years. We found a well-resolved separation within haplotype C (> 95%); the other haplotypes are also well resolved (> 95%).

Our study confirms the genetic diversity found in previous papers [2, 33, 34]. We found a comparable number of haplotypes, although we added many more sequences of diverse origin but mainly from non-indigenous areas. For example, Ståhls et al. [34] found 56 distinct haplotypes with 418 CO1 sequences, and we found a comparable number of haplotypes based on the analysis of 677 CO1 sequences. These results suggest we have now captured a major part of the extant BSF worldwide diversity outside of the genetic hotspots identified previously [2, 34], namely Latin America. Indeed, in South America, we found 30 haplotypes among the 55 sequences sampled that open up the possibility of a still unknown genetic diversity in this part of the world.

With the maximum diversity found in Latin America, the hypothesis of the geographical origin of BSF [1] is supported by our study. Analysis of the complete mitochondrial genome supports the hypothesis of several distinct introductions on all non-native continents. Considering the estimated divergence times between the different haplotypes, it seems impossible that this is due to a single introduction followed by diversification in the territory.

(See figure on next page.)

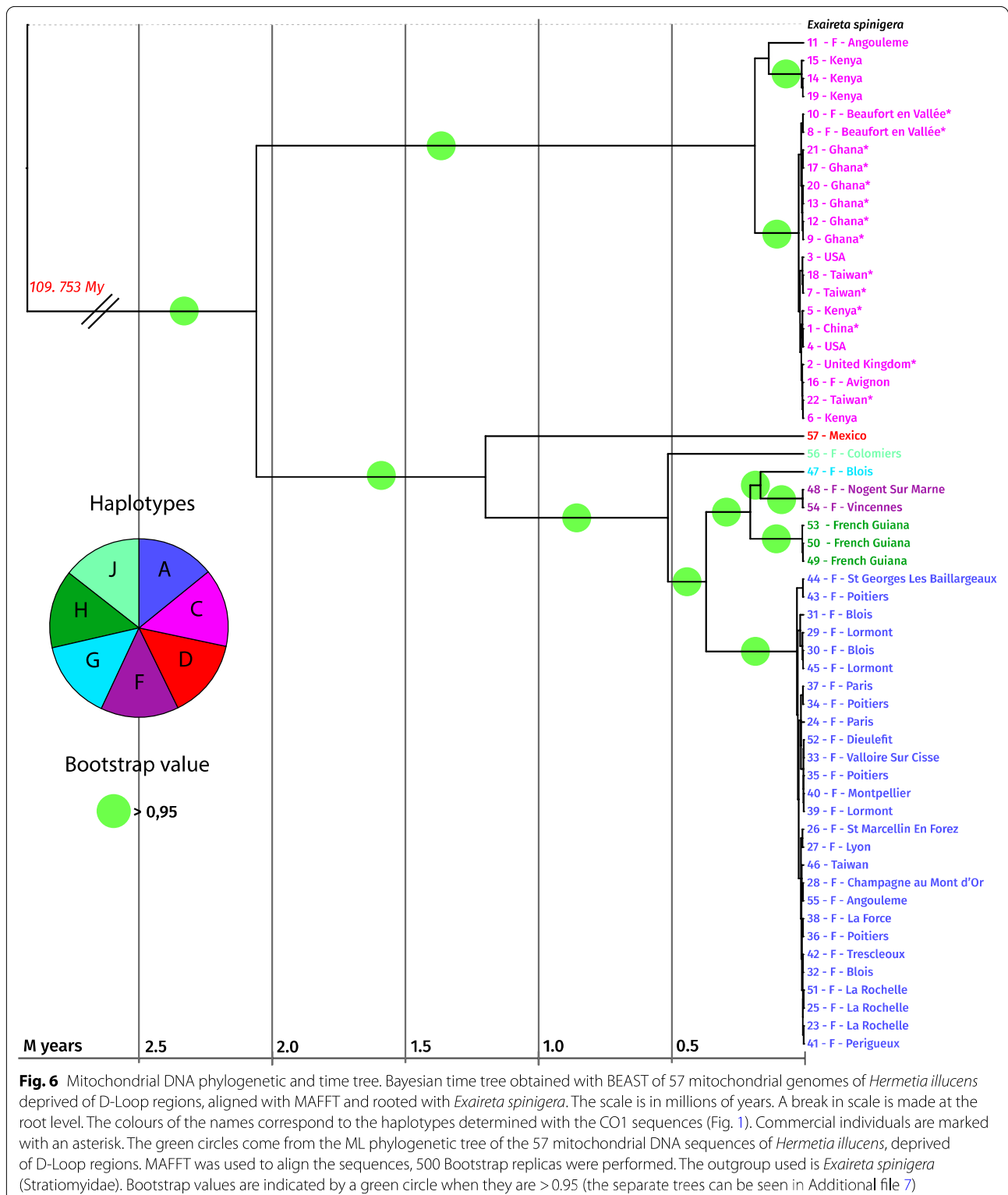
**Fig. 5** Mitochondrial DNA alignment and sequence conservation between *Hermetia illucens* and *Exaireta spinigera*. **A** Multiple alignments of mitochondrial DNA sequences from *Hermetia illucens*. Alignment was performed with the MAFFT aligner. The colours of the names come from their haplotypes determined with the CO1 sequences. Each black mark represents a difference with the reference (1—China). Above the alignment, there is a linear representation of the mitochondrial genes (a circular one can be seen in Additional file 4). The three last sequences are those of *Exaireta spinigera* analysed in the same way as *Hermetia illucens* (we kept one that we used as an outgroup because it was the closest complete mitochondrial sequence phylogenetically that we had). The sequences are numbered to be found in Fig. 6. **B** Sequence conservation of mitochondrial genes of *Hermetia illucens* to *Exaireta spinigera*. The sequences are grouped according to haplotypes. The A haplotype (**B1**), the C haplotype (**B2**), and the other haplotypes (**B3**). The sequences inside each graph are represented in the alignment (5A). Sequences from 23 to 46 plus sequences 51, 52, and 55 for haplotype A (**B1**); from 1 to 22 for haplotype C (**B2**); and from 47 to 57 (except sequences 51, 52, and 55) for the other haplotypes (**B3**)



On the other hand, according to Kaya et al. [2], an alternative but complementary scenario is also possible: bridgehead populations, consisting of a mixture of haplotypes, have been introduced on different continents.

Indeed, taking Europe as an example, the eight haplotypes found in Europe can be explained by the admixture scenario proposed by Kaya et al. [2], who argue that central European BSF populations originated from





admixture between Australian and Southeast African populations, which themselves can be traced back to earlier independent admixture events. Furthermore, Kaya

et al. [2] suggested some gene flows with Mediterranean (altogether originating from Latin America) and particularly East African populations, which interestingly seems

to be corroborated in our study, as exemplified by the F haplotype. It would be extremely useful in the future to combine nuclear data and mitochondrial data to better resolve the potentially diverging patterns of the different markers and to precisely estimate relative impacts of historic genetic admixtures versus recent introductions on regional population structures. France contains on its territory a great diversity of haplotypes, whose origin can be explained either by the historical introduction around the year 1950 [11] in addition to more recent and repetitive introductions. The haplotype A, which is by far the most common haplotype in wild BSF found in France (71.1%) and in Europe, may correspond to the initial introduction. On the other hand, our data suggest that the presence of haplotype C in France may be due to a more recent introduction caused by BSF industries.

As suggested in a recent article on the structure and demography of the BSF based on microsatellite data [2], the captive populations used for breeding in Europe and North America derive from a common, genetically related strain. Our analysis based on CO1 and complete mitochondrial genome sequences confirm this result: the haplotype we named C contains all the sequences of industrial origin with very little genetic diversity. Interestingly, this commercial group does not have any South American relatives. Based on nuclear genetic data, Kaya et al. report wild North American populations being most closely related to the most prevalent managed strains, which we infer to be dominated by haplotype C. This result reinforces the hypothesis of a North American origin of the BSF commercial stock.

Moreover, specific individuals in the wild belonging to this group were collected in the vicinity (around 5 km) of industries breeding BSF. For example, if we focus on the commercial haplotype analysis in France, the wild BSF belonging to haplotype C may have escaped from industries in Poitiers, Bordeaux, Avignon, and Montpellier. This last point raises the question of the biosecurity of BSF farms worldwide. We can assume that the priority of companies involved in BSF farming is not biosecurity and escape control. Indeed, costs in terms of security to prevent insects from escaping are high, and the regulation is late at this level. In that case, the continuous flow of escaping BSF of industrial origin in large numbers could either result in an establishment in territories where it was not yet present or trigger losses of genetic integrity of locally resident wild populations via continuous one-directional introgression. This would lead inexorably to a loss of the local genetic diversity that may prevent future adaptations to changing environmental conditions.

However, it is interesting to note that in Ståhls et al. [34], in contrast to our study, some genetic diversity in the breeding strains in Europe and North America has

been evidenced (10 haplotypes). This may be due to a broader sampling effort of commercial strains. Indeed, Ståhls [34] had the opportunity to sample many small farms leading to 292 BSF sequences from rearing cultures. This allowed them to access a broader genetic diversity not found in our study. It is also possible that smaller farms did not obtain BSF from the dealers who constituted most of the more prominent industries but directly from the wild. Nevertheless, our results confirm that the industrial actors work with the same and unique haplotype.

We also evidence that the mitochondrial genome is ideally suited to resolve long-term BSF phylogenetic relationships. It allows us to date the natural history of the BSF, bringing a sum of details that were impossible to distinguish based only on the CO1. For example, we report puzzling variations in the level of conservation of the ATP8 genes for some samples (haplotypes C and D). Interestingly, these variations have already been found in other insect species (for example, in Coleoptera, see Zhang et al. [63], and in grasshoppers, see Li et al. [64]). The ATP8 gene is small (160 bp on average in insects) and encodes for a subunit of the complex V of the ATP synthase. In *Hermetia illucens*, the initiator codon of the ATP8 gene is ATT, is 167 bp long, and it overlaps with the ATP6 gene. It has been assumed in other species that the modification of the sequence of the ATP8 gene can affect the structure of mitochondrial complex V and thus its function.

Furthermore, it has been shown that changes within ATP synthase can be explained by adaptation to different ecological environments [65–67]. This indicates that some local adaptation driven by the ATP8 genes may exist in BSF, as it has already been shown in other insects. This can lead to an adaptation to survive in high-altitude environments, i.e., at a lower oxygen level and low temperature [68]. Even if these variations did not change the phylogeny of the BSF, it could be a clue to some local adaptation signal. The fact that all commercial strains (plus USA plus Mexico) show a high level of conservation of the ATP8 gene compared to all other wild strains may suggest some adaptation to rearing.

Complete mitochondrial genome data allows us to make a robust and well-resolved time tree of the species. The divergence times between the different haplotypes are surprisingly long (more than 2 million years for the most divergent lineages), and the differences observed between the sequences prove that there is still unsuspected and hidden genomic diversity. Comparing it to *Drosophila*, 2 million years correspond to many speciation events leading to different reproductively isolated species [69]. By contrast, Ståhls et al. [34] indicated that BSF belonging to phylogenetically divergent haplotypes

can still cross and generate offspring, indicating no apparent reproductive isolation in laboratory conditions. Interestingly, our data evidence that the cohabitation of different haplotypes at the same locality could lead to crosses between haplotypes. According to our results, the 2 million years of divergence between the C haplotypes and the other haplotypes could correspond to the North American lineages separating from the ancestral BSF population in South America upon their northwards range expansion beyond the Isthmus of Panama, which appears 2.8 million years ago [70]. This seems especially true since the C haplotype has not been found in South America but only in North America. This ancestral and natural colonization of North America more than 2 million years ago may also explain why the C haplotype group displays a higher degree of mitochondrial genetic diversity than the other groups, a process already observed in other insects as ants for example [71]. Further investigations are thus needed to understand the possible genetic barriers between highly divergent BSF groups but also the possible patterns of introgression in natural BSF populations. This phenomenon might considerably confuse the elucidation of the population structure and the natural history of these species.

Finally, it is interesting to recall that, if indeed, most industrial actors use the same haplotype for the industry, the other haplotypes present in France are highly divergent from this one. Strikingly, we found in one compost bin in the region of Paris flies belonging to two different haplotypes (A and F). Thus, crossbreeding between these two individuals belonging to these two haplotypes would bring a solid genetic mix that could benefit breeding. Associated with this significant genetic diversity, BSF may also have some significant phenotypic diversity that might be useful for specific industrial usage and selection processes to increase performances or ameliorate peculiar traits or behaviors [72].

## Conclusion

Through the analysis of 60 mitochondrial genomes during this study, we proposed a robust and well-resolved time tree of the BSF that confirms the vast genetic diversity within *Hermetia illucens*. Our result elucidates the phylogenetic relationship and time divergence between the BSF lineages. By adding almost 20% of CO1 sequences, mainly from non-native areas, we did not significantly increase the number of haplotypes previously found [34]. This indicates that a significant part of the diversity from non-native areas has already been found and that we should now concentrate our efforts on native areas. Furthermore, the combination of our results with the nuclear analysis already performed [2] gives an interesting insight into the invasion routes of the BSF and the

divergence times between the different groups. We have also shown that many BSF industrial companies work with flies derived from a unique haplotype. We would also like to recall that the apparent biosecurity deficiency within some BSF industrial companies probably led to a continuous flow of escapement and possible local introduction/replacement of natural populations of BSF. Finally, our data indicated that it becomes fascinating to look at the nuclear genomes' evolution of the BSF to better understand the specific adaptations of the different lineages that might be useful for industrial needs and initiate the selection process for peculiar traits.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-022-02025-6>.

**Additional file 1.** List of all individuals used during the study. Origin and accession number.

**Additional file 2.** CO1 gene sequencing protocol. Primers used and Tm.

**Additional file 3.** Phylogenetic tree of the CO1 gene.

**Additional file 4.** Annotated mitochondrial genome of *Hermetia illucens*.

**Additional file 5.** Paired blasts of the mitochondrial genomes used in the study.

**Additional file 6.** Phylogeny based on the complete mitochondrial genome of *Hermetia illucens* with the ATP8 gene sequence removed.

**Additional file 7.** Mitochondrial DNA phylogenetic tree and mitochondrial DNA time tree.

## Acknowledgements

We want to thank Florian Laville and Frederic Marion-Poll for their continuous support and help during this project. We warmly thank the members of the Réseau Compost Citoyen for their precious help during the sampling in France. We also thank Matan Shelomi (National Taiwan University), Meng-Kun Wu, Shu-Min Chen, Jing-Jiun Huang (Yi Mi Community College, Taiwan) for flies from Taiwan. We thank the International Centre of Insect Physiology and Ecology (ICIPE) from Nairobi City, Kenya. We thank Mathieu Chouteau from CNRS LEESIA, in French Guiana. We finally thank Benoit Gilles for his help at the beginning of this project and the two anonymous reviewers for their constructive comments.

## Author contributions

JG carried out the sampling, performed data curation, and wrote the original draft. NP performed data curation and reviewed the manuscript. GB helped with the data curation and reviewed the manuscript. JF conceptualized and supervised the study, performed data curation, and reviewed and finalized the manuscript. All authors read and approved the final manuscript.

## Funding

This work was funded in the framework of a partnership between the CNRS and the CycleFarms company. A CIFRE PhD supports JG. Grant (CIFRE 2018/1095) from the Agence Nationale pour la Recherche et la Technologie (ANRT).

## Availability of data and materials

The dataset generated during the study has been deposited in the European Nucleotide Archive database under Accession Numbers PRJEB48031 Bioprotect. Alignment and raw data can be downloaded following this link: [https://figshare.com/articles/dataset/sequences\\_and\\_trees\\_zip/19589335](https://figshare.com/articles/dataset/sequences_and_trees_zip/19589335). The supplementary can be downloaded following this link: [https://figshare.com/articles/dataset/Supplementary\\_zip/19589332](https://figshare.com/articles/dataset/Supplementary_zip/19589332).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The private company Cycle Farms partially granted this project. The company has not influenced the design and the analysis of the results and has not influenced the writing of the manuscript and its conclusions. No commercial products have been developed based on the outputs of this article. None of the authors has any competing interests.

### Author details

<sup>1</sup>EGCE, CNRS, 12 rue 128, 91190 Gif sur Yvette, France. <sup>2</sup>Cycle Farms, 6 boulevard de entrepreneurs, 49250 Beaufort En Anjou, France. <sup>3</sup>IDEEV - Laboratoire Evolution, Génomes Comportement, Ecologie CNRS Université Paris Sud UMR 9191, IRD UMR 247, 12 rue 128, 91198 Gif sur Yvette, France.

Received: 11 January 2022 Accepted: 13 May 2022

Published online: 01 June 2022

## References

- Hauser M, Woodley M. The historical spread of the black soldier fly, *Hermetia illucens* (L.) (Diptera, Stratiomyidae, Hermetiinae), and its establishment in Canada. *J Kansas Entomol Soc.* 2015;146:51–4.
- Kaya C, Generalovic TN, Ståhls G, Hauser M, Samayoa AC, Nunes-Silva CG, et al. Global population genetic structure and demographic trajectories of the black soldier fly, *Hermetia illucens*. *BMC Biol.* 2021;19(1):94.
- Rozkosný R. A biosystematic study of the European Stratiomyidae (Diptera): volume 2—Clitellariinae, Hermediinae, Pachygasterinae and bibliography. Amsterdam: Springer Netherlands; 1983. (Series entomologica). <https://www.springer.com/gp/book/9789061931355>. Accessed 29 Dec 2020.
- Mason F, Rozkošný R, Hauser M. A review of the soldier flies (Diptera: Stratiomyidae) of Sardinia. *Zootaxa.* 2009;2318(1):507–30. <https://tinyurl.com/4f6nbe2t>.
- Wang Y-S, Shelomi M. Review of black soldier fly (*Hermetia illucens*) as an animal feed and human food. *Foods.* 2017;6(10):91. <https://www.cbd.int/invasive/WhatareIAS.shtml>.
- Picker M, Griffiths C, Weaving A. Field guide to insects of South Africa. Cape Town: Penguin Random House South Africa; 2002. p. 452.
- Lindner E. Die amerikanische *Hermetia illucens* L. im Mittelmeergebiet: (Stratiomyiidae, Dipt.). Akadem. Verlagsges; 1936. p. 2.
- Benelli G, Canale A, Raspi A, Fornaciari G. The death scenario of an Italian Renaissance princess can shed light on a zoological dilemma: did the black soldier fly reach Europe with Columbus? *J Archaeol Sci.* 2014;49:203–5.
- Barbier J. Introduction en France d'un Diptère Stratiomyide américain. *Bull Soc Entomol France.* 1952;57(7):108.
- Dauphin P. Présence de *Hermetia illucens* (LINNÉ, 1758) dans le sud-ouest de la France (Diptera Stratiomyiidae). p. 3.
- Philippe R. Sur la présence d'*Hermetia illucens* (Linnaeus, 1758) (Diptère Stratiomyidae) dans la région lyonnaise/présence of *Hermetia illucens* (Linnaeus, 1758) (Diptera Stratiomyidae) in the region of Lyons. *Bull mensuel de la Société linnéenne de Lyon.* 2009;78:137–8.
- Pierre-Olivier M, Denis R, Willems J. First record of the black soldier fly, *Hermetia illucens*, in the Western regions of France (Vendée, Loire-Atlantique, Ille-et-Vilaine) with notes on its worldwide repartition (Diptera, Stratiomyidae). *Bull Soc Entomol France.* 2020;125:13–8.
- Leclercq M. Dispersion and transport of noxious insects: *Hermetia illucens* L. (Diptera Stratiomyidae). *Bulletin des recherches agronomiques de Gembloux.* 1969. [https://scholar.google.com/scholar\\_lookup?title=Dispersion+and+transport+of+noxious+insects%3A+Hermetia+illucens+L.%28Diptera+Stratiomyidae%29&author=Leclercq%2C+M.&publication\\_year=1969](https://scholar.google.com/scholar_lookup?title=Dispersion+and+transport+of+noxious+insects%3A+Hermetia+illucens+L.%28Diptera+Stratiomyidae%29&author=Leclercq%2C+M.&publication_year=1969). Accessed 9 Nov 2021.
- Craig Sheppard D, Larry Newton G, Thompson SA, Savage S. A value added manure management system using the black soldier fly. *Bioresour Technol.* 1994;50(3):275–9.
- Lord WD, Goff ML, Adkins TR, Haskell NH. The black soldier fly *Hermetia illucens* (Diptera: Stratiomyidae) as a potential measure of human post-mortem interval: observations and case histories. *JFS.* 1994;39(1):215–22.
- Gold M, Tomberlin JK, Diener S, Zurbrügg C, Mathys A. Decomposition of biowaste macronutrients, microbes, and chemicals in black soldier fly larval treatment: a review. *Waste Manag.* 2018;82:302–18.
- Hoc B, Genva M, Fauconnier M-L, Lognay G, Francis F, Caparros MR. About lipid metabolism in *Hermetia illucens* (L. 1758): on the origin of fatty acids in prepupae. *Sci Rep.* 2020;10(1):11916.
- Ewald N, Vidakovic A, Langeland M, Kiessling A, Sampels S, Lalander C. Fatty acid composition of black soldier fly larvae (*Hermetia illucens*)—possibilities and limitations for modification through diet. *Waste Manag.* 2020;102:40–7.
- Giannetto A, Oliva S, Ceccon Lanes CF, de Araújo Pedron F, Savastano D, Baviera C, et al. *Hermetia illucens* (Diptera: Stratiomyidae) larvae and prepupae: biomass production, fatty acid profile and expression of key genes involved in lipid metabolism. *J Biotechnol.* 2019;307:44–54.
- Sprangers T, Ottoboni M, Klootwijk C, Ovyen A, Deboosere S, De Meulenaer B, et al. Nutritional composition of black soldier fly (*Hermetia illucens*) prepupae reared on different organic waste substrates. *J Sci Food Agric.* 2017;97(8):2594–600.
- van Huis A. Edible insects: future prospects for food and feed security (FAO forestry paper). Rome: Food and Agriculture Organization of the United Nations; 2013. p. 187.
- Leong SY, Kuty SRM, Malakahmad A, Tan CK. Feasibility study of biodiesel production using lipids of *Hermetia illucens* larva fed with organic waste. *Waste Manag.* 2016;47(Pt A):84–90.
- Surendra KC, Tomberlin JK, van Huis A, Cammack JA, Heckmann LHL, Khanal SK. Rethinking organic wastes bioconversion: evaluating the potential of the black soldier fly (*Hermetia illucens* (L.)) (Diptera: Stratiomyidae) (BSF). *Waste Manag.* 2020;117:58–80.
- Heuel M, Sandrock C, Leiber F, Mathys A, Gold M, Zurbrügg C, et al. Black soldier fly larvae meal and fat can completely replace soybean cake and oil in diets for laying hens. *Poult Sci.* 2021;100(4): 101034.
- Schiavone A, Dabbou S, Petracci M, Zampiga M, Sirri F, Biasato I, et al. Black soldier fly defatted meal as a dietary protein source for broiler chickens: effects on carcass traits, breast meat quality and safety. *Animal.* 2019;13(10):2397–405.
- Neumann C, Velten S, Liebert F. N balance studies emphasise the superior protein quality of pig diets at high inclusion level of algae meal (*Spirulina platensis*) or insect meal (*Hermetia illucens*) when adequate amino acid supplementation is ensured. *Animals.* 2018;8(10):172.
- Biasato I, Renna M, Gai F, Dabbou S, Meneguz M, Perona G, et al. Partially defatted black soldier fly larva meal inclusion in piglet diets: effects on the growth performance, nutrient digestibility, blood profile, gut morphology and histological features. *J Anim Sci Biotechnol.* 2019;10:12.
- Khamis FM, Ombura FLO, Akutse KS, Subramanian S, Mohamed SA, Fiaboe KKM, et al. Insights in the global genetics and gut microbiome of black soldier fly, *Hermetia illucens*: implications for animal feed safety control. *Front Microbiol.* 2020. <https://doi.org/10.3389/fmicb.2020.01538/full>.
- Erickson MC, Islam M, Sheppard C, Liao J, Doyle MP. Reduction of *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Enteritidis in chicken manure by larvae of the black soldier fly. *J Food Prot.* 2004;67(4):685–90.
- Liu Q, Tomberlin JK, Brady JA, Sanford MR, Yu Z. Black soldier fly (Diptera: Stratiomyidae) larvae reduce *Escherichia coli* in dairy manure. *Environ Entomol.* 2008;37(6):1525–30.
- Park S, Choi H, Choi JY, Jeong G. Population structure of the exotic black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae) in Korea. *Korean J Environ Ecol.* 2017;31(6):520–8.
- Ståhls G, Meier R, Sandrock C, Hauser M, Šašić Zorić L, Laiho E, et al. The puzzling mitochondrial phylogeography of the black soldier fly (*Hermetia illucens*), the commercially most important insect protein species. *BMC Evol Biol.* 2020;20(1):60.
- Qi Y, Xu J, Tian X, Bai Y, Gu X. The complete mitochondrial genome of *Hermetia illucens* (Diptera: Stratiomyidae). *Mitochondrial DNA Part B.* 2017;2(1):189–90.



36. Generalovic TN, McCarthy SA, Warren IA, Wood JMD, Torrance J, Sims Y, et al. A high-quality, chromosome-level genome assembly of the black soldier fly (*Hermetia illucens* L.). bioRxiv. 2020. <https://doi.org/10.1093/g3journal/jkab085>.
37. Zhan S, Fang G, Cai M, Kou Z, Xu J, Cao Y, et al. Genomic landscape and genetic manipulation of the black soldier fly *Hermetia illucens*, a natural waste recycler. *Cell Res*. 2020;30(1):50–60. <https://reseaucompost.org/>.
38. <https://tinyurl.com/46a7xejr>.
39. Desjardins P, Conklin D. NanoDrop microvolume quantitation of nucleic acids. *J Vis Exp*. 2010;45:2565. <https://tinyurl.com/evaah44c>.
40. <https://www.geneious.com>.
41. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
42. Maddison DR, Swofford DL, Maddison WP. NEXUS: an extensible file format for systematic information. *Syst Biol*. 1997;46(4):590–621.
43. Leigh JW, Bryant D. popart: full-feature software for haplotype network construction. *Methods Ecol Evol*. 2015;6(9):1110–6.
44. Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res*. 2010;38(6):1767–71.
45. Hansen KD, Brenner SE, Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res*. 2010;38(12): e131.
46. Erlich Y, Mitra PP, Delabastide M, McCombie WR, Hannon GJ. Alta-cyclic: a self-optimising base caller for next-generation sequencing. *Nat Methods*. 2008;5(8):679–82.
47. Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, et al. Synthetic spike-in standards for RNA-seq experiments. *Genome Res*. 2011;21(9):1543–51.
48. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol*. 2013;20(9):1131–9.
49. Meng G, Li Y, Yang C, Liu S. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualisation. *Nucleic Acids Res*. 2019;47(11): e63.
50. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31(10):1674–6.
51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
52. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol*. 2013;69(2):313–9.
53. Leinonen R, Sugawara H, Shumway M. International nucleotide sequence database collaboration. The sequence read archive. *Nucleic Acids Res*. 2011;39(Database issue):D19–21.
54. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–9.
55. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49(W1):W293–6.
56. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 2018. <https://doi.org/10.1093/ve/vey016>.
57. Abrami G, Mehler A, Lücking A, Rieb E, Helfrich P. TextAnnotator: a flexible framework for semantic annotations. In: Proceedings of the 15th joint ACL-ISO workshop on interoperable semantic annotation (ISA-15). Association for Computational Linguistics, London; 2019. p. 1–12.
58. Paradis E. Analysis of haplotype networks: the randomised minimum spanning tree method. *Methods Ecol Evol*. 2018;9(5):1308–17.
59. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol*. 2015;32(4):835–45.
60. Tomberlin JK, van Huis A. Black soldier fly from pest to 'crown jewel' of the insects as feed industry: an historical perspective. *J Insects Food Feed*. 2020;6(1):1–4.
61. Zhang S, Shu J, Wang Y, Liu Y, Peng H, Zhang W, et al. The complete mitochondrial genomes of two sibling species of camellia weevils (Coleoptera: Curculionidae) and patterns of Curculionini speciation. *Sci Rep*. 2019;9(1):3412.
62. Li X-D, Jiang G-F, Yan L-Y, Li R, Mu Y, Deng W-A. Positive selection drove the adaptation of mitochondrial genes to the demands of flight and high-altitude environments in grasshoppers. *Front Genet*. 2018;9:605.
63. Peng Q, Tang L, Tan S, Li Z, Wang J, Zou F. Mitogenomic analysis of the genus *Pseudois*: evidence of adaptive evolution of morphological variation in the ATP synthase genes. *Mitochondrion*. 2012;12(5):500–5.
64. Gu P, Liu W, Yao Y, Ni Q, Zhang M, Li D, et al. Evidence of adaptive evolution of alpine pheasants to high-altitude environment from mitogenomic perspective. *Mitochondrial DNA Part A*. 2016;27(1):455–62.
65. Zhang Q-L, Zhang L, Zhao T-X, Wang J, Zhu Q-H, Chen J-Y, et al. Gene sequence variations and expression patterns of mitochondrial genes are associated with the adaptive evolution of two *Gynaephora* species (Lepidoptera: Lymantriinae) living in different high-elevation environments. *Gene*. 2017;610:148–55.
66. Luo Y, Yang X, Gao Y. Mitochondrial DNA response to high altitude: a new perspective on high-altitude adaptation. *Mitochondrial DNA*. 2013;24(4):313–9.
67. Russo CAM, Mello B, Frazão A, Voloch CM. Phylogenetic analysis and a time tree for a large drosophilid data set (Diptera: Drosophilidae). *Zool J Linn Soc*. 2013;169(4):765–75.
68. O'Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, Cione AL, et al. Formation of the isthmus of Panama. *Sci Adv*. 2016;2(8): e1600883.
69. Winston ME, Kronauer DJC, Moreau CS. Early and dynamic colonization of Central America drives speciation in Neotropical army ants. *Mol Ecol*. 2017;26(3):859–70.
70. Zhou F, Tomberlin JK, Zheng L, Yu Z, Zhang J. Developmental and waste reduction plasticity of three black soldier fly strains (Diptera: Stratiomyidae) raised on different livestock manures. *J Med Entomol*. 2013;50(6):1224–30.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)



**Titre :** Génétique, génomique et histoire évolutive d'*Hermetia illucens*.

**Mots clés :** Génétique ; *Hermetia illucens* ; évolution ; agro-industrie ; écologie ; bio-informatique.

**Résumé :** *Hermetia illucens*, la mouche soldat noire, est un diptère de la famille des Stratiomyidae. Cet insecte détritivore est de plus en plus utilisé en élevage de masse pour réduire les biodéchets et produire de la nourriture à destination des élevages aquacoles et avicoles. Elle est originaire du nord de l'Amérique Latine et s'est répandue sur une grande partie de la planète. Depuis une dizaine d'années, un nombre croissant d'entreprises élèvent et transforment des larves dans une optique d'économie circulaire. Cette thèse CIFRE a pour but d'améliorer les connaissances génétiques et génomiques sur la mouche soldat noire pour l'amélioration des performances et de la sélection. Nous nous sommes concentrés, dans un premier temps, sur l'obtention et l'analyse d'un génome de référence, pour l'espèce en utilisant du séquençage hybride. Nous avons également travaillé sur la diversité génétique et l'histoire invasive de la BSF. Nous l'avons fait au moyen d'un échantillonnage

mondial en utilisant à la fois le gène CO1 et l'assemblage d'une soixantaine de génomes mitochondriaux complets. Nous avons montré la complexité des scénarios d'invasion de la mouche soldat noire combinant à la fois un processus ancien de dispersion naturelle sur le continent Américain ainsi que, l'importance des introductions humaines, notamment en raison des élevages industriels. Enfin nous nous sommes également intéressés aux variations structurales des génomes, induites par les éléments transposables dans les populations naturelles d'*Hermetia illucens*, afin de déterminer les éléments les plus actifs dans les génomes ainsi que leurs dynamiques évolutives. Ce travail a révélé une étonnante et inattendue diversité génétique et génomique au sein de l'espèce. Il ouvre de nouvelles perspectives vers des expérimentations de sélection et d'amélioration sur la mouche soldat noire pour l'industrie entomocole.

**Title:** Genetic, genomic, and evolutionary history of *Hermetia illucens*.

**Keywords:** Genetics; *Hermetia illucens*; evolution; agro-industry; ecology; bioinformatics.

**Abstract:** *Hermetia illucens*, the black soldier fly, is a dipteran of the family Stratiomyidae. This detritivorous insect is increasingly used in mass production to reduce bio-waste and produce food for aquaculture and poultry farms. It originates from the north of Latin America and has spread over much of the planet. For the last ten years, many companies have been breeding and processing larvae from a circular economy perspective. This CIFRE thesis aims to improve genetic and genomic knowledge on the black soldier fly for performance improvement and selection. We focused, in a first step, on obtaining and analyzing a reference genome for the species using hybrid sequencing. We also worked on the genetic diversity and invasive history of BSF. We did this through global sampling using the CO1 gene and

assembling about 60 complete mitochondrial genomes. We have shown the complexity of the black soldier fly's invasion scenarios combining an ancient process of natural dispersal on the American continent and the importance of human introductions, notably due to industrial breeding. Finally, we also looked at the structural variations of the genomes induced by transposable elements in natural populations of *Hermetia illucens* to determine the most active elements in the genomes and their evolutionary dynamics. This work revealed the species' surprising and unexpected genetic and genomic diversity. It opens new perspectives towards selection and improvement experiments on the black soldier fly for the entomoculture industry.