



HAL
open science

Génomique de la résistance aux insecticides chez la punaise de lit *Cimex lectularius*

Chloé Haberkorn

► **To cite this version:**

Chloé Haberkorn. Génomique de la résistance aux insecticides chez la punaise de lit *Cimex lectularius*. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université Claude Bernard - Lyon I, 2023. Français. NNT : 2023LYO10046 . tel-04560765

HAL Id: tel-04560765

<https://theses.hal.science/tel-04560765>

Submitted on 26 Apr 2024

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.



THÈSE DE DOCTORAT DE L'UNIVERSITÉ CLAUDE BERNARD LYON 1

École doctorale n°341
Évolution, Écosystèmes, Microbiologie, Modélisation

Spécialité de doctorat: Biologie
Discipline: Génomique évolutive

Soutenue publiquement le 31/03/2023, par :

Chloé Haberkorn

Génomique de la résistance aux insecticides chez la punaise de lit *Cimex lectularius*

Devant le jury composé de :

Dr. Cristina VIEIRA-HEDDI , Pr., UCBL Lyon 1, LBBE	Présidente
Dr. Pierrick LABBÉ , Pr., Université de Montpellier, ISEM	Rapporteur
Dr. Gaëlle LE GOFF , CR, INRAE, Institut Sophia Agrobiotech	Rapportrice
Dr. Christoph VORBURGER , Pr., ETZürich, Eawag	Rapporteur
Dr. Claire MÉROT , CR, CNRS, ECOBIO	Examinatrice

Dr. Julien VARALDI , MC, UCBL Lyon 1, LBBE	Directeur de thèse
Dr. Fabrice VAVRE , DR, CNRS, LBBE	Co-directeur de thèse
Dr. Romain LASSEUR , IZInovation	Co-directeur de thèse

“Where do pesticides fit into the picture of environmental disease? We have seen that they now contaminate soil, water, and food, that they have the power to make our streams fishless and our gardens and woodlands silent and birdless. Man, however much he may like to pretend the contrary, is part of nature. Can he escape a pollution that is now so thoroughly distributed throughout our world?”

Rachel Carson, 1962

Résumé

Une forte résurgence de la punaise de lit, *Cimex lectularius*, est observée à l'échelle globale. L'évolution de la résistance aux insecticides est un des phénomènes expliquant cette résurgence, et en particulier concernant les insecticides de la famille des pyréthrinoïdes, la plus utilisée pour lutter contre cet insecte hématophage. Un des enjeux du contrôle de la punaise de lit consiste à identifier les régions génomiques impliquées dans cette résistance. Pour cela, nous avons effectué une analyse de différenciation génétique (génomique des populations) entre deux populations de punaises résistantes et deux populations sensibles aux insecticides. Ces données ont mis en évidence un superlocus contenant de nombreux gènes connus pour être impliqués dans la résistance aux insecticides chez d'autres insectes. Ce superlocus contient de nombreux SNPs fortement différenciés en fonction du statut de résistance des souches, mais également de potentielles duplications en tandem et inversions. Dans un second temps, nous avons réalisé une étude de l'expression des gènes de la punaise de lit, afin de détecter les gènes différentiellement exprimés en fonction du phénotype de résistance et de l'exposition à un pyréthrinoïde. Si un grand nombre de gènes étaient sur-exprimés suite au traitement insecticide, un nombre bien plus restreint a été détecté comme étant régulés différentiellement de manière constitutive entre les souches résistantes et sensibles. Enfin, une forte mortalité a affecté notre élevage de punaises, entravant la réalisation du projet initial. A travers des approches de séquençages d'ADN et de cDNA métagénomiques, nous avons recherché les bactéries et virus potentiellement à l'origine de cette mortalité, dans l'objectif de faciliter les études futures de ce modèle, et de détecter un potentiel pathogène de cet insecte.

Abstract

A strong resurgence of the common bed bug, *Cimex lectularius*, is observed on a global scale. The evolution of resistance to insecticides is one of the phenomena explaining this resurgence, and in particular concerning insecticides of the pyrethroid family, the most commonly used to control this hematophagous insect. One of the challenges for controlling bed bugs is to identify the genomic regions involved in this resistance. Hence, we performed a genetic differentiation analysis (population genomics) between two insecticide-resistant and two insecticide-susceptible populations of bed bugs. These analyses highlighted a superlocus containing numerous genes known to be involved in insecticide resistance in other insects. This superlocus contains many SNPs that are highly differentiated according to the resistance status of the strains, but also potential tandem duplications and inversions. In a second step, we conducted a gene expression study of the bed bug, in order to detect genes differentially expressed according to the resistance phenotype and exposure to a pyrethroid. While a large number of genes were overexpressed following insecticide treatment, a much smaller number were detected as constitutively differentially regulated between resistant and susceptible strains. Finally, a high mortality rate affected our bedbug rearing, hindering the realisation of the initial project. Through metagenomic DNA and cDNA sequencing approaches, we searched for bacteria and viruses potentially responsible for this mortality, with the objective of facilitating future studies of this model, and detecting a potential pathogen of this insect.

Table des matières

Résumé	4
Abstract	5
Table des figures	11
Liste des tableaux	13
1 Introduction générale	23
1.1 La punaise de lit, un modèle d'étude cauchemardesque	26
1.1.1 Un grand voyage : des grottes à la résurgence moderne	26
Biogéographie des Cimicidae	26
Hôtes successifs	26
Association à l'humain	28
Facteurs de recrudescence	29
Dynamique contemporaine	30
1.1.2 Portrait d'un parasite redoutablement efficace	31
<i>Cimex</i> versus <i>Homo</i>	32
Cycle de vie	34
Hématophagie	35
Cortège bactérien et viral	35
Caractéristiques du génome	36
1.1.3 Élevage en laboratoire	37
1.1.4 Les méthodes de contrôle	39
La traque	39
Lutte physique et mécanique	40
Lutte naturelle	40
(Més)usages des insecticides conventionnels	42
1.2 Problématique de la résistance aux pyréthri-noïdes	45
1.2.1 Les insecticides pyréthri-noïdes	45
1.2.2 Détecter la résistance	46
1.2.3 Définir la résistance	47
1.2.4 Mécanismes de résistance	48
Comportementale	48
Cuticulaire	49
Métabolique	50
Mutation de cible	53
Autres mécanismes	54
1.2.5 Résistances : un pour tous et tous pour un	56
Résistance croisée	56
Résistance polygénique	57
1.2.6 Coût de la résistance	58
Impact sur la valeur sélective	58
Différents coûts pour différentes mutations	59
Mutations de compensation	60
1.3 Génomique de l'adaptation	61
1.3.1 Bases génétiques de l'adaptation	61

Variants ponctuels	62
Variants structuraux	63
1.3.2 Détection des signatures adaptatives	66
1.4 Objectifs et organisation de la thèse	70
1.4.1 Projet initial	70
1.4.2 Réorientation du sujet	72
1.4.3 Plan de la thèse	73
2 Génomique de la résistance aux insecticides chez la punaise de lit	75
ARTICLE 1 : Un superlocus majeur de 6 Mb est impliqué dans la résistance aux pyréthriinoïdes chez la punaise de lit <i>Cimex lectularius</i> .	78
3 Réponse transcriptomique à l'exposition aux pyréthriinoïdes	115
ARTICLE 2 : Réponse transcriptomique au traitement pyréthriinoïdes chez des souches de punaises de lit résistantes et sensibles.	118
4 Métagénomique infectieuse	141
4.1 Contexte - encore une sombre histoire d'élevage	142
4.2 Exploration de l'ADN hors punaise de lit	144
4.2.1 Assemblage et analyse des contigs	144
4.2.2 Identification taxonomique	146
4.3 Analyse des punaises de lit au phénotype oxydé	148
4.3.1 Extraction et séquençage des acides nucléiques viraux putatifs	149
Insectes	149
Broyage et filtration	149
Digestion des acides nucléiques de <i>C. lectularius</i>	150
Extraction des acides nucléiques restants	150
Digestion enzymatique des échantillons obtenus	151
4.3.2 Assemblage et alignement des contigs	153
4.3.3 Identification taxonomique	154
Contigs ADN	154
Contigs ARN	157
4.4 Validation de candidats	159
4.4.1 Préparation des échantillons	159
4.4.2 Réalisation de la PCR	160
4.4.3 Migration sur gel d'agarose	161
4.5 Discussion & Perspectives	163
4.5.1 Potentiels agents pathogènes détectés chez la punaise de lit . .	163
4.5.2 Exploration des bases de données SRA	164
4.5.3 Gestion de la mortalité en laboratoire	165
4.6 Matériel supplémentaire	168
5 Discussion générale	169
5.1 Résumé des principaux résultats de la thèse	170
5.2 Distribution des candidats associés au phénotype de résistance le long du génome	174
5.2.1 Un superlocus impliqué dans la résistance aux pyréthriinoïdes .	174
5.2.2 Superlocus en auto-stop ou évolution d'un supergène?	176
5.2.3 Organisation du génome	180
5.3 Mécanismes de résistance aux pyréthriinoïdes chez <i>C. lectularius</i> . .	184
5.3.1 Rares SNPs non-synonymes	185
Importance de la mutation <i>kdr</i>	185

Variations des séquences codantes de gènes induits par l'exposition aux pyréthrinoïdes	186
5.3.2 Régulation transcriptionnelle	188
Différences constitutives d'expression entre les souches résistantes et sensibles	188
Quid du cluster de P450 détecté sur le superlocus?	190
5.3.3 Régulation post-transcriptionnelle	191
Explorer l'implication des ARN non-codants dans la résistance aux insecticides	192
Des ARNnc dans le supergène?	193
5.4 Comprendre pour faciliter la lutte de demain	195
5.4.1 Développement de tests diagnostics	195
5.4.2 La recherche dans de beaux draps!	198
5.5 Conclusion générale	200

Bibliographie	201
----------------------	------------

Table des figures

1.1	Phylogénie de la famille des punaises de lit (Cimicidae).	27
1.2	La punaise de lit et l'humain au fil des siècles.	29
1.3	Cycle de vie de <i>Cimex lectularius</i> .	34
1.4	Système Hemotek pour l'élevage des punaises de lit.	38
1.5	Exemples de solutions naturelles de lutte contre les punaises de lit.	41
1.6	Structure des pyréthriinoïdes de type I et II.	45
1.7	Exemple de méthodologie d'exposition des punaises de lit aux insecticides.	47
1.8	Mécanismes hypothétiques de résistance aux insecticides chez la punaise de lit.	48
1.9	Mécanismes métaboliques entraînant la résistance aux insecticides.	51
1.10	Mutations de cible entraînant la résistance aux insecticides.	53
1.11	Voies de régulation transcriptionnelle de l'expression des gènes de détoxification chez les insectes.	55
1.12	Différents types de variants structuraux.	63
1.13	Signaux de détection d'un variant structural.	64
1.14	Impact des balayages sélectifs sur la variation génétique.	68
1.15	Projet initial : design des croisements et protocole de détection des SNPs par RAD-seq.	70
1.16	Extraits d'une vidéo prise sur le terrain lors d'un échantillonnage de punaises de lit.	72
4.1	Apparence des punaises de lit touchées par le phénomène de mortalité anormal au sein des élevages.	143
4.2	Pipeline de traitement des reads pour la détection d'organismes hors punaise de lit.	145
4.3	Tailles cumulées des contigs détectés pour chaque espèce au sein des données pool-seq, hors alignement sur le génome de <i>C. lectularius</i> .	146
4.4	Interactions décrites entre le nématode <i>Steinernema</i> , la bactérie <i>Alcaligenes</i> et la larve de lépidoptère <i>Galleria</i> .	147
4.5	Migration des échantillons d'ARN viral et ADN viral sur gel d'agarose.	151
4.6	Migration des échantillons d'ADN génomique et d'ADN viral sur gels d'agarose, avec ou sans digestion enzymatique.	152
4.7	Algorithme MMseq2 d'attribution taxonomique.	153
4.8	Diagramme de Sankey des contigs ADN détectés.	154
4.9	Tailles cumulées des contigs détectés pour chaque espèce au sein des données ADN suite à la purification virale.	157
4.10	Diagramme de Sankey des contigs ARN détectés.	158
4.11	Tailles cumulées des contigs détectés pour chaque espèce au sein des données ARN suite à la purification virale.	159
4.12	Migration sur gel des produits de PCR pour différentes amorces : <i>Cimex lectularius</i> , <i>Wolbachia</i> , et BEV-like proteobacterium.	161
4.13	Migration sur gels des produits de PCR pour différentes amorces : BEV-like proteobacterium, <i>Alcaligenes faecalis</i> et réovirus.	162
4.14	Transmission horizontale du virus de l'hypertrophie des glandes salivaires chez <i>G. pallidipes</i> .	166
4.1	Intensité de l'échantillon en fonction de sa taille.	168

5.1 Schéma de l’articulation des différents chapitres présentant les travaux réalisés dans le cadre de cette thèse.	170
5.2 Schéma bilan du chapitre 2 - analyses génomiques.	171
5.3 Schéma bilan du chapitre 3 - analyses transcriptomiques.	172
5.4 Résultats extraits des travaux de Fountain <i>et al.</i> (2016)	174
5.5 Variété de phénotypes alternatifs pouvant être régulés par les supergènes d’après Schwander <i>et al.</i> (2014)	177
5.6 Haplotypes putatifs (non-exhaustifs) du superlocus inversé pouvant (co-)exister par rapport au génome de référence, en fonction de la présence de l’allèle muté <i>kdr</i> du <i>VGSC</i>	182
5.7 Impact transcriptionnel des différents mécanismes potentiellement impliqués dans le phénotype de résistance aux insecticides.	184
5.8 Architecture putative du cluster de cytochromes P450 détecté.	191
5.9 Test diagnostic du supergène putatif.	195
5.10 Propagation des punaises de lit entre foyers suivant leur statut de résistance.	196

Liste des tableaux

4.1	Concentration et nombre de reads séquencés pour les échantillons d'ARNv et ADNv extraits des punaises de lit présentant un phénotype oxydé.	152
4.2	Concentration des échantillons d'ADN génomique et d'ARN total extraits des punaises de lit présentant un phénotype sain ou oxydé.	160
4.3	Résumé des principales espèces d'intérêt à génome ADN détectées dans les données métagénomiques.	164
4.1	Amorces utilisées pour les PCR de validation des virus et bactéries détectées dans les données métagénomiques.	168

Remerciements

Merci à mon trinôme d'encadrants, en commençant par **Julien Varaldi** : merci de m'avoir fait confiance pour ce projet et de m'avoir accompagnée au bout. J'ai appris à compter sur ton efficacité redoutable et tes expertises variées ! Merci de m'avoir considérée comme une vraie jeune chercheuse, et de m'avoir toujours poussée à aller gratter un peu plus loin. Merci **Fabrice Vavre** pour ta bienveillance, et ta capacité presque magique à sentir quand ça n'allait pas trop et te rendre instantanément disponible pour faire le point. Merci de m'avoir poussée dans mes retranchements sur ces questions d'adaptation, avec toujours ce regard global tellement difficile à aller chercher quand on a le nez dans le guidon. Enfin, merci **Romain Lasseur** de m'avoir fait découvrir le monde de l'entreprise, et aidée à prendre confiance en moi dans cette jungle en me proposant de nombreux projets : j'en ressors grandie. J'ai beaucoup appris à votre contact à tous les trois, merci d'avoir accepté de faire ce chemin avec moi.

Merci aux membres du jury qui ont accepté d'évaluer cette thèse : **Gaëlle Le Goff**, **Pierrick Labbé**, **Christoph Vorburger**, **Claire Mérot** et **Cristina Vieira** – votre présence m'a honorée. Merci pour vos rapports et les discussions très intéressantes qui en ont découlé.

Merci aux membres de mon comité de suivi pour votre disponibilité, vos encouragements, et vos commentaires, un savant mélange d'expertises et de bienveillance – **Laurent Duret**, **Tristan Lefébure**, **Isabelle Amat**, **Jean-Philippe David**, et mon tuteur **Guillaume Minard**, quelle chance de vous avoir eu. Un merci un peu particulier à Jean-Philippe qui m'a prise sous son aile pour mes premiers stages de recherche au LECA, et qui m'a fait l'honneur de continuer à suivre mes travaux, entre autres en faisant partie de ce comité. Merci à tous les trois, avec **Thierry Gaude** et **Julien Cattel**, de m'avoir fait découvrir le monde de la recherche, et de m'avoir encouragée à le parcourir un peu plus.

Merci à l'**équipe GEI**, et à tou.te.s les collègues qui font vivre ce groupe, en particulier **Natacha Kremer** et au tout nouveau gang des punaisologues, **Jean-Marie Delpuech** qui m'a guidée pour réaliser mes premières expositions aux insecticides au LBBE, et **Sylvain Charlat** avec qui on aura beaucoup discuté de

plein de choses différentes (même de maths)!

Merci également aux équipes du pôle administratif, pôle informatique et pôle biotechnologique, sans qui la recherche n'existerait tout simplement pas. Merci à **Hélène Henri**, la maman des thésard.e.s de ce labo, que dis-je, la queen du multitasking, qui arrive à trouver le temps d'offrir une oreille attentive. Merci à **David Lepetit** d'avoir partagé les secrets de la purification virale avec moi.

Merci à toute l'équipe de l'entreprise IZInovation qui m'a accueillie. Un merci tout particulier à **Christine Lasseur** qui s'emploie à créer un environnement de travail chaleureux, **Aurélien Guicherd** pour m'avoir accueillie au labo IZI en attendant le Symbiotron, **Jérémy Soret** pour avoir réalisé entre autres la vidéo du repas de sang, et **Lucile Rametti** qui a suivi ma thèse de très près (et plus encore!).

Merci aux ami.e.s de l'équipe : **Zaïnab** (mais si tu comptes dans l'équipe!!), faire des manips avec toi a été un plaisir, merci pour ton aide si précieuse et ta gentillesse à toute épreuve, **Sylvain**, toujours là pour nous derrière ton air sarcastique, **Camille**, d'avoir organisé toutes ces petites choses qui nous permettent d'avoir cette belle cohésion de groupe, **Sara** pour ta bonne humeur constante, **Vincent** pour tes multiples conseils de bio-informaticien aguerri, **Pierre** d'avoir été un si bon papa-thésard pour nous (quelle patience), **Marius** avec qui on a pu partager nos déboires sur ce charmant petit insecte, **Anaïs**, et tou.te.s les autres! Mais il y a aussi des gens incroyables en dehors de l'équipe (c'est fou) – **Théo**, qui fait clairement partie des meubles et est toujours prêt à donner un coup de main (de manière très littérale aussi) pour grimper plus haut, **Mary** pour ta douceur et les meilleurs conseils de tatoueurs, **Émilie**, **Alex**, **Mélodie**, **Thibault**, **Gaspard**, et tou.te.s les autres! Un merci tout particulier aussi aux acolytes qui ont relu des morceaux (parfois imposants) de thèse – **Beber**, **Julien J.**, **Benjamin**, vos conseils ont vraiment été importants dans la toute dernière ligne droite. Un petit coucou aussi aux personnes incroyables rencontrées à l'ESEB 2022 à Prague, avec vous l'avenir de la recherche est entre de très bonnes mains.

Merci à toute la bande du lycée de Brochon : **Gillien** (tu seras très vite un docteur incroyable), **Dele**, **Pierrick**, **Chloé Dom**, **Pauline**, **Julien**, et bien sûr mon **Amélia**, j'ai tellement de chance de vous avoir encore dans ma vie! Merci à mon duo préféré de la fac de bio de Dijon, **Manon & Antoine**, si vous saviez comme je vous aime. Merci aux autres auto-proclammés blaireaux et blairelles du master de Grenoble, et tout particulièrement à **Théo**, j'espère que tu trouveras le même à la hauteur, et à ma **Billur**, notre amitié m'est si chère.

Merci aux ami.e.s de la CO et surtout **Solène**, **Florence**, et toute la bande de l'Adoc (gros gros cœur sur vous). Merci à ma traileuse préférée **Paula** bien sûr, même en étant physiquement de l'autre côté du monde tu n'es finalement jamais bien loin.

Merci à mes colocs, dans l'ordre d'apparition - **Samson**, pour ta bonne humeur constante, tes associations de légumes hasardeuses, et m'avoir partagé ton obsession plutôt saine pour les trucs à deux roues sans moteurs ; **Clémence**, pour ta passion, ton énergie, et nos grandes discussions politiques tisanes en main.

Merci à mon coloc de bureau, **Benjamin** : 3 ans et demi sans se lâcher d'une semelle, l'aventure aurait été bien différente sans toi. Merci d'avoir été un binôme de thèse aussi incroyable, avec toujours un bon conseil ou une vidéo improbable sous le coude pour faire redescendre la pression.

Merci à toute la famille d'Olivier ! À **Christine** et **Pierre**, merci pour tout, et bravo pour m'avoir supportée pendant les confinements et de multiples week-ends.

Merci à toute ma famille pour leur soutien, les moments de détente plus ou moins sportifs partagés. À mes parents **Valérie** & **Bruno** et à mon grand-petit frère **Guilhem** : vous êtes mon socle. Je ne serais jamais arrivée jusque là sans vous, merci pour votre patience et vos encouragements dans tout ce que j'entreprends.

Merci **Olivier**, tu sais déjà tout étant donné que tu as la *chance* de suivre cette aventure au jour le jour (presque heure par heure, vu que je parle beaucoup). J'ai du mal à concevoir que sur six années ensembles, plus de la moitié auront été pendant cette thèse. Merci pour ton soutien indéfectible et notre amour !

Enfin, merci à toutes celles et ceux que je n'ai pas cité directement, ou que j'ai momentanément oublié !

Laissez-moi vous raconter une histoire. Imaginez une chauve-souris, fouettant l'air de ses ailes membraneuses à l'aurore, cherchant un abri. Imaginez maintenant un insecte, troublant le repos de cette chauve-souris au fond d'une grotte obscure en se repaissant de son sang. Imaginez maintenant un humain, attiré dans cette même grotte pour se cacher d'un prédateur ou d'une tempête - moins poilu, le sang tout aussi chaud, et la promesse d'une surface comestible autrement plus élevée. Imaginez cet humain, au fil des millénaires, se construire un foyer, le chauffer, s'offrir le confort d'un matelas. Prendre le bateau, le train, l'avion ; dormir dans des auberges, à l'hôtel, en AirBnb. Et toujours ce parasite avec lui, qu'il dissémine partout. Appliquer des plantes réduites en poudre, mettre le feu, asphyxier au Zyklon B, détruire au DDT. Et encore ce parasite avec lui. Recommencer. Est-ce que ça ne fonctionne plus... ?

Alors, ça commence à vous démanger ? Laissez-moi vous raconter une histoire de punaises de lit !



Photo très contractuelle - © C.Haberkorn

Avant-propos

Cette thèse a pu voir le jour grâce à un partenariat entre le LBBE (Laboratoire de Biométrie et Biologie Évolutive), et la société IZInnovation, réunis autour du projet Idex Micro-be-have¹ avec quatre autres laboratoires partenaires, spécialistes des relations insectes-microorganismes. La thèse bénéficie d'un financement CIFRE, pour Convention Industrielle de Formation par la Recherche. Ce dispositif, créé en 1981, “subventionne toute entreprise de droit français qui embauche un.e doctorant.e pour le/la placer au cœur d'une collaboration de recherche avec un laboratoire public. (...) Les Cifre sont intégralement financées par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation qui en a confié la mise en œuvre à l'ANRT (Association Nationale Recherche Technologique)”².

La société IZIGroup qui a soutenu cette thèse propose trois grands types d'activités : un centre de tests et une plateforme d'innovation (IZInnovation), un centre de formations techniques, réglementaires et d'inspection de sites (IZIpest), et une agence marketing (Influence Pest Control). Ainsi, l'entreprise partenaire ne commercialise pas de produits, et n'a pas d'intérêt financier direct lié à la présente thèse. Je tenais donc à réaffirmer en introduction l'absence de conflit d'intérêt ayant pu influencer ces travaux.

1. <https://microbehave.universite-lyon.fr>

2. <https://www.anrt.asso.fr/fr/cifre-35654>

1

Introduction générale

1.1 La punaise de lit, un modèle d'étude cauchemardesque	26
1.1.1 Un grand voyage : des grottes à la résurgence moderne . . .	26
Biogéographie des Cimicidae	26
Hôtes successifs	26
Association à l'humain	28
Facteurs de recrudescence	29
Dynamique contemporaine	30
1.1.2 Portrait d'un parasite redoutablement efficace	31
<i>Cimex</i> versus <i>Homo</i>	32
Cycle de vie	34
Hématophagie	35
Cortège bactérien et viral	35
Caractéristiques du génome	36
1.1.3 Élevage en laboratoire	37
1.1.4 Les méthodes de contrôle	39
La traque	39
Lutte physique et mécanique	40
Lutte naturelle	40
(Més)usages des insecticides conventionnels	42
1.2 Problématique de la résistance aux pyréthrinoïdes . . .	45
1.2.1 Les insecticides pyréthrinoïdes	45
1.2.2 Détecter la résistance	46
1.2.3 Définir la résistance	47
1.2.4 Mécanismes de résistance	48
Comportementale	48
Cuticulaire	49
Métabolique	50
Mutation de cible	53
Autres mécanismes	54
1.2.5 Résistances : un pour tous et tous pour un	56
Résistance croisée	56
Résistance polygénique	57

1.2.6 Coût de la résistance	58
Impact sur la valeur sélective	58
Différents coûts pour différentes mutations	59
Mutations de compensation	60
1.3 Génomique de l'adaptation	61
1.3.1 Bases génétiques de l'adaptation	61
Variants ponctuels	62
Variants structuraux	63
1.3.2 Détection des signatures adaptatives	66
1.4 Objectifs et organisation de la thèse	70
1.4.1 Projet initial	70
1.4.2 Réorientation du sujet	72
1.4.3 Plan de la thèse	73

La punaise de lit est un insecte rampant de la grande famille des Cimicidae, qui se nourrit exclusivement du sang de ses victimes - chauves souris, oiseaux, ou humains. Deux espèces en particulier parasitent l'humain : la punaise de lit tropicale, *Cimex hemipterus*, et la punaise de lit commune, *Cimex lectularius* (Usinger, 1966). Si la punaise de lit a toujours fait partie du cortège de parasites connus chez l'humain, sa résurgence récente la place plus que jamais sous le feu des projecteurs. Cette augmentation est imputée à la mondialisation, au marché de la seconde main, mais également à un phénomène de résistance aux produits appliqués : les insecticides chimiques.

Au cours de cette thèse, nous nous concentrerons sur la punaise de lit commune, *Cimex lectularius*, afin de décrypter les bases génomiques de la résistance aux insecticides de la famille des pyréthriinoïdes, mais également d'identifier de potentiels pathogènes naturels, en vue de proposer de nouvelles pistes de lutte plus respectueuses de l'environnement.

Dans un premier temps, nous suivrons l'histoire de la dissémination de la punaise de lit, de son changement d'hôte à l'époque des cavernes à sa résurgence moderne. Nous décrirons ensuite ce modèle d'étude par le prisme de ses interactions avec l'humain, du cycle de vie, de l'hématophagie, du cortège bactérien et viral qui l'accompagne et des connaissances générales de son génome. Après avoir passé en revue les différentes techniques d'élevage, nous ferons ensuite l'inventaire des méthodes de contrôle à notre disposition. Pour ce faire, nous décrirons les techniques de lutte physique et chimique, des poudres miracles mystérieuses aux insecticides pyréthriinoïdes modernes.

Dans un second temps, nous reviendrons sur la problématique de la résistance aux pyréthriinoïdes. Après avoir décrit les insecticides de cette famille, le phénomène de résistance sera défini de manière exhaustive, ainsi que les principaux mécanismes sous-jacents. Les différents types de résistance, mais également les coûts associés et les compensations potentielles seront également abordés.

Enfin, la notion d'adaptation redeviendra centrale, en explorant son évolution d'un point de vue génomique. Les origines des mécanismes adaptatifs, mais également leurs bases génétiques seront abordées, en nous attardant plus particulièrement sur les mutations ponctuelles ainsi que les variants structuraux.

1.1 La punaise de lit, un modèle d'étude cauchemardesque

La punaise de lit est parfaitement adaptée au mode de vie humain, ce qui semble rendre particulièrement difficile le fait de s'en séparer. Pourquoi, et comment, en sommes-nous arrivés là ?

1.1.1 Un grand voyage : des grottes à la résurgence moderne

L'histoire de la punaise de lit commence bien avant la nôtre. Au sein de l'ordre des Hemiptera, la famille des Cimicidae regroupe une centaine d'espèces, réparties en une vingtaine de genres. Ces genres se déclinent en fonction de leurs hôtes principaux - chauve-souris, oiseau, humain - et de leur répartition géographique.

Biogéographie des Cimicidae

Les premiers fossiles de Cimicidae ont été datés d'au moins 100 millions d'années (Engel, 2008). Cela a été confirmé par une phylogénie moléculaire calibrée à partir des données fossiles, qui estime la branche des Cimicidae à 115 millions d'années (Roth *et al.*, 2019). La punaise de lit *C. lectularius* fait partie de la sous-famille des Cimicinae, inféodée au Paléarctique, qui englobe l'Europe, l'Afrique du Nord, et l'Asie au nord de l'Himalaya (Figure 1.1, en vert). Les Cimicinae auraient un ancêtre commun avec les *Afrocimicinae*, une sous-famille inféodée à l'Afrique (Figure 1.1, en rose), et seraient arrivés en Europe lors de la collision entre les plaques Africaines et Européennes, il y a 50 millions d'années.

Hôtes successifs

Au sein des hétéroptères dont font partie les Cimicidae, on retrouve des insectes généralistes, principalement phytophages, qui auraient évolué vers l'hématophagie suite à des repas de sang accidentels (Weirauch *et al.*, 2019). La chauve-souris (Chiroptère) a longtemps été présumée être l'hôte primitif des Cimicidae, jusqu'à ce que les plus anciens fossiles des Chiroptères soient relégués environ 50 millions d'années après les plus anciens fossiles des Cimicidae (Simmons *et al.*, 2008).

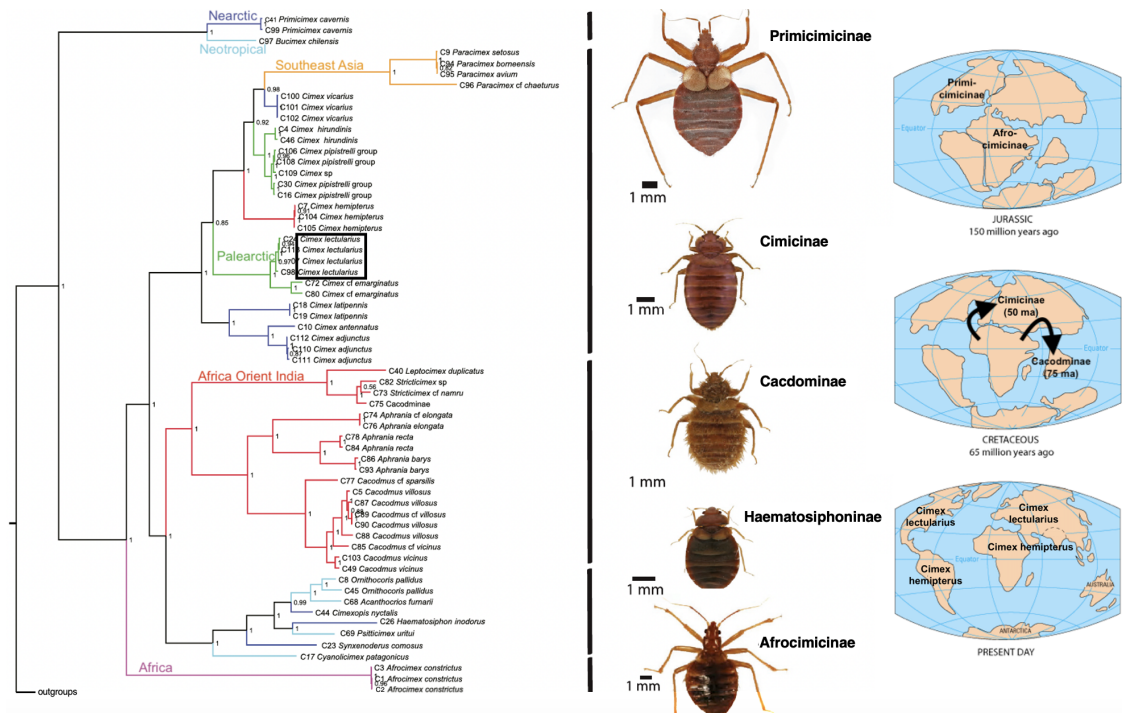


Figure 1.1 – Phylogénie de la famille des punaises de lit (*Cimicidae*) (Adapté de Roth et al. (2019)). À gauche, un arbre de consensus Bayésien basé sur les séquences de cinq gènes, montrant la distribution biogéographique (couleurs des branches) et la taxonomie à l'échelle des sous-familles. A côté de chaque nœud, la valeur de probabilité postérieure est indiquée. L'échelle de longueur des branches représente le nombre de substitutions nucléotidiques estimées par site. La punaise de lit commune, *Cimex lectularius*, est encadrée en noir. Chaque sous-famille est représentée par la photo d'un spécimen. À droite, la distribution des sous-familles d'intérêt de *Cimicidae* sur les continents au Jurassique et au Crétacé, et la répartition des espèces de punaises de lit communes (*C. lectularius*) et tropicales (*C. hemipterus*) aujourd'hui.

Grâce à la reconstruction phylogénétique de Roth et al. (2019), il semblerait que les Cimicidae étaient des hématophages spécialistes, dont certaines lignées ont par la suite évolué vers des comportements plus généralistes au fil des opportunités. Les quatre grandes lignées ancestrales semblent en effet toutes avoir été des parasites spécialistes des chauves-souris. Si la datation des plus anciens fossiles de Chiroptères est exacte (postérieure à l'apparition des Cimicidae donc), cela implique que le parasitisme spécifique des chauves-souris ait évolué plusieurs fois de manière indépendante. Néanmoins, l'hôte ancestral antérieur à l'apparition des Chiroptères n'est pas connu.

Au sein des Cimicidae, seules quelques espèces peuvent parasiter les humains, dont deux principales : *Cimex lectularius*, la punaise de lit commune, et *Cimex hemipterus*, inféodée aux zones tropicales. Toujours d'après ces récents travaux, si l'ancêtre commun de *C.lectularius* et *C.hemipterus* parasitait spécifiquement les chauves souris, la lignée dont provient *C.hemipterus* aurait changé une première

fois d'hôte en parasitant les oiseaux, avant de se diversifier de nouveau, *C.hemipterus* étant une espèce considérée comme généraliste. *C.lectularius* est également une espèce généraliste, capable de parasiter la majorité des oiseaux (Galloanseres et Neoaves, les Chiroptères (chauves-souris), et les humains (Roth *et al.*, 2019). Sur le terrain, certains professionnels ont ainsi observé des punaises de lit qui passaient d'un appartement au nid de pigeons bisets (*Columba livia*) à proximité, par les interstices de la fenêtre.

Concernant la divergence entre *C. lectularius* et *C. hemipterus*, les études divergent également : si l'hypothèse d'Ashford (2000) prédit qu'elle a eu lieu il y a 1,6 million d'années, coïncidant avec la séparation entre les clades de l'*Homo sapiens* et de l'*Homo erectus*, Roth *et al.* (2019) ont finalement daté cette divergence à 47 millions d'années. Or, le plus ancien fossile du genre *Homo* découvert à ce jour à été daté de 2,8 millions d'années (Villmoare *et al.*, 2015), soit bien après cette séparation. Les Cimicidae auraient donc parasité plus d'une fois, et de manière indépendante, l'humain. Bien que les circonstances de la première rencontre entre *Cimex* et *Homo* ne soient pas connues, une hypothèse principale a émergé et semble faire consensus.

Association à l'humain

D'après Usinger (1966), les humains auraient été parasités pour la première fois par des punaises de lit dans des grottes de la région méditerranéenne (Figure 1.2). Bien plus tard, à la fin de la préhistoire, la sédentarisation aurait permis aux punaises de lit de s'installer de manière pérenne avec l'humain, passant d'une maison à l'autre. La plus ancienne trace de punaise de lit associée à l'humain date de plus de 3 500 ans, et a été découverte dans une tombe égyptienne située près du Caire (Panagiotakopulu and Buckland, 1999). Chauffer les demeures fait aussi partie des facteurs ayant permis aux punaises de lit de prospérer à l'année. D'après les écrits, d'abord présente uniquement en Eurasie, la punaise de lit a colonisé le monde, cachée dans le repli d'un vêtement ou la doublure d'un sac de voyage, accordant un peu plus de répit aux contrées insulaires, moins accessibles. Les punaises de lit seraient par exemple arrivées en Angleterre seulement au 16^e siècle, 300 ans après avoir infesté la France (Potter, 2011). Certains "exterminateurs" s'y sont fait un nom, et ont même publié les premiers guides pour se débarrasser de la punaise de lit (Southall, 1730).

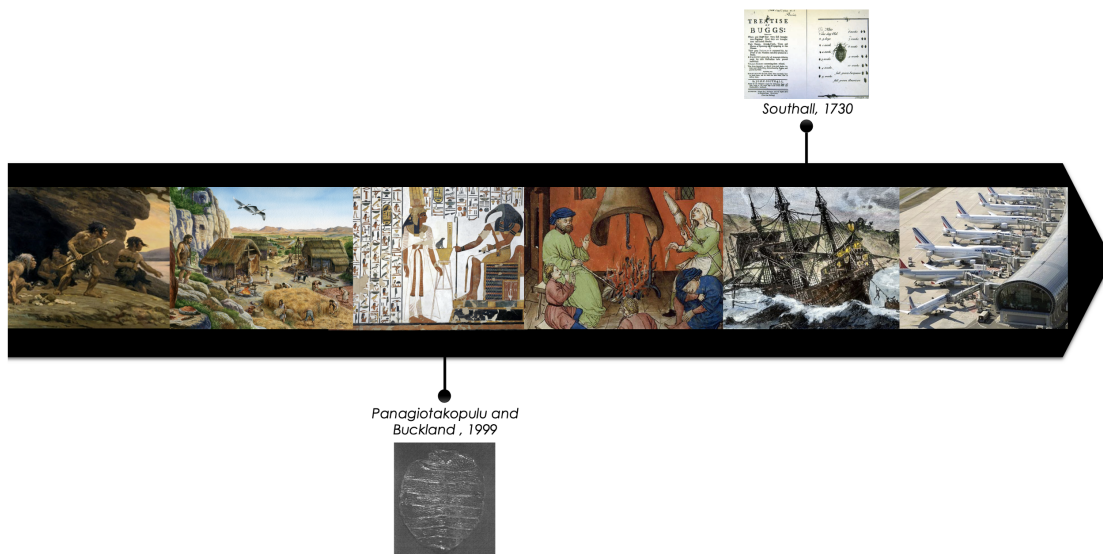


Figure 1.2 – *La punaise de lit et l’humain au fil des siècles. Descriptions des illustrations : 1 - L’Homme préhistorique, devant une grotte dans laquelle il aurait pu rencontrer la punaise de lit; 2 - Sédentarisation humaine; 3 - Tombeau égyptien, époque à laquelle les premières punaises de lit associées à l’humain ont été figées; 4 - Chauffage dans une habitation au Moyen-Âge; 5 - Conquête maritime permettant la propagation insulaire de la punaise comme en Angleterre, où de premiers guides de lutte dont celui de Southall sont édités; 6 - Aéroport, témoin de la mondialisation en partie responsable de la résurgence de cet insecte.*

Facteurs de recrudescence

Avec la mondialisation, les flux de personnes et de biens n’ont jamais été aussi intenses, facilitant la propagation de nombreuses espèces invasives. À cela s’ajoutent d’autres conditions propices à la résurgence de la punaise de lit, telles que l’essor du marché de la seconde main (meubles anciens avec pléthore de recoins sombres, vêtements, linge de maison), et la résistance aux insecticides (Doggett *et al.*, 2004; Romero *et al.*, 2007).

Depuis les années 90, les punaises de lit sont en pleine recrudescence dans la majorité des pays occidentaux, tels que les États-Unis (Zhu *et al.*, 2010), l’Australie (Doggett *et al.*, 2004), le Canada (Hwang *et al.*, 2005), de nombreux pays d’Europe et la Russie (Naylor *et al.*, 2018).

En France, le nombre d’infestations de punaises de lit aurait augmenté de 76% en 2020, d’après la Chambre Syndicale 3D. Un sondage IPSOS a comptabilisé 7% de la population, soit 4,7 millions de Français-e-s, comme ayant été touché-e-s entre 2016 et 2021 par les punaises de lit (Laffargue *et al.*, 2021). Les enquêtes se multiplient dans les médias : l’UFC Que Choisir y a même consacré un numéro

spécial, entre sondages et tests d'efficacité¹. Pour 1/4 des lecteurs-trices ayant eu une infestation de punaise de lit, celle-ci n'était pas terminée, alors même que pour près de la moitié de ces victimes l'infestation durait depuis plus de trois mois. Les infestations décrites touchaient principalement des appartements (69%), et la région de l'île de France (43%). Lorsque les personnes connaissaient l'origine de l'infestation, il s'agissait en majorité des voyages (40%). Une personne sur deux a fait appel à une société spécialisée pour s'en débarrasser, qui utilisera des insecticides dans 91% du temps, pour un coût moyen de 1207€.

Un plan d'action interministériel de lutte contre les punaises de lit a été lancé en février 2022², avec pour objectifs (i) d'informer pour éviter les infestations ou savoir les identifier de manière précoce, (ii) de donner accès à une liste d'entreprises qualifiées, ainsi que (iii) la mise en place d'un numéro d'information et d'assistance juridique, notamment en cas de conflit avec le bailleur. En effet, la loi a également évolué, imposant au bailleur de prendre en charge les frais de détection et de désinfestation, en se référant à l'obligation de décence du logement (article 142 de la loi ELAN, 2018). Toujours dans le cadre de ce plan d'action, un groupe de travail "Punaises de lit" a été constitué par l'ANSES en août 2021, à l'initiative de la Direction générale de la Prévention des risques et la Direction générale de la Santé³. L'objectif est de réaliser une expertise scientifique des caractéristiques des punaises de lit, des facteurs de prolifération, des méthodes de lutte, et des bénéfices et risques associés à ces méthodes.

Dynamique contemporaine

Permettre au grand public de comprendre les mécanismes menant à une infestation de punaises de lit est crucial pour lui permettre de l'éviter. En effet, *C. lectularius* peut se déplacer de manière passive, grâce à son hôte : accrochée à ses vêtements posés sur un fauteuil à l'hôtel, à l'intérieur d'un matelas revendu sur un site de seconde main, dans son sac de voyage... Mais la punaise de lit peut également se propager de manière active au sein d'un bâtiment, entre deux chambres voire appartements par des trous dans les murs, des conduites d'eau ou encore d'anciens vide-ordures (Booth *et al.*, 2012). Concernant les hôtels, la formation des agents d'entretien est primordiale pour éviter la contamination de

1. Voir les résultats du sondage en ligne : <https://www.quechoisir.org/actualite-punaises-de-lit-infographie-vos-solutions-pour-en-venir-a-bout-n89858/>

2. Site gouvernemental du plan d'action : <https://www.ecologie.gouv.fr/punaises-lit-letat-vous-accompagne>

3. Voir l'appel à candidatures : https://www.anses.fr/fr/system/files/APC_Punaisesdelit_roleetmission.pdf

plusieurs chambres *via* leur matériel ou le linge (How and Lee, 2010).

Ces mécanismes de dispersion vont directement impacter et façonner la structure des populations de *C. lectularius*. En effet, le parasitisme de l'humain facilite à la fois la colonisation de nouvelles zones (patches) vides, et l'extinction d'autres au bout d'une lutte (souvent) acharnée. La plupart des études montrent une diversité génétique très faible au sein d'une infestation, soutenant l'hypothèse d'un très petit nombre de colonisateurs à l'origine de chaque infestation (Booth *et al.*, 2012; Saenz *et al.*, 2012; Fountain *et al.*, 2014). On parle alors d'effet de fondation (Le Corre and Kremer, 1998). En revanche, la diversité inter-populations mesurée dans ces mêmes études est très élevée.

Une étude menée sur 14 populations de punaises de lit échantillonnées dans un hôtel et plusieurs appartements individuels en France a également montré une diversité génétique très faible au sein des populations (Akhoundi *et al.*, 2015). Cette faible diversité intra-population suggère également que chaque colonisation a débuté avec un faible nombre d'individus, voire une seule femelle fécondée. La diversité génétique était plus élevée pour les populations échantillonnées au sein de l'hôtel, et pourrait s'expliquer par plusieurs épisodes d'introduction, ou de dispersion entre les chambres.

Colonisatrice hors pair, la punaise de lit semble donc capable de résilience lorsqu'un petit groupe d'individus crée une nouvelle sous-population, malgré les effets de la consanguinité. Pour comprendre la mise en place de la dynamique actuelle qui fait le succès de ce parasite, il faut se pencher sur ses caractéristiques physiologiques ainsi que ses traits d'histoire de vie, contribuant à la production et à la survie de ses descendants.

1.1.2 Portrait d'un parasite redoutablement efficace

Aujourd'hui, la punaise de lit, comme son nom l'indique, est un parasite qui est le plus fréquemment associé à la chambre à coucher. En effet, son mode vie nocturne est parfaitement adapté à notre état statique de repos, et peut engendrer de nombreux désagréments.

Cimex versus *Homo*

Cimex lectularius est un parasite qui profite de la vulnérabilité de son hôte endormi pour se nourrir de son sang. Les dommages pour l'être humain peuvent être physiques, du fait de la piqûre en elle-même et de la réaction allergique qu'elle peut engendrer, mais également psychologiques.

Les piqûres de punaises de lit sont souvent alignées, ce qui les rends assez caractéristiques. Des cas d'infection des piqûres, mais également de réactions allergiques allant jusqu'à l'œdème de Quincke dans de très rares cas ont été relevés (Delaunay *et al.*, 2010). Pour certaines fortes infestations, un trop grand nombre de piqûres peut également provoquer une anémie chez l'hôte (Paulke-Korinek *et al.*, 2012). Une étude américaine récente a montrée que les patients ayant une infestation de punaises de lit à leur domicile étaient plus susceptibles d'être anémiques (59,5% vs 36,9%) que les autres patients (Sheele *et al.*, 2021).

De nombreuses études ont été également réalisées afin de comprendre si, à l'instar d'autres insectes hématophages, comme les moustiques (Achee *et al.*, 2019), la punaise de lit pouvait transmettre des pathogènes à l'humain. La compétence vectorielle en laboratoire (soit la capacité à recevoir, multiplier et retransmettre un pathogène) est étudiée séparément de la capacité vectorielle (reproduire la compétence vectorielle *in natura*). Plusieurs micro-organismes pouvant déclencher des maladies chez l'humain ont été détectés dans des populations de punaises de lit sur le terrain (Delaunay *et al.*, 2011). C'est le cas de *Coxiella burnetii*, à l'origine de la fièvre Q ou coxiellose chez l'humain. Bien que la transmission entre stades ait été observée, ainsi que cette bactérie ait été identifiée dans les fèces des insectes infectés, la capacité des punaises de lit à être vecteur de *C. burnetii* doit encore être étudiée (Dařter, 1963). Les mêmes observations ont été faites pour le virus de l'hépatite B, sans que des preuves formelles d'une transmission à l'humain n'existent (Ghoda and Shah, 2005). Enfin, la présence de *Trypanosoma cruzi*, responsable de la maladie de Chagas, a été détectée dans des populations de terrain (Villagrán *et al.*, 2008). *T. cruzi* serait capable de se transmettre entre punaises de lit de différents stades, de se répliquer à l'intérieur de l'insecte et d'être ensuite détecté dans sa salive et ses fèces, arguments soutenant une potentielle capacité vectorielle (Delaunay *et al.*, 2011). Néanmoins, aucun cas avéré de punaise de lit transmettant *T. cruzi* à l'humain n'a été observé. En réalité, la punaise de lit serait un bien piètre vecteur dans les appartements, se nourrissant toujours des mêmes hôtes. Néanmoins,

l'impact pourrait être tout autre dans le cas d'une population détectée au sein d'un hôtel.

Aux désagréments physiques peuvent s'ajouter de réels dommages psychologiques suite à une infestation de son domicile. Une étude compilant des témoignages en ligne rapporte que, dans plus de 80% des cas, les personnes infestées ont un ou plusieurs symptômes associés au syndrome de stress post-traumatique. Parmi ces symptômes, on retrouve des flashbacks de l'événement traumatique (ici, l'infestation) *via* des souvenirs intrusifs ou des cauchemars, une hypervigilance, des insomnies, ou encore de l'anxiété et de la paranoïa (Goddard and de Shazo, 2012). Des comportements d'isolation sociale, à cause des stigma et la honte ressentie, ont également été relevés (Burrows *et al.*, 2013). La honte est liée au fait qu'une infestation de punaises de lit, de la même manière que les poux par exemple, est associée à tort à un manque d'hygiène. Si un nettoyage fréquent - et plus spécifiquement d'endroits stratégiques tels que derrière le lit - peut permettre la détection précoce d'une infestation, la saleté n'attire pas les punaises, contrairement à d'autres nuisibles qui pourraient se repaître de restes de nourriture, tels que les cafards. Ainsi, des beaux quartiers aux logements les plus précaires, tout le monde peut être touché par la punaise de lit. Néanmoins, le coût du traitement ou l'absence d'agent d'entretien formé à la détection précoce peuvent être des facteurs discriminants les foyers les plus pauvres.

Des maladies psychiatriques déclenchées par la présence de punaises de lit, comme la dépression, ont également été décrites (Goddard and de Shazo, 2012). Le cas le plus grave à ce jour décrit le suicide d'une femme atteinte d'un trouble bipolaire aux USA, 6 semaines après le début de l'infestation de son appartement. Dans son dernier mail, elle écrivait *"Je suis sûre que les vampires sont de retour et je ne peux pas supporter de vivre dans la peur d'être dévorée vivante"* (Burrows *et al.*, 2013).

En effet, les cas d'infestations non ou mal traitées peuvent rapidement sembler insurmontables. La punaise de lit a un temps de génération très court (environ 40 jours), lui permettant de se multiplier en un temps record : on considère que l'introduction d'une seule femelle préalablement fécondée dans un appartement peut être à l'origine d'une population florissante en seulement quelques mois. Pour mieux comprendre ce potentiel invasif, il nous faut donc s'attacher à comprendre le cycle de vie de la punaise de lit.

Cycle de vie

La punaise de lit est un insecte hétéroptère, ayant 5 stades juvéniles très semblables au stade adulte. Le passage à chaque stade suivant se fait nécessairement suite à un repas de sang, *via* une mue. Les juvéniles sont aisément repérables à l'œil nu de part leur taille, mais également leur coloration, qui passe d'une transparence quasi totale de la cuticule à une opacité brun foncé.

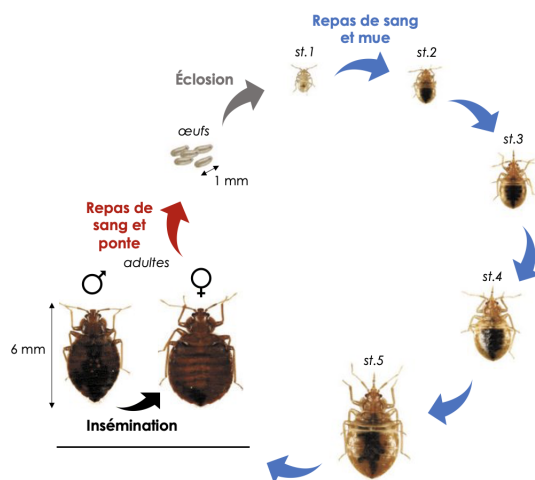


Figure 1.3 – Cycle de vie de *Cimex lectularius*.

Une fois la dernière mue - dite imaginale - réalisée, un dimorphisme sexuel marqué est alors relevé : les femelles sont plus grosses et ont l'extrémité de l'abdomen rond, tandis qu'il est pointu chez les mâles. Mâles et femelles auront encore besoin de faire des repas de sang réguliers, qui permettront la ponte chez les femelles. Néanmoins, une punaise de lit dans les conditions adéquates peut survivre plus d'un an sans se nourrir. Bien que le record de survie en laboratoire sans repas de sang soit de 580 jours (Johnson, 1940), des études plus récentes montrent qu'il s'agirait plutôt d'une centaine de jours pour les punaises des souches sensibles, et d'une cinquantaine pour les souches résistantes aux insecticides (Polanco *et al.*, 2011b).

Lors de l'accouplement, les mâles pratiquent une insémination dite traumatique, en perforant l'abdomen de la femelle avec leur appareil génital externe. En réalité, cette perforation doit être réalisée au niveau d'une incurvation d'un des sternites (en face ventrale donc) de la femelle, située juste au dessus du mésospermagène. De là, le sperme migrera aux ovaires, où la fertilisation aura lieu. Cette reproduction a un coût pour les femelles, en diminuant leur longévité (Stutt and Siva-Jothy, 2001).

Environ une semaine après la fécondation, les femelles commenceront à pondre

entre 5 et 15 œufs par jour. La longueur de chaque étape du cycle peut varier en fonction des conditions environnementale, mais l'éclosion ainsi que les mues successives s'enchaînent toutes les semaines avec des repas de sang réguliers, pour une durée totale du cycle (d'œuf à œuf) d'une quarantaine de jours.

Hématophagie

La punaise de lit est un insecte lucifuge qui se nourrit exclusivement de sang (hématophagie stricte). Sa croissance et sa reproduction dépendent donc de ces repas de sang, réalisés sur ses proies lors d'attaques nocturnes.

Les pièces buccales de la punaise de lit sont parfaitement adaptées à leur régime hématophage. La piqûre est réalisée grâce au rostre piqueur, composé de stylets capables de perforer la peau, encerclant les canaux alimentaires et salivaires. La salive injectée permet d'anesthésier la proie, avant que le sang ne soit aspiré via le canal alimentaire. Une femelle adulte prélève un repas sanguin moyen de 7 μL en une quinzaine de minutes, doublant son poids corporel.

En plus des repas de sang, la punaise de lit a également besoin de vitamine B pour assurer son développement et sa reproduction. Cette vitamine, absente du sang de vertébrés, est produite par un symbiote obligatoire : la bactérie *Wolbachia* (Hosokawa *et al.*, 2010).

Cortège bactérien et viral

De nombreux insectes dépendent d'une symbiose mutualiste pour survivre, une relation dont chaque partie engagée tire des bénéfices. Chez *C. lectularius*, un peu comme une auberge, où l'on payerait pour le gîte, la bactérie *Wolbachia* évolue dans des cellules dédiées appelées bactériocytes (le gîte), et fournit comme monnaie d'échange la vitamine B, essentielle au bon développement de l'hôte.

Les bactériocytes sont des cellules situées à l'intérieur d'un organe permettant l'accueil de potentiels symbiotes appelé bactériome. Ces structures sont proches des gonades, facilitant leur transmission verticale. En effet, la bactérie *Wolbachia* a été identifiée dans les ovaires des femelles *C. lectularius*, puis dans les oocytes, et enfin dans la région du futur organe symbiotique chez l'individu nouvellement formé (Hosokawa *et al.*, 2010). En l'absence de *Wolbachia* suite à l'administration d'antibiotique, les travaux d'Hosokawa ont également montré une diminution de la fécondité des punaises, compensée par un apport externe en vitamine B. Une

γ -protéobactérie BEV-like de la famille des Enterobacteriaceae, ayant un rôle présumé de symbiote secondaire, a également été décrite (Meriweather *et al.*, 2013). Enfin, une bactérie du genre *Rickettsia* a été détectée chez certaines populations de punaises de lit (Potts *et al.*, 2020; Thongprem *et al.*, 2020). Des études approfondies seront nécessaires pour comprendre leurs implications respectives. Il semble néanmoins hautement improbable que leur rôle soit vital, étant donné que ni l'une ni l'autre n'a été détectée chez tous les individus échantillonnés.

Contrairement aux bactéries, très peu d'études ont été menées sur le virome de la punaise de lit à notre connaissance. Une analyse en microscopie a permis de mettre en évidence la présence d'un réovirus dans le tube digestif de la punaise de lit (Eley *et al.*, 1987). Plus récemment, un séquençage ARN de plusieurs punaises a également permis de mettre en évidence un réovirus du genre *Orbivirus*, ainsi qu'un *Picornaviridae* (Stanojević *et al.*, 2020). Néanmoins, on peut supposer que de nombreux autres virus et/ou bactéries peuvent infecter les punaises de lit. Caractériser la diversité de ces micro-organismes potentiellement entomopathogènes pourrait permettre, à plus long terme, d'ouvrir de nouvelles pistes de bio-contrôle.

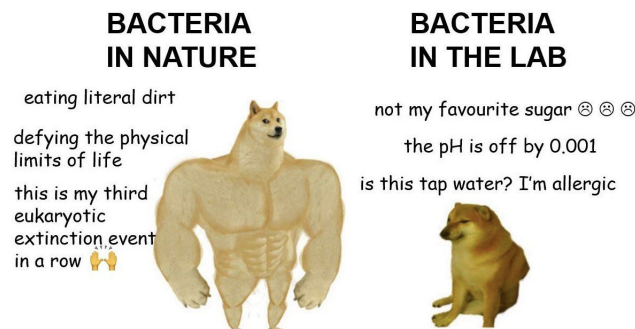
Caractéristiques du génome

Le génome de la punaise de lit a été assemblé pour la première fois en 2016 (Rosenfeld *et al.*, 2016; Benoit *et al.*, 2016). La version RefSeq la plus récente du génome (décembre 2017) a été réalisée dans le cadre de l'initiative ik5, qui visait à séquencer le génome de 5000 arthropodes (<https://i5k.nal.usda.gov>). Il s'agit de la version Clec_2.1, ou GCF_000648675.2, utilisée dans le cadre de cette thèse (Thomas *et al.*, 2020). Ce génome de 510,8 Mb est composé de 1574 régions génomiques (1573 régions d'ADN nucléaire et une d'ADN mitochondrial), appelées scaffolds (assemblage de contigs, eux même composés des lectures de séquençage appelées reads). La taille médiane de ces scaffolds est de 1,6 Mb (N50), et 14,647 gènes y sont annotés. Ce génome a été réalisé à partir d'un séquençage en lectures courtes (bibliothèque "mate-pair") d'un individu mâle de la souche Harlan, provenant du premier élevage de *C. lectularius*. La mise en place de cet élevage remonte à 1973, lorsque Harold Harlan, un entomologiste américain, décida d'échantillonner les insectes infectant un baraquement militaire, puis de les nourrir sur son bras. La souche Harlan fait désormais office de référence sensible, n'ayant pas été en contact avec les insecticides modernes. Les efforts de séquençage de ce génome doivent cependant être poursuivis, pour passer d'un assemblage en scaffolds à un

assemblage en chromosomes.

Le nombre de chromosomes attendus est connu, grâce à plusieurs études ayant réalisé des caryotypes de la punaise de lit (Grozeva *et al.*, 2010; Sadílek *et al.*, 2013). Ainsi, le caryotype le plus courant chez les mâles de *C.lectularius* semble être $2n=26+X_1X_2Y$, avec néanmoins une grande variabilité quant au nombre de chromosomes sexuels, qui peuvent être fragmentés pour un total de 29 à 47 chromosomes (étude réalisée sur $n=61$ punaises de lit) (Sadílek *et al.*, 2013). La détermination sexuelle chez la punaise de lit suit le système XY, où le sexe masculin est codé par des gènes présents sur le chromosome Y, chromosome qu'ils sont les seuls à posséder. Chez la femelle, on décrit le caryotype complémentaire comme étant $2n=26+2X_{1-n}$, avec n le nombre de chromosomes X additionnels.

Toutes ces études génomiques s'appuient sur des insectes prélevés sur le terrain, puis potentiellement élevés en laboratoire pour plus de praticité. Néanmoins, réussir un élevage d'insectes (ou de tout autre organisme) n'est jamais chose aisée, malgré des conditions artificielles contrôlées et supposément idéales.



Mème, source : internet

1.1.3 Élevage en laboratoire

La punaise de lit est un insecte qui prolifère très facilement dans les habitations, et qui peut théoriquement être élevé en laboratoire de manière relativement aisée. Plusieurs paramètres doivent néanmoins être considérés, dont les facteurs physiques (température, hygrométrie, photopériode) et physiologiques (fréquence des repas et type de sang) (Cannet *et al.*, 2015).

Les conditions physiques idéales pour minimiser le temps de génération (et donc maximiser la prolifération des punaises de lit) sont de 26°C, 70% d'humidité relative, et des cycles jour/nuit de 12h chacun. Concernant le repas de sang, deux principales méthodes se dégagent : le gorgement sur être vivant, ou sur du sang

déjà prélevé et chauffé artificiellement. La convention Européenne pour la protection des animaux vertébrés utilisés à des fins expérimentales (CETS 123) stipule que la seconde option doit être privilégiée. Ainsi, dans le cadre de cette thèse, le système Hemotek®⁴ (Figure 1.4), a été choisi. L'Hemotek permet de chauffer le sang à 37°C afin de mimer la température corporelle, et de varier les sangs et les membranes utilisées. Les punaises de lit sont placées dans des bocaux en plastique transparent fermés par de l'organza, un tissu au maillage suffisamment fin pour ne pas laisser d'œufs ou de stades 1 s'échapper. Chaque bocal est ensuite placé sur un réservoir de l'Hemotek remplis de sang et scellé avec du parafilm (Montes *et al.*, 2009). Les punaises vont alors pouvoir se nourrir en piquant successivement à travers l'organza puis le parafilm dans le réservoir⁵. Le sang utilisé est du sang humain qualifié (caractérisation du groupe sanguin et recherche de maladies ou agents transmissibles préalable), prélevé dans le cadre d'un partenariat avec l'Établissement Français du Sang (CODECOH approuvé par le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation). L'ajout d'un anti-coagulant au sang est nécessaire (Feldlaufer *et al.*, 2018), et l'ESF utilise à ce titre de l'héparine de sodium (170 UI soit 0,03 mL d'héparine de sodium dans 10 mL de sang).

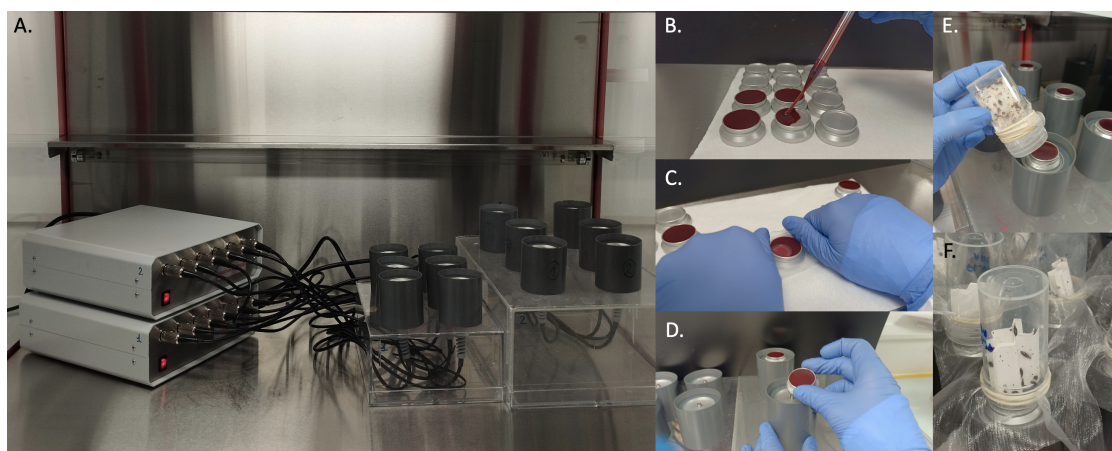


Figure 1.4 – *Système Hemotek pour l'élevage des punaises de lit. A. Deux systèmes centraux Hemotek, reliés à 12 unités chauffantes. B. Remplissage des réservoirs de sang avec une pipette stérile à usage unique (volume : 0,5 µL). C. Scellage d'un réservoir avec du parafilm étiré. D. Installation du réservoir sur une unité chauffante en le vissant. E. Le pot contenant les punaises est placé avec la face fermée par l'organza sur le réservoir. F. Les punaises peuvent prélever le sang dans le réservoir à travers l'organza puis le parafilm. Photos : Élodie Rosinski (LBBE) et Chloé Haberkorn.*

J'ai été la première à monter un élevage de punaises de lit au sein du LBBE. Après quelques mois de croissance, cet élevage s'est brusquement effondré, avec

4. <http://hemotek.co.uk>

5. Pour visualiser un gorgement de punaises de lit réalisé au laboratoire : <https://www.youtube.com/watch?v=zZISi7ySzwY>

un phénotype de mortalité caractéristique "d'oxydation" suite au repas de sang. Ce phénotype est décrit dans le Chapitre 4, et l'hypothèse d'un agent infectieux potentiellement responsable de cette mortalité y est explorée. Cette piste pouvant mener à la découverte de nouvelles possibilités de lutte biologique contre la punaise de lit, elle est particulièrement intéressante. En effet, les produits de lutte utilisés contre les punaises de lit sont en grande majorité dommageables, à la fois pour l'humain et pour l'environnement. Il semble donc primordial d'identifier de nouvelles méthodes de bio-contrôle qui seraient spécifiques, c'est à dire sans effet sur les organismes non-cibles (en incluant l'humain) et efficaces.

1.1.4 Les méthodes de contrôle

Il existe de nombreuses méthodes applicables à la lutte contre la punaise de lit, aux degrés d'efficacité variables. À la lutte mécanique se sont jointes les molécules naturelles, puis les molécules synthétisées chimiquement. Mais avant de traiter, il est nécessaire de s'assurer que le nuisible présent est bien la punaise de lit.

La traque

Le premier contrôle est donc visuel et olfactif. La punaise de lit ayant un comportement grégaire, elle se regroupe avec ses semblables durant la journée. Mues, œufs ou punaises de lit doivent être recherchés à proximité des lits et canapés (sous les matelas, sommiers, rideaux, plinthes, tapis...). Après avoir digéré son repas de sang, la punaise produit également des déjections sous forme de petites traces rondes noires très reconnaissables. Néanmoins, le contrôle visuel peut être insuffisant ou difficile, notamment à cause du caractère cryptique de la punaise de lit, cachée dans des endroits difficiles d'accès. Pour les habitués, la punaise de lit a une odeur d'amande amère caractéristique, d'autant plus pour le flair d'un chien entraîné. Les chiens spécialisés dans la détection de punaises de lit peuvent signaler la présence d'un seul individu ou d'un lot d'œuf (n=5) dans 97,5% des cas, et les distinguer de punaises de lit mortes, mues et fèces dans 95% du temps (Pfiester *et al.*, 2008). En France, l'ACDPL⁶ - Association des Chiens Détecteurs de Punaises de Lit - a pour but d'encadrer et crédibiliser la pratique, bien que leur certification n'offre pas la reconnaissance d'un diplôme gouvernemental.

La mise au point de pièges a été également explorée dans de nombreuses études,

6. Site internet de l'Association des Chiens Détecteurs de Punaises de Lit : <https://acdpl.fr>

afin notamment de faciliter la surveillance et la détection précoce de punaises de lit dans les hôtels. Les punaises de lit sont attirées par la présence de leurs semblables, qui déposent un cocktail de cinq molécules qualifiées de phéromone d'attraction (Gries *et al.*, 2015). L'humain est également un attractant, *via* le CO₂ relâché et certaines kairomones (Singh *et al.*, 2013). Néanmoins, le nombre élevé de produits disponibles sur le marché et la grande variabilité de leurs prix ne permet pas à l'un de ces pièges de se démarquer (Crawley and Borden, 2021). Sur le terrain, les professionnels n'utilisent que rarement ces pièges, ou privilégient les pièges de type surfaces collantes à placer autour ou sous les pieds du lit.

Lutte physique et mécanique

Une fois l'insecte formellement identifié, il est conseillé de nettoyer la zone avec une brosse à sec avant de l'aspirer méthodiquement, afin de retirer un maximum d'insectes et d'œufs. L'aspirateur devra être à son tour soigneusement nettoyé, voire traité avec un insecticide afin d'éviter que les punaises ne s'en échappent. Aucun recoin ne doit être négligé : certaines fortes infestations nécessitent de changer un papier peint partiellement décollé, ou de condamner les plinthes par des joints. Le linge potentiellement contaminé doit être congelé minimum 48h à -20°C, ou lavé en machine à 60°C. Il existe également des techniques plus onéreuses telles que le nettoyage vapeur à 120°C, qui demande beaucoup de minutie, le chauffage d'une pièce entière à plus de 80°C pendant plusieurs heures, ou du mobilier dans une grande étuve.

Lutte naturelle

Il n'existe que peu de solutions naturelles contre la punaise de lit commercialisées en France. La terre de diatomée, ou dioxyde de silicium, en est une (Figure 1.5.A). Cette poudre composée de fossiles d'algues diatomées microscopiques (10–200 µm) est extrêmement abrasive pour les insectes (Akhoundi *et al.*, 2019). Lors d'un contact, la terre de diatomée adhère à la cuticule et dégrade sa couche protectrice lipidique, pénétrant dans le corps de l'insecte. À chaque mouvement de l'insecte, la poudre de silice provoquera ensuite des lésions dans les organes et ainsi une perte en eau. La punaise de lit mourra de dessiccation après une dizaine de jours en moyenne, en fonction des concentrations. Les formulations en spray ou en gel sont préconisées pour éviter que cette poudre très volatile ne soit inhalée, étant dangereuse pour les poumons, tout particulièrement chez les enfants et les animaux domestiques.

Certaines huiles essentielles, telles que le géraniol, sont préconisées contre la punaise de lit. L'EcoRaider, un produit composé d'huiles essentielles de géraniol, de cèdre et d'un détergent a même montré des résultats similaires à des insecticides chimiques de la famille des néonicotinoïdes et des pyréthrinoïdes (Wang *et al.*, 2014). Les pyrèthres naturels, extraits de la plante *Tanacetum cinerariaefolium* sont également utilisés pour leur action neurotoxique (Figure 1.5.B). Leur synthèse en laboratoire a permis de stabiliser et ainsi d'augmenter la rémanence de ces molécules, donnant naissance à la famille des pyréthrinoïdes chimiques (Matsuo, 2019).



Figure 1.5 – Exemples de solutions naturelles de lutte contre les punaises de lit. A. Punaises de lit adultes mortes après exposition à la terre de diatomée (Doggett and Russell, 2009); B. Planche botanique du pyrèthre de Dalmatie, *Tanacetum cinerariaefolium* (Franz Eugen Köhler, *Köhler's Medizinal-Pflanzen*, 1897); C. Punaises de lit adultes mortes après exposition au champignon *Beauvaria bassiana* (photos prises par Émeline Lemarié, INRAE Marseille, 2021).

Parmi les nouvelles pistes explorées, le potentiel de plusieurs bactéries entomopathogènes a été testé, en les ajoutant au sang avant que les punaises ne s'en nourrissent. Si la mortalité ne dépassait jamais les 50%, *Serratia marcescens*, *Pseudomonas fluorescens* et *Bacillus thuringiensis israelensis* (Bti) ont tous les trois montré un potentiel pour le bio-contrôle de la punaise de lit (Pietri and Liang, 2018).

Concernant les champignons, plusieurs espèces ont également été à l'essai, telles que *Isaria fumosoroseus*, *Lecanicillium muscarium* et *Beauvaria bassiana* (Aak *et al.*, 2018). Lors de ces essais, seule l'espèce *Beauvaria bassiana* (Figure 1.5.C) a eu un effet significatif sur la mortalité de la punaise de lit, de manière concomitante avec d'autres études (Barbarin *et al.*, 2012). Appliquée sur des souches de punaises de lit résistantes aux pyréthrinoïdes et aux néonicotinoïdes, *B. bassiana* permet de restaurer le phénotype de sensibilité avec une mortalité approchant les 100%, contre moins 40% avec un produit à base de deltaméthrine (Barbarin *et al.*, 2017). Néanmoins, cette mortalité n'est atteinte qu'à 8-10 jours post-traitement. Un produit contenant *B. bassiana*, Aprehend™, est

commercialisé depuis fin 2017 aux USA et au Canada.

Néanmoins, la plupart des traitements sont réalisés avec des insecticides chimiques, considérés comme à la fois plus efficaces, mais aussi plus économiques. En effet, les traitements thermiques sont très onéreux, même si le fait de pouvoir les utiliser plutôt que de jeter un matelas ou un canapé neuf (à éviter absolument de traiter en chimique, du fait du contact direct avec la peau et de la rémanence des produits) peut permettre de rentabiliser ce coût (environ 35€ le m² pour un canon à chaleur, contre moins de 10€ le m² pour un traitement chimique⁷). Quant au delta concernant l'efficacité, il a été souligné plus haut qu'il fallait compter une dizaine de jours pour que la terre de diatomée ou les champignons détruisent une population de punaises (le temps de la propagation puis de la croissance du champignon par exemple). *A contrario*, un insecticide chimique peut engendrer la mort de l'insecte en à peine quelques heures d'exposition, lorsqu'il est correctement utilisé. Enfin, privilégier l'utilisation de techniques "naturelles" est un choix clivant, que toutes les entreprises ne proposent pas.

(Més)usages des insecticides conventionnels

Les insecticides, comme les antibiotiques, ne devraient pas être automatiques. Néanmoins, peut être du fait de l'absence de rime, leur usage a longtemps été trop peu précautionneux. L'innovation galopante permettait un remplacement rapide des insecticides dès que leur efficacité était amoindrie, sans trop se soucier de la cause sous-jacente. À la fin du XX^e siècle, un ralentissement de l'innovation devenue trop coûteuse est observé, ainsi que la mise en place de tests pour évaluer la toxicité sanitaire et environnementale des produits avant de permettre leur commercialisation. Ces deux facteurs ont limité le champs des possibles, et encouragé secteurs publics & privés à se pencher plus sérieusement sur ce phénomène de résistance (Ruaux, 2013).

Avant cela, on pouvait utiliser contre la punaise de nombreux produits également létaux pour l'humain. Dans les années 1800, l'arsenic et le mercure étaient ainsi utilisés en insecticides de contact (Kinsley, 1893). Des fumigations ont été réalisées avec des bougies au soufre (Hockenyos, 1940), ou du cyanure d'hydrogène, y compris sous sa forme la plus tristement célèbre : les cristaux de Zyklon B, utilisés durant l'Holocauste (Potter, 2011).

7. Tarifs constatés sur <https://allopunaise.com/pages/prix-traitement-punaise-de-lit-tarif>

À la même période, les propriétés insecticides du dichlorodiphényl-trichloroéthane (DDT), un organochloré, sont découvertes et rapidement testées pour être appliquées aux baraquements américains. Ce produit semble alors inoffensif pour l'humain, tout en étant redoutablement efficace contre les punaises de lit. Sa rémanence dans le temps est tout particulièrement impressionnante : des résidus de produit ont pu provoquer la mortalité d'insectes jusqu'à trois ans après leur application (Mallis, 1954). Pourtant, alors que la punaise de lit semblait disparaître, les premières populations résistantes au DDT sont identifiées dès 1947 à la base navale américaine de Pearl Harbor (Johnson and Hill, 1948). Il faudra attendre une décennie après le succès de *Silent Spring*, le livre de Rachel Carson alertant quant aux effets du DDT sur l'environnement (Carson, 1962), pour que le DDT soit interdit dans la plupart des pays occidentaux. Il reste néanmoins encore aujourd'hui utilisé ponctuellement pour lutter contre les maladies vectorielles (et donc contrôler les populations de moustiques) dans de nombreux pays d'Afrique.

L'histoire du malathion, un insecticide de la famille des organophosphorés qui s'est substitué aux organochlorés dès les années 70, est similaire. Malgré son interdiction en France en 2008, une dérogation a été accordée à la Guyane en 2014, également dans le cadre de la lutte anti-vectorielle (Anses, 2014). Les carbamates, tout comme les organophosphorés, inhibent l'activité cholinestérase au niveau du système nerveux, ce qui les rend également toxiques pour les mammifères. La plupart des insecticides de cette famille (aldicarbe, carbofuran...) ont donc été interdits en France à partir de 2007, malgré les réticences de la filière betteravière (Langrand *et al.*, 2018). Remplaçons-les alors encore, avec les néonicotinoïdes cette fois, ciblant les récepteurs nicotiniques de l'acétylcholine. Interdits depuis 2018 en Union Européenne, notamment pour leur action mortifère qui leur a valu d'être surnommés les pesticides "tueurs d'abeille", ils ont été temporairement ré-autorisés en France pour les cultures de betterave en 2020, puis en 2022 (Le Monde, 2022). S'il existe des exceptions au sein de chaque famille, il semblerait qu'aucune ne déroge au schéma suivant - introduction, utilisation extensive, identification de toxicité et/ou baisse d'efficacité, remplacement.

Aujourd'hui, les familles d'insecticides pouvant encore être au moins partiellement utilisées contre la punaise de lit sont les pyréthrinoïdes, les régulateurs de croissance, les néonicotinoïdes, les carbamates, les pyrroles et les silicates (Centre National d'Expertise sur les Vecteurs (Institution), 2015). Les pyréthrinoïdes sont actuellement majoritaires, sur le marché comme dans les

usages. Ces insecticides ne tuant pas les œufs, la méthode de traitement préconisée implique deux voir trois passages à 7 à 10 jours d'intervalle, pour atteindre les larves nouvellement nées. Néanmoins, même en respectant ce protocole, il arrive que les traitements soient inefficaces.

Les mésusages des insecticides sont principalement réalisés par le grand public, n'ayant pas de formation adéquate, et contribuent à la mise sous cloche progressive des produits (comme concernant la mise à jour d'octobre 2019 du code de l'environnement⁸). Dans ce contexte plus que jamais, la formation des professionnel·le·s du secteur joue un rôle essentiel. Les formations se doivent d'insister tout particulièrement sur le fait que différents produits peuvent avoir le même mode d'action, notamment s'ils font partie d'une même famille. Les mauvaises pratiques de traitement, comme les applications successives d'insecticides de la même famille alors qu'une inefficacité est constatée, peuvent en effet favoriser la sélection de populations résistantes.

8. Article R522-16-3, décret n°2019-1052. (...) Il est interdit de céder directement en libre-service à des utilisateurs non professionnels : les produits (...) pour lesquels des données établissent qu'ils sont fréquemment utilisés en méconnaissance des règles visant à préserver la santé humaine ou l'environnement, figurant dans leur autorisation de mise sur le marché ou dans la notice élaborée par leur fabricant. Voir : <https://www.legifrance.gouv.fr/codes/id/LEGIARTI000039226824/2019-10-17>

1.2 Problématique de la résistance aux pyréthrinoïdes

La résistance la plus couramment observée et étudiée chez la punaise de lit est celle aux pyréthrinoïdes, étant la famille d'insecticide la plus utilisée. Pour comprendre la résistance aux pyréthrinoïdes, il nous faut d'abord revenir sur l'histoire de cette famille d'insecticides, avant de définir clairement le phénomène de résistance. Nous verrons ensuite comment détecter la résistance, et quels mécanismes peuvent entrer en jeu.

1.2.1 Les insecticides pyréthrinoïdes

Les pyrèthres naturels sont extrêmement photosensibles, avec une demi-vie de 5h au soleil. L'objectif premier de synthétiser des pyrèthres en laboratoire était donc de stabiliser la molécule, et d'augmenter la toxicité pour les insectes tout en diminuant celle pour les mammifères (Lopez and Fernandez-Bolanos, 2011). En 1949, le premier pyréthrinoïde de synthèse, l'alléthrine, est commercialisé. Quelques années plus tard, la perméthrine sera commercialisée pour une utilisation en agriculture, étant suffisamment peu photosensible.

Plusieurs nomenclatures coexistent pour classer les pyréthrinoïdes, dont une portant sur les manifestations de l'intoxication chez les mammifères, et une autre sur la structure de la molécule (Soderlund, 2012). Chez les rongeurs, on distingue en effet une symptomatologie de type T (tremblements, ataxie, excitabilité, hypersensibilité), et une symptomatologie de type CS (choréo-athétose, salivation, tremblements, convulsions). Les molécules peuvent également être séparées entre les pyréthrinoïdes de type I et de type II, le type II ayant un groupement α -cyané (cyano-3-phénoxybenzyle) absent chez le type I (Figure 1.6).

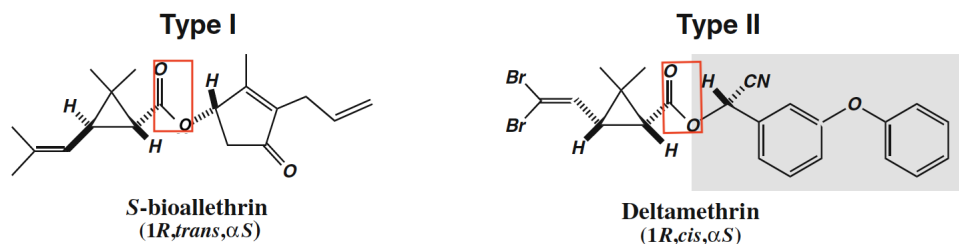


Figure 1.6 – Structure des pyréthrinoïdes de type I et II. Exemples de la *S*-bioallethrin (Type I) et de la deltaméthrine (Type II) (Soderlund, 2012). La fonction ester est encadrée en rouge. Le groupement α -cyano-3-phénoxybenzyle, caractéristique des molécules de type II, est encadré en gris.

Toujours selon Soderlund (2012), il est généralement admis que la symptomatologie de type T correspond au type I, et la symptomatologie CS au type II, même s'il y a des exceptions. Les pyréthriinoïdes de type I englobent l'alléthrine, la perméthrine, ou la tétraméthrine, tandis qu'on retrouve chez les pyréthriinoïdes de type II la cyperméthrine, la cyhalothrine, la deltaméthrine ou le fenvalérate. Les deux types perturbent l'ouverture des canaux sodium voltage-dépendants en augmentant le temps passé en position ouverte à raison de 20 ms de plus pour le type I et >200 ms de plus pour le type II, entraînant une perturbation dans la transmission de l'influx nerveux. Cet effet hyperexcitateur des pyréthriinoïdes de type II explique leur toxicité plus élevée que le type I.

Les pyréthriinoïdes ont également, dans une moindre mesure, un effet sur les canaux chlorure et calcium voltage-dépendants, ainsi que les récepteurs de l'acide γ -aminobutyrique (GABA), et les canaux activés par les GABA (Soderlund, 2012). Enfin, les pyréthriinoïdes peuvent être métabolisés par l'organisme à petite dose, en commençant par le clivage des esters (encadrés en rouge fig. 1.6), réalisé par des carboxylestérases ou par des cytochromes P450 (*CYP*).

Cette famille d'insecticide à la fois toxique et rémanente a donc été massivement appliquée sur le terrain, pour lutter contre une multitude d'insectes. Néanmoins, les premiers échecs de traitement des punaises de lit aux pyréthriinoïdes ont très vite été observés, et se sont rapidement propagés, en témoignent certaines études faites à large échelle comme aux États Unis (Romero *et al.*, 2007). La résistance est un phénomène observé d'abord sur le terrain, lorsqu'un produit appliqué sur une population d'insectes ne semble plus faire effet. Le produit peut alors être concentré, mais dans une certaine mesure seulement, chaque molécule pouvant atteindre une limite de solubilité, jusqu'à ce qu'appliquer le produit pur sur l'insecte n'ait plus aucun effet sur sa survie (Cáceres *et al.*, 2019). À partir d'un constat sur le terrain, la recherche vise à formaliser la détection de cette résistance, pour mieux surveiller son évolution dans le temps.

1.2.2 Détecter la résistance

Afin de déterminer si une population est résistante, la méthode la plus simple est de réaliser des bioessais, en comparant la courbe de mortalité réalisée avec des doses croissantes d'insecticides d'une population donnée à celle d'une population sensible (Barrès *et al.*, 2016). Il est ainsi possible de calculer un ratio de résistance (RR), correspondant au rapport entre deux doses létales pour un pourcentage donné

d'individus (par exemple 95%, soit la LD_{95} de la population résistante divisée par celle de la population sensible). Une population est considérée comme faiblement résistante avec un RR inférieur à 5, modérément résistante entre 5 et 10, et très résistante au delà de 10 (Mazzarri and Georghiou, 1995).

Concernant l'exposition aux insecticides, il existe deux techniques : le contact forcé, ou l'application topicale. Pour un contact forcé, du papier buvard est imprégné de la dilution voulue, avant de transvaser les insectes sur les résidus secs (Seong *et al.*, 2010). Dans le cas d'une application topicale, les insectes sont tout d'abord immobilisés, sur glace ou en les faisant inhaler du CO_2 (Figure 1.7.A), et une seringue en verre est utilisée afin de délivrer une dose donnée directement sur la cuticule de l'insecte (Figure 1.7.B). Dans les deux cas, les insectes doivent être conservés dans des conditions stables avant de comptabiliser la mortalité (Figure 1.7.C). Une fois que le statut de résistance d'une population est observé, il a fallu définir la résistance d'un point de vue évolutif et adaptatif, afin de formaliser le phénomène.

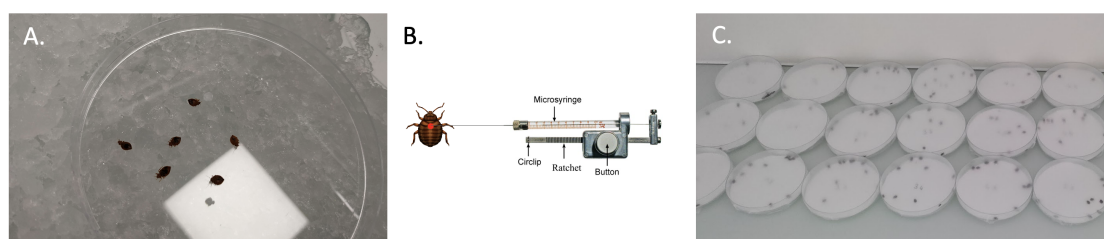


Figure 1.7 – Exemple de méthodologie d'exposition des punaises de lit aux insecticides. A. Engourdissement des insectes dans une boîte de Petri posée sur de la glace pilée. B. Application topicale de l'insecticide grâce à une seringue montée sur un distributeur à répétition. Le point rouge correspond à la zone qui doit être visée pour l'application, entre les coxae en face ventrale (adapté de Caffrey and Cherezov (2009), pas à l'échelle). C. Boîtes de Petri contenant les punaises sur le papier buvard propre, scellées au parafilm en attendant de mesurer le taux de mortalité.

1.2.3 Définir la résistance

La résistance aux insecticides est une diminution de la sensibilité héréditaire observée chez les populations exposées par rapport aux non-exposées au cours du temps.

Une mutation conférant un avantage face à l'application d'un insecticide augmentera les chances de survie et de reproduction de l'individu la portant. Ainsi, la fréquence de cette mutation adaptative augmentera dans la population sous la pression de sélection produite par les applications répétées de cet

insecticide.

La résistance est à distinguer de la tolérance aux insecticides, qui décrit la capacité naturelle d'un insecte à supporter l'effet de l'insecticide, sans être le résultat de changements génétiques récents dus à l'utilisation massive d'insecticide (Zhu, 2008). Chez la punaise de lit, les changements physiologiques suite à un repas de sang peuvent par exemple conférer une tolérance accrue aux pyréthrinoïdes (DeVries *et al.*, 2015).

1.2.4 Mécanismes de résistance

Il existe quatre grands mécanismes de résistance : la résistance comportementale, cuticulaire, métabolique, et par modification de la cible (Figure 1.8). Dans un souci de concision, nous détaillerons ces mécanismes uniquement pour la résistance aux pyréthrinoïdes.

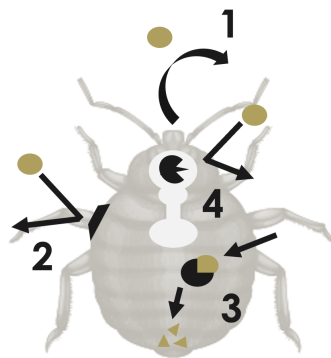


Figure 1.8 – Mécanismes hypothétiques de résistance aux insecticides chez la punaise de lit. 1 - Résistance comportementale : évitement ; 2 - Résistance cuticulaire : épaissement ou modification de la cuticule, diminuant la pénétration de l'insecticide ; 3 - Résistance métabolique : dégradation de l'insecticide par des enzymes, comme les cytochromes P450 ou les estérases ; 4 - Résistance par modification du site cible : mutation au niveau du système nerveux empêchant l'insecticide de s'y fixer.

Comportementale

La résistance comportementale est principalement décrite comme étant qualitative, avec un évitement temporel, spatial ou trophique (relatif à la nutrition) du contact avec les insecticides (Carrasco *et al.*, 2019). De manière plus conceptuelle, Carrasco *et al.* (2019) mentionne également une résistance quantitative, caractérisée par des comportements de fuite, de thermorégulation, ou encore "d'auto-médication" à la suite d'un contact avec l'insecticide.

Un décalage temporel d'activité a pu par exemple être observé dans des populations de moustiques du genre *Anopheles* exposées aux insecticides. L'utilisation de moustiquaires imprégnées, ou la pulvérisation des insecticides en intérieur est en effet mortelle pour les moustiques se déplaçant à la recherche de leur proie lorsqu'elle se trouve dans son logement, pendant la nuit (Moiroux *et al.*, 2012). *A contrario*, les moustiques prédatant l'humain au lever du jour ont plus de chance de survivre et de se reproduire, entraînant donc une sélection de ce comportement permettant la résistance aux pyréthrinoïdes (Moiroux *et al.*, 2012). Un évitement strictement spatial, avec une augmentation des piqûres à l'extérieur du logement, a été observé chez diverses espèces d'*Anopheles* (Musiba *et al.*, 2022). Un évitement trophique, par la prédation d'autres animaux non-humains, a également pu être mis en évidence dans des population de *Culex quinquefasciatus* résistantes à la perméthrine, et également dans de nombreuses espèces d'*Anopheles* (Bøgh *et al.*, 1998; Ndenga *et al.*, 2016).

À ce jour, aucune résistance comportementale n'a été mise en évidence chez *C. lectularius* (voir review Dang *et al.* 2017), bien que certains comportements pouvant limiter l'effet des pyréthrinoïdes aient pu être observés. La deltaméthrine semble en effet avoir un effet excitant-répulsif sur les punaises de lit, avec une augmentation des déplacements quand l'insecte est en contact avec cet insecticide, ou la préférence d'une tente non-traitée aux insecticides à une tente traitée pour se cacher (Romero *et al.*, 2009a). De part leur mode de vie cryptique et nocturne, l'évitement temporel ou spatial décrit chez les moustiques est difficilement envisageable chez la punaise de lit. En revanche, un évitement trophique, en se nourrissant sur un animal de compagnie dont le panier serait dans une pièce non-traitée, ou un nid d'oiseau à proximité semble théoriquement possible.

Cuticulaire

La résistance cuticulaire peut être caractérisée par un épaississement ou une modification de la structure cuticulaire, diminuant le taux de pénétration des insecticides (Balabanidou *et al.*, 2018).

Pour quantifier la pénétration de l'insecticide, un dosage du taux d'insecticide ayant pénétré relativement à celui resté à l'extérieur de la cuticule peut être réalisé. Une diminution de la pénétration a ainsi été observée chez une population d'*Helicoverpa armigera* résistante à la deltaméthrine (Ahmad *et al.*, 2006). Les raisons de cette diminution de la pénétration peuvent être diverses : chez *A. gambiae*, l'épaississement de la cuticule d'une population résistante à plusieurs

insecticides a été lié à une surexpression en cytochrome P450s, engendrant un enrichissement en hydrocarbures de l'épicuticule (Balabanidou *et al.*, 2016). Au sein d'une autre population résistante à la deltaméthrine d'*A. gambiae*, cet épaissement a été corrélé à la surexpression de protéines cuticulaires (Yahouédo *et al.*, 2017). Une surexpression de laccase 2 peut également modifier la structure moléculaire de la cuticule, la rendant moins pénétrable, comme chez une souche de *C. pipiens* résistante au fenvalérate (Pan *et al.*, 2009). Enfin, les transporteurs de type ABC peuvent acheminer des lipides cuticulaires (hydrates de carbone) dans l'épiderme, réduisant la pénétration des insecticides (lipophiles) (Balabanidou *et al.*, 2018), bien qu'il reste à préciser les mécanismes exacts conduisant à la résistance.

Une étude morphologique a confirmé que la cuticule était plus épaisse chez une souche de punaises de lit résistante aux pyréthrinoïdes collectée en Australie, par rapport à une souche sensible (Lilly *et al.*, 2016a). Un épaissement accru de la cuticule est corrélé avec un temps plus long précédant la paralysie (knock-down) de l'insecte suite à une exposition aux insecticides. D'autres études basées sur des analyses RT-qPCR ont montré que les gènes des protéines cuticulaires, tels que la chitine synthase (CHS) ou la protéine cuticulaire larvaire/pupale (L/PCP), présentaient des niveaux de transcription plus élevés chez les souches de punaises de lit résistantes aux pyréthrinoïdes (Mamidala *et al.*, 2012; Koganemaru *et al.*, 2013).

Métabolique

Après la pénétration de la barrière cuticulaire, de nombreuses voies métaboliques peuvent être mobilisées pour dégrader ou excréter un insecticide. Ces mécanismes de détoxification auraient été sélectionnés chez les insectes phytophages pour leur permettre de résister aux toxines produites par les plantes, durant des millions d'années de course aux armements (Després *et al.*, 2007). Bien que la punaise de lit ne soit plus en contact avec le règne végétal aujourd'hui, du fait de son hématophagie stricte et de sa vie en intérieur, on pense que cette hématophagie a évolué à partir d'insectes phytophages (Weirauch *et al.*, 2019). La mise en place ancestrale de ces mécanismes de détoxification aurait donc également permis la sélection de leurs versions les plus efficaces de manière rapide au sein des souches de punaises de lit résistantes.

Les mécanismes potentiellement impliqués dans la résistance métabolique ont été décrits par Barrès *et al.* (2016) (Figure 1.9) comme étant :

- la surexpression d'une enzyme ou l'amélioration de son efficacité (grâce à une meilleure affinité pour le substrat par exemple), augmentant ainsi la dégradation de l'insecticide ;
- l'efflux hors de la cellule *via* des transporteurs ABC ;
- la séquestration de l'insecticide (moléculaire ou intracellulaire) ;
- et dans le cas où l'enzyme atteindrait sa cible, la protection contre les effets cytotoxiques potentiellement générés, ou la compensation par une voie ou enzyme alternative de la cible.

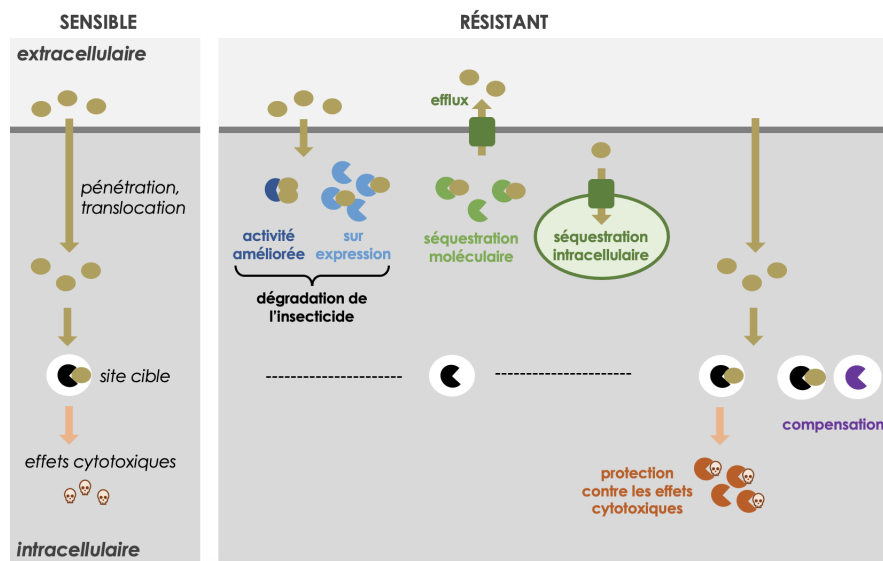


Figure 1.9 – Mécanismes métaboliques entraînant la résistance aux insecticides. Les enzymes dégradant l'insecticide peuvent être surexprimées, ou avoir plus d'affinité pour l'insecticide. L'excrétion par les transporteurs en dehors de la cellule (efflux) peut être accrue, ainsi que la séquestration de l'insecticide (moléculaire ou intracellulaire). Dans le cas où l'enzyme atteindrait sa cible, certaines enzymes peuvent conférer une protection contre les effets cytotoxiques potentiellement générés, ou remplacer la cible dans ses fonctions, compensant ainsi son occupation par l'insecticide.

La dégradation de l'insecticide peut être réalisée *via* les enzymes de phase I, réalisant principalement des réactions d'oxydation grâce aux monooxygénases à cytochromes P450 (P450s) ou d'hydrolyse par les estérases. Leur action peut être suivie par celle des enzymes de phase II, pour les étapes par exemple de glutathionylation avec les glutathion-S-transférases (GSTs) ou de glucuronidation *via* les UDP-glucuronosyltransferases (UGTs) (Xu *et al.*, 2005). Néanmoins, une seule enzyme de la phase I ou II peut suffire dans certains cas à dégrader l'insecticide (Amezian *et al.*, 2021). La surexpression d'un P450, le *CYP6A51*, semble ainsi impliquée dans la résistance aux pyréthriinoïdes chez la mouche méditerranéenne des fruits, *Ceratitidis capitata*, (Arouri *et al.*, 2015).

La phase III correspond à l'excrétion enzymatique de l'insecticide hors de la

cellule ou dans la vacuole. Cette étape d'efflux est principalement réalisée par des transporteurs ABC (Gott *et al.*, 2017). Chez *Aedes aegypti* par exemple, certains transporteurs ABC sont surexprimés dans les souches résistantes aux pyréthri-noïdes (Bariami *et al.*, 2012). Plusieurs enzymes peuvent être impliquées également dans la séquestration de l'insecticide, dont les esté-rases, comme cela a été observé avec le pyréthri-noïde esfenvalérate chez la noctuelle *Helicoverpa armigera* (Gunning *et al.*, 1996).

Enfin, chez la cicadelle brune *Nilaparvata lugens*, une activité élevée des GSTs confère une résistance à la perméthrine et à la λ -cyhalothrine, en limitant la peroxydation lipidique que ces insecticides induisent, et qui cause des dommages oxydatifs aux tissus (Vontas *et al.*, 2001). D'autres enzymes connues pour leur rôle protecteur du stress oxydatif peuvent également être impliquées dans la résistance aux insecticides (Zhou *et al.* 2019, voir 1.2.4).

La plupart de ces mécanismes de résistance métabolique semblent donc dus à une surexpression enzymatique. Quand on s'intéresse de plus près aux séquences sous-jacentes, une surexpression enzymatique peut être due à (i) une mutation de la séquence codante (Campbell *et al.*, 1998) ou de son promoteur (Pang *et al.*, 2014), (ii) une amplification génique (Bass and Field, 2011), ou (iii) un changement au niveau des éléments régulateurs (Wilding 2018, détaillé en section 1.2.4).

Dans les populations résistantes aux pyréthri-noïdes de punaises de lit, une activité accrue de plusieurs enzymes telles que les cytochromes P450 (*CYP*) (Romero *et al.*, 2009b) et les esté-rases (Lilly *et al.*, 2016b) a été observée. Des expressions plus élevées des glutathion-S-transfé-rases (*GST*) ont également été détectées chez les stades immatures de punaises de lit résistantes, mais ce mécanisme reste à identifier chez les adultes (Mamidala *et al.*, 2011). Des souches résistantes ont été également exposées à des pyréthri-noïdes en synergie avec des inhibiteurs enzymatiques spécifiques, afin de déterminer si cela permet de restaurer la sensibilité aux insecticides, et ainsi évaluer l'implication de ces enzymes dans le phénotype résistant. Le piperonyl butoxyde (PBO), un inhibiteur des cytochromes P450, a diminué de manière significative la résistance à la deltaméthrine (Romero *et al.*, 2009b). De même, un inhibiteur des esté-rases a augmenté la mortalité des punaises de lit par rapport à l'utilisation de la λ -cyhalothrine seule (Lilly *et al.*, 2016b). Enfin, les gènes codant pour les transporteurs ABC *Abc8*, *Abc9*, *Abc10* et *Abc11* ont été détectés comme étant surexprimés chez des populations de *C. lectularius* résistantes aux pyréthri-noïdes (Zhu *et al.*, 2013). L'inhibition par RNAi des *Abc8* et *Abc9* au cours de la même

étude a diminué significativement la résistance de ces souches à la β -cyhalothrine.

Mutation de cible

Si certaines enzymes de détoxification peuvent être polyvalentes et dégrader plusieurs familles d'insecticides, les mutations de la cible de l'insecticide sont en revanche, par définition, plus spécifiques. Deux principaux mécanismes liés à des mutations de cible ont été décrits chez les populations résistantes (Figure 1.10) :

- la surexpression de la cible, permettant de maintenir la fonction neurologique en compensant la présence de l'insecticide qui l'entrave ;
- et la modification de la cible, diminuant son affinité pour l'insecticide et prévenant la liaison de ce dernier.

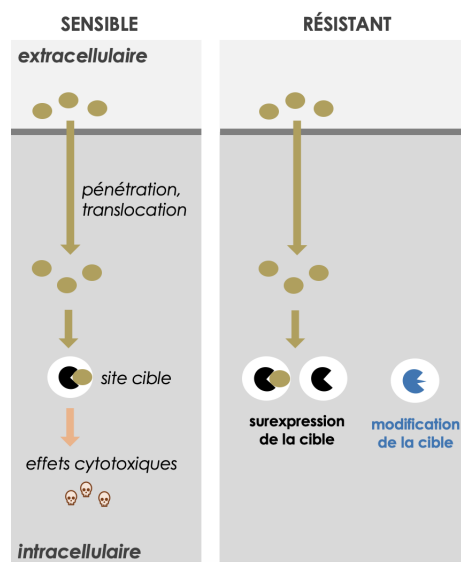


Figure 1.10 – Mutations de cible entraînant la résistance aux insecticides. La cible de l'insecticide peut être surexprimée, ou mutée de sorte que l'insecticide ne puisse plus s'y fixer.

Les pyréthrinoïdes ciblent les canaux sodium voltage-dépendants du système nerveux de l'insecte, provoquant sa paralysie (effet knock-down). Une mutation non synonyme d'un seul nucléotide peut modifier la conformation de ces canaux sodiques et, par conséquent, empêcher la liaison des pyréthrinoïdes. Cette mutation conférant une résistance à l'effet "knock-down", elle est appelée mutation *kdr*. La première mutation *kdr* (L1014F, pour leucine à phénylalanine en position 1014) a été identifiée chez des populations de mouche domestique, *Musca domestica*, modérément résistantes au DDT et aux pyréthrinoïdes (Williamson *et al.*, 1993). Une mutation supplémentaire appelée "super-*kdr*" (M918T) a été également détectée chez des populations à la résistance plus

élevée, portant également toujours *kdr*. À quelques variations près, la mutation *kdr* a été observée chez de nombreuses espèces d'insectes, dont la blatte germanique *Blattella germanica* (Dong and Scott, 1994), la mouche *Haematobia irritans* (Jamroz *et al.*, 1998), ou encore le moustique *A. aegypti* (Saavedra-Rodriguez *et al.*, 2007).

La mutation *kdr* a été étudiée chez de nombreuses populations de punaises (Dang *et al.*, 2014, 2015; Akhondi *et al.*, 2015; Balvín and Booth, 2018). Trois mutations *kdr* ont été associées au phénotype de résistance aux pyréthrinoides chez *C. lectularius* : V419L (de la valine 419 à la leucine) et L925I (de la leucine 925 à l'isoleucine), observées au début des années 2000 dans de nombreuses villes d'Eurasie et des États-Unis (Zhu *et al.*, 2010, 2013), et I936F (de l'isoleucine 936 à la phénylalanine), pendant la même période en Australie (Dang *et al.*, 2015). Selon (Durand *et al.*, 2012), toutes les punaises de lit échantillonnées à Paris (n =124) en 2011 étaient homozygotes pour la mutation L925I.

Autres mécanismes

D'autres mécanismes, moins bien décrits, peuvent être également impliqués dans la résistance aux insecticides. Trois en particulier sont détaillés ici : les facteurs de transcription (Wilding, 2018), les gènes impliqués dans l'équilibre red-ox (Champion and Xu, 2018), et les symbioses bactériennes (Kikuchi *et al.*, 2012).

En amont des gènes codant pour les enzymes de résistance métabolique, les facteurs de transcription peuvent jouer un rôle majeur (Wilding, 2018). Les insecticides peuvent être reconnus par de nombreux récepteurs, enclenchant une cascade de réactions moléculaires qui permettra à un facteur de transcription de se lier en amont de son gène cible afin d'activer sa transcription. Cinq voies de signalisation impliquées dans l'expression de gènes de détoxification ont ainsi été décrites (Amezian *et al.*, 2021) : GPCR (G protein-coupled receptor), MAPKs-CREB (Mitogen-activated protein kinases-cAMP response element-binding protein), AhR/ARNT (Aryl hydrocarbon receptor/Aryl hydrocarbon receptor nuclear translocator), HR96 (Hormone receptor-like in 96) et CncC/MAF (Cap'n'collar isoforme/Musculoaponeurotic fibrosarcoma) ou Keap1 (Kelch-like ECH-associated protein 1) (Figure 1.11). Les récepteurs couplés à la protéine G (GPCR) sont par exemple surexprimés chez les insectes résistants de nombreuses espèces, et seraient impliqués dans la régulation des cytochromes P450 (Liu *et al.*, 2021). C'est le cas aussi de l'activation par MAPK

du facteur de transcription CREB, qui régule l'expression des P450 responsables de la résistance aux néonicotinoïdes chez *Bemisia tabaci* (Yang *et al.*, 2020), ou de la voie CncC/Maf, responsable de la surexpression des P450 chez *Spodoptera exigua* (Hu *et al.*, 2021). Ces études étant toutes récentes, l'impact des facteurs de transcription reste sûrement encore largement sous-estimé.

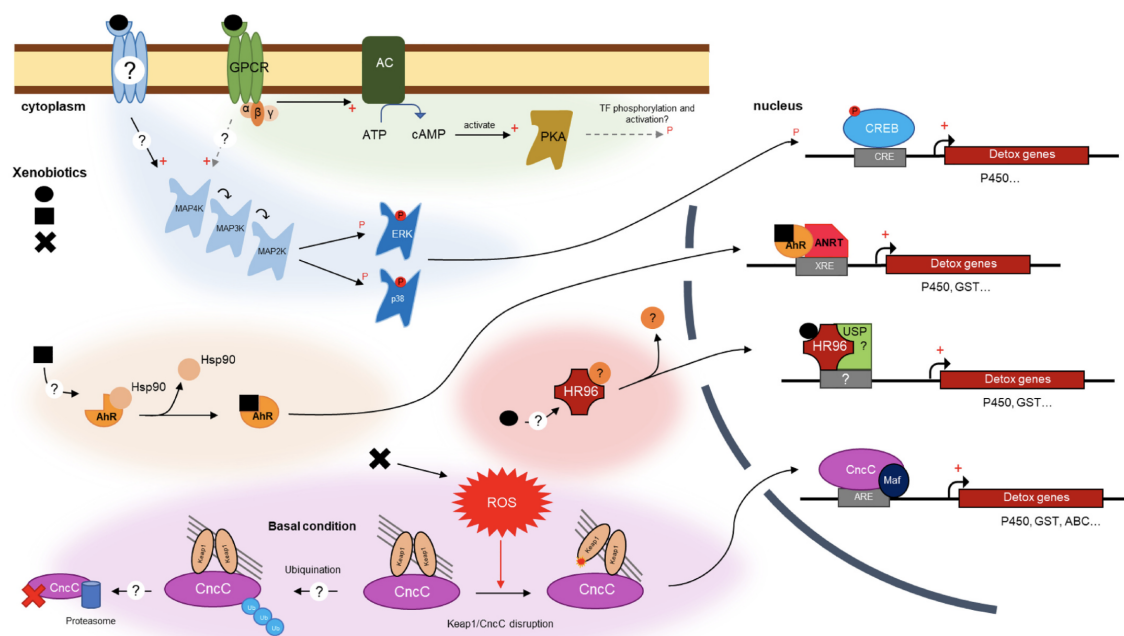


Figure 1.11 – Voies de régulation transcriptionnelle de l'expression des gènes de détoxification chez les insectes. Cinq voies de signalisation conduisant à la régulation de la transcription de gènes de détoxification potentiellement impliqués dans la résistance aux insecticides ont été décrites : GPCR (vert), MAPKs-CREB (bleu), AhR/ARNT (orange), HR96 (rouge) et CncC/Keap1 (violet) (Amezian *et al.*, 2021).

Le système redox (réduction-oxydation) est essentiel au maintien de l'homéostasie des organismes. Chez plusieurs espèces de moustique du genre *Anopheles*, l'inhibition d'enzymes de défense oxydative telles que la superoxyde dismutase (SOD) Cu-Zn et les catalases augmente la sensibilité aux insecticides (Champion and Xu, 2018). Les punaises de lit étant hématophages, elles sont soumises à un stress oxydatif intense après un repas de sang, et donc adaptées pour éliminer les radicaux libres à l'origine de ce stress. Les pyréthriinoïdes générant un fort stress oxydatif (Wang *et al.*, 2016), on peut émettre l'hypothèse que ces adaptations peuvent être également avantageuses pour leur défense contre l'exposition à cet insecticide.

Certains symbiotes peuvent également influencer le phénotype de résistance. L'hémiptère *Riptortus pedestris* possède une bactérie symbiotique du genre *Burkholderia* qui peut lui conférer une résistance au fénitrothion, un organophosphoré, en le dégradant (Sato *et al.*, 2021). Une étude très récente chez

C. hemipterus a montré une augmentation de la sensibilité aux insecticides fénitrothion (un organophosphoré) et imidaclopride (néonicotinoïde) mais pas à la deltaméthrine, suite à un traitement antibiotique (Soh and Veera Singham, 2022). L'hôte a été supplémenté en vitamine B en parallèle à ce traitement, permettant ainsi de maintenir fécondité, taux d'éclosion, ainsi que développement larvaire de manière similaire aux insectes sans traitement. Les symbiotes bactériens pourraient donc jouer un rôle dans la résistance à certains insecticides de l'hôte chez la punaise de lit tropicale. À notre connaissance, ces mécanismes n'ont pas encore été explorés chez la punaise de lit commune, *C. lectularius*.

1.2.5 Résistances : un pour tous et tous pour un

Certains gènes impliqués dans la résistance peuvent être polyvalents, en interagissant avec différentes familles d'insecticides - on parle alors de résistance croisée (un pour tous, ou tout du moins pour plusieurs). Lorsqu'il s'agit d'un ensemble de gènes qui est impliqué dans la résistance, on parle de résistance polygénique (tous pour un). Pour ne rien faciliter, l'identification d'une résistance polygénique n'exclue pas l'implication de résistance croisée.

Résistance croisée

La résistance croisée peut entraver la mise en place de bonnes pratiques de lutte. En effet, certains mécanismes peuvent conférer une résistance à différents insecticides, appartenant à une ou plusieurs familles, bien que n'ayant été sélectionnés suite à une exposition à un seul de ces insecticides. La résistance croisée peut alors rendre inefficace une rotation entre différentes molécules.

Une résistance croisée a tout d'abord été détectée chez des souches de *Drosophila melanogaster* résistantes au DDT, ainsi qu'à des insecticides d'autres familles (Weiner and Crow, 1951). Cette résistance a été imputée à un gène codant pour le cytochrome P450 Cyp6g1, sur-exprimés chez les souches résistantes, et capable de métaboliser de nombreux xénobiotiques, comme le néonicotinoïde imidaclopride (Joußen *et al.*, 2008; Le Goff and Hilliou, 2017).

Le DDT a également été largement plébiscité pour lutter contre la punaise de lit durant la seconde guerre mondiale. Cet insecticide ciblant les canaux sodium voltage-dépendants, tout comme les pyréthriinoïdes, une résistance croisée à ces deux familles a été détectée chez *C. lectularius* dès la fin des années 50 (Busvine,

1958). La mutation *kdr* des canaux sodium voltage-dépendants, en empêchant l'insecticide de s'y fixer, peut être responsable de cette résistance croisée. Des modélisations de liaisons moléculaires ont également montrées que les P450s pouvaient octroyer une résistance croisée à plusieurs insecticides chez *C. lectularius*, tels que les organochlorés, les pyréthrinoïdes, et les néonicotinoïdes (Mamidalala *et al.*, 2011).

Résistance polygénique

La résistance multiple est une accumulation de mutations conférant différents avantages. On peut observer des cas de synergie, avec par exemple une mutation du site de fixation de l'insecticide qui rendrait le gène moins efficace, et une mutation compensatoire pour fixer tout de même le substrat originel. Ce type de réponse, impliquant plusieurs gènes, est appelée polygénique.

Une étude menée sur le doryphore de la pomme de terre, *Leptinotarsa decemlineata*, a ainsi mis en évidence 42 gènes candidats de résistance aux insecticides chez différentes populations de cette espèce (Crossley *et al.*, 2017).

Globalement, les populations d'insectes de laboratoire, dont les élevages sont créés à partir de peu d'individus, montrent plutôt une réponse polygénique à la sélection à un insecticide, tandis qu'elle semble monogénique *in natura* (Georghiou, 1972; Roush and McKenzie, 1987; Ffrench-Constant *et al.*, 1990). L'hypothèse principale pour expliquer ce patron s'appuie sur les grandes tailles des populations de terrain, qui permettraient à la sélection d'agir sur une infinité de mutations rares. *A contrario*, les populations de laboratoire étant très consanguines, l'application d'insecticide ne peut que sélectionner un ensemble de traits restreints, et donc de manière probable plusieurs mutations dont les effets cumulés conféreront une résistance (Ffrench-Constant *et al.*, 1990, 2004).

De manière intéressante, le mode de dispersion de la punaise de lit - introduction avec une très petit propagule, voir une seule femelle - doit rendre cet insecte très résilient face aux effets fondateurs. Un effet fondateur se produit lorsqu'un petit nombre d'individus quitte une population pour aller en fonder une nouvelle (Mayr, 1942). Cet effet se traduit par une perte de diversité génétique au sein de la nouvelle population, étant donné que le nombre de reproducteurs capables d'échanges géniques diminue drastiquement par rapport à la population originelle. On peut donc imaginer des résistances qui évolueraient de manière polygénique au sein de grandes populations de punaise de lit, avant d'être réduite

à une faible diversité allélique lors de la colonisation d'un nouvel environnement, de manière passive ou active. L'absence de traitement dans cet environnement peut également jouer sur le nombre d'allèles impliqués dans la résistance, étant donné que ceux-ci ont des coûts plus ou moins élevés, et peuvent donc être contre-sélectionnés s'ils ne sont plus utiles dans l'immédiat.

1.2.6 Coût de la résistance

En biologie, on parle de compromis (trade-off) quand le gain apporté par un avantage tel que la résistance aux insecticides va de pair avec la diminution partielle ou totale d'une autre compétence (fécondité, fertilité, capacité de déplacement,...) (Kliot and Ghanim, 2012). Des mécanismes de résistance variés sont employés pour réduire la toxicité engendrée par un xénobiotique, tels que la surexpression d'enzymes de détoxification ou la mutation du site ciblé par l'insecticide, présentés ci-dessus. Ces mécanismes peuvent en effet être coûteux, et affecter négativement diverses compétences de l'insecte (Trdan, 2016). Les individus sensibles aux insecticides peuvent alors avoir un avantage reproductif, ce qui peut faire baisser le niveau de résistance de la population, notamment si la pression de sélection liée aux insecticides diminue également. En revanche, si cette pression reste constante, des mutations de compensation peuvent apparaître, équilibrant les coûts liés à la résistance (Labbé *et al.*, 2014).

Impact sur la valeur sélective

Pour évaluer l'impact des mutations de résistance sur la valeur sélective des individus, plusieurs traits peuvent être comparés entre des lignées sensibles et résistantes aux insecticides élevées en laboratoire. Une diminution de la longévité et de la fécondité (nombre moyen d'œufs produits par femelle) a ainsi été observée chez des lignées de moustiques *Culex pipiens pallens* et *Aedes aegypti* résistantes aux pyréthrinoïdes (Jaramillo-O. *et al.*, 2014; Li *et al.*, 2002). Le temps de développement larvaire semble, *a contrario*, être allongé chez les lignées résistantes aux pyréthrinoïdes de *C. quinquefasciatus* et *A. aegypti*, notamment en présence de la mutation *kdr* (Li *et al.*, 2002; Martins *et al.*, 2012). Un temps de développement plus long implique un risque plus élevé d'être en contact avec des pathogènes ou des prédateurs, ou de voir les conditions abiotiques du milieu se dégrader.

Chez la punaise de lit, une étude a été réalisée sur les effets d'une seule

génération de sélection avec un produit contenant de la β -cyfluthrine (insecticide pyréthriinoïde) et de l'imidaclopride (néonicotinoïde) (Gordon *et al.*, 2015). Si la longévité des lignées de *C. lectularius* résistantes ne semblait pas affectée, le temps de génération, le taux de reproduction (nombre de descendants femelles produits par individu par génération) et la durée de ponte ont été significativement diminués. La résistance au sein de ces souches est, au moins en partie, due à la production métabolique d'enzymes de détoxification, les cytochromes P450s.

Différents coûts pour différentes mutations

En fonction des mécanismes de résistance impliqués, les coûts associés ne sont pas les mêmes. La surexpression d'enzymes de détoxification, et plus largement la résistance métabolique, engendre donc un coût en consommant des ressources et de l'énergie au détriment des processus métaboliques et de développement. La résistance aux pyréthriinoïdes chez le moustique *A. funestus* est conférée par la surexpression des P450s CYP6P9a (Weedall *et al.*, 2019). Tout comme chez *C. lectularius*, cette résistance semble avoir un coût métabolique : les moustiques portant l'allèle de résistance ont moins de chance de pondre des œufs que les moustiques sensibles (Tchouakui *et al.*, 2020). Toujours d'après cette étude, le taux d'éclosion, ainsi que le nombre de larves étaient également inférieurs chez les moustiques homozygotes résistants. La production de ces enzymes utiliserait des ressources qui ne pourraient alors plus être allouées à certains traits, comme la fécondité de l'insecte.

Les mutations de cible permettant la résistance peuvent diminuer l'affinité de la cible pour son substrat, altérant la fonction d'origine du gène. Les gènes pouvant avoir des rôles pléiotropiques, déterminant plusieurs traits phénotypiques, une mutation peut en effet affecter d'autres traits, impactant potentiellement la survie ou la reproduction de l'individu. C'est le cas chez le puceron vert du pêcher *Myzus persicae*, pour lequel la mutation du site cible *kdr* est associée à une diminution de la réponse à une phéromone d'alarme (Foster *et al.*, 2003). Autrement dit, la présence de mutations de résistance semble perturber la perception de certains stimuli environnementaux. La mutation *kdr* affectant les fonctions neuronales de l'insecte, il peut sembler cohérent qu'elle altère certains comportements de l'insecte.

Mutations de compensation

Lorsque la fonction d'origine d'un gène est empêchée par une mutation de résistance, des mutations de compensation peuvent être sélectionnées. Chez *C. pipiens* et *Anopheles gambiae*, une seule mutation nucléotidique (G119S) du gène codant pour l'acétylcholinestérase *ace-1* peut conférer une résistance aux organophosphorés et aux carbamates (Weill *et al.*, 2002). Néanmoins, cette mutation a un coût élevé sur la valeur sélective de l'individu en l'absence d'insecticide (Lenormand *et al.*, 1998), probablement dû au fait que l'enzyme est moins efficace pour hydrolyser l'acétylcholine, un neurotransmetteur essentiel au fonctionnement du système nerveux. Les duplications du locus *ace-1* observées, combinant une copie résistante *ace-1R* et une copie sensible *ace-1S*, pourraient alors permettre de compenser le coût d'*ace-1R*. Le coût de cette résistance (mesurée avec la mortalité pré-imaginale) semble en effet corrélé chez *C. pipiens* à la proportion relative de copies *ace-1S* et *ace-1R* (Labbé *et al.*, 2014).

Une sorte de compensation peut également avoir lieu au sein d'un même locus au cours du temps : on peut alors parler d'amélioration "directe" (Labbé *et al.*, 2009). Ainsi, chez le moustique *C. pipiens*, le locus *Ester* code pour des enzymes de détoxification de type estérases, dont la surproduction joue un rôle majeur dans la résistance aux organophosphorés chez cette espèce (Raymond *et al.*, 2001). Cet allèle de résistance a été remplacé à deux reprises par un allèle au coût plus faible, tout en conférant une résistance *a minima* équivalente (Labbé *et al.*, 2009). Le remplacement d'allèle par une version compensant un coût trop élevé permet ici d'améliorer l'adaptation de l'organisme à la présence d'insecticides.

Toutes ces mutations, impliquées dans la résistance aux insecticides et sélectionnées par la forte pression que ceux-ci exercent sur le vivant, relèvent de mécanismes adaptatifs. Différents mécanismes peuvent être à l'œuvre, en témoignent la variété des bases génomiques de l'adaptation. Ainsi, cette thèse s'inscrit dans un contexte plus global de compréhension de l'adaptation des organismes aux environnements dans lesquels ils évoluent.

1.3 Génomique de l'adaptation

L'adaptation peut être définie comme l'ajustement d'une population au milieu dans lequel elle se trouve. On peut reprendre l'exemple de fortes pressions de sélection liées aux applications d'insecticides, qui perturbent le milieu de vie de l'insecte en y ajoutant un danger mortel. Les individus porteurs de mutations conférant un avantage face à ce danger ont alors plus de chance de survivre, et donc de se reproduire et de répandre cette allèle de résistance. La fréquence des individus adaptés, et donc des allèles adaptatifs qu'ils portent, va alors augmenter au sein d'une population.

L'étude des variations de fréquence des différents allèles peut donc donner la possibilité de détecter l'adaptation. La génomique permet quant à elle d'avoir une vision d'ensemble du matériel génétique, en étudiant la diversité intra ou inter-populationnelle. Ensemble, ces méthodes permettent donc d'isoler les variants alléliques impliqués dans un mécanisme adaptatif de la majorité des mutations étant neutres ou délétères.

Afin de détecter de potentiels variants alléliques adaptatifs, les organismes étudiés doivent tout d'abord être séquencés. Depuis les années 2000, les approches de séquençage nouvelle génération (NGS, pour next generation sequencing) ont révolutionné les études génomiques, en permettant d'obtenir un grand nombre de séquences en très peu de temps et à très faible coût ([Behjati and Tarpey, 2013](#)). On peut ainsi réaliser des études d'association pangénomiques (GWAS, genome-wide association study), qui consistent à génotyper séparément les individus des phénotypes résistants et sensibles, puis à rechercher les variants génétiques associés aux différents phénotypes ([Uffelmann *et al.*, 2021](#)). Bien que le séquençage à longues lectures (long-reads) ouvre de nouvelles possibilités de détection des variants impliqués dans l'adaptation et d'améliorer l'assemblage, notamment en produisant des fragments plus grands que les séquences répétées présentes dans le génome, la détection des variants ne sera abordée que sous le prisme du séquençage en lectures courtes (short-reads) dans cette introduction.

1.3.1 Bases génétiques de l'adaptation

L'implication d'un gène dans un phénotype adaptatif peut être dû à plusieurs types de mutations sous-jacentes. Les variations entre les génomes peuvent être classées en fonction de leurs tailles, des mutations ponctuelles (d'un seul

nucléotide) à des variants structuraux (SV) pouvant englober jusqu'à l'entièreté d'un chromosome (Conrad and Hurler, 2007).

Le choix du génome de référence, à partir duquel seront détectées ces mutations, est crucial : dans le cas de *C. lectularius*, le génome de référence a été séquencé à partir d'individus de la souche Harlan, sensible aux pyréthrinoides (Thomas *et al.*, 2020). Ainsi, les allèles impliqués dans la résistance à ces insecticides sont considérés à l'état "ancestral" dans cette souche, et par opposition "mutés" s'ils diffèrent chez une souche montrant un phénotype de résistance.

Pour détecter ces événements au sein des populations, le séquençage en pool (groupe d'individus) peut être privilégié (Schlötterer *et al.*, 2014). S'il est très efficace afin de réduire les coûts, il complique en revanche le calcul des fréquences alléliques, en ne permettant par exemple pas de connaître le statut homozygote ou hétérozygote des individus.

Variants ponctuels

L'ADN est organisé en séquences codantes, régions traduites en protéines, et en séquences non-codantes. Les séquences non-codantes peuvent être impliquées dans de nombreux processus, dont la régulation de la transcription d'ADN en ARN, ou n'avoir aucun rôle. L'ARN non codant (ARNnc), longtemps considéré comme n'ayant justement aucune fonction particulière, peut influencer la régulation des gènes au niveau post-transcriptomique (Djebali *et al.*, 2012).

Les variants ponctuels de l'ADN englobent les insertions, délétions (indels) ou substitutions d'un nucléotide par un autre (SNP, pour polymorphisme nucléotidique d'un seul site). Lorsque ces mutations ont lieu dans la séquence codante, on peut distinguer les SNPs synonymes qui ne modifient pas la séquence protéique, des non-synonymes qui l'altèrent.

La mutation de cible *kdr* décrite plus haut est bon exemple de SNP non-synonyme adaptatif, en favorisant la survie des organismes les portant suite à une exposition aux pyréthrinoides. Chez la mouche *Musca domestica*, on peut retrouver *kdr* en position 1014 du gène *VGSC*, avec la substitution d'une leucine en phénylalanine (Williamson *et al.*, 1993).

Les indels quant à eux peuvent altérer le cadre de lecture s'ils ne sont pas des multiples de trois, et être donc contre-sélectionnés dans les régions codantes (Lin

et al., 2017). Néanmoins, les indels peuvent aussi être avantageux lorsqu'ils apparaissent dans des séquences non-codantes. La présence d'indels dans les régions promotrices des gènes peut par exemple moduler l'expression de ces gènes, et ainsi leur conférer un caractère adaptatif (Zhang *et al.*, 2008; Mullaney *et al.*, 2010). L'insertion d'indels dans l'espace intergénique entre deux estérases chez *C. quinquefasciatus* a ainsi introduit des séquences régulatrices, qui semblent agir comme des promoteurs en modulant l'expression de ces gènes (Vaughan *et al.*, 1997; Hawkes and Hemingway, 2002).

Variants structuraux

Les variants structuraux (SV, pour structural variant) sont divisés en deux principaux types : équilibrés (inversion, insertion d'élément mobile⁹ et translocation) et déséquilibrés (délétion, duplication et insertion d'un nouveau fragment¹⁰) (Figure 1.12).

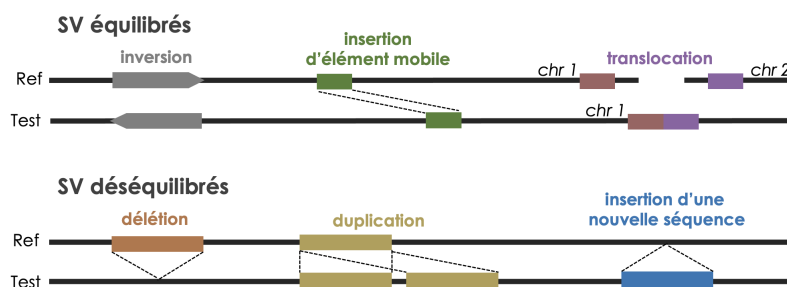


Figure 1.12 – Différents types de variants structuraux. Les variants structuraux (SV) équilibrés (inversion, insertion d'élément mobile et translocation) sont représentés au dessus des déséquilibrés (délétion, duplication et insertion d'une nouvelle séquence) tels qu'ils sont détectés dans un génome test par rapport au génome de référence.

Les variants déséquilibrés impliquent une variation du nombre de copie (Copy number variant, CNV) des gènes qu'ils englobent. Certains types de SVs peuvent également conduire à des combinaisons plus ou moins complexes, comme dans le cas des duplications qui peuvent être en tandem ou inversées. Les SVs peuvent être générés par plusieurs mécanismes, dont la mobilisation d'éléments transposables, la recombinaison de l'ADN, la réplication (événements de glissement de la polymérase), et les processus associés à la réparation (Carvalho *et al.*, 2010; Balachandran and Beck, 2020).

9. Par insertion d'élément mobile, on fait ici référence aux éléments transposables de classe II, ou transposons, capables d'excision et de réinsertion.

10. Les éléments transposables de classe I sont capables de transcription et de rétrotranscription, insérant ainsi une nouvelle "copie" de leur séquence dans l'ADN de l'organisme.

Afin de détecter des SVs suite au séquençage d'un génome test et au mapping de ces reads sur un génome de référence, plusieurs signaux peuvent être utilisés (Figure 1.13). Le premier concerne les lectures uniques à cheval sur l'évènement qui seront découpées en deux (split-reads, flèches oranges). Certaines paires de lectures se trouvent également au niveau des jonctions de l'évènement (flèches bleues), et pourront alors être alignées l'une par rapport à l'autre dans un sens différent de la norme, avec une taille d'insert les séparant plus longue. Enfin, on peut observer une différence de couverture par rapport au reste du génome au niveau du SV.

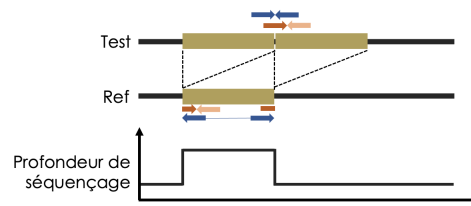


Figure 1.13 – Signaux de détection d'un variant structural. Exemple de détection de la duplication d'un segment (en doré) dans un génome test par rapport au génome de référence, via trois signaux. (i) Une lecture chevauchant les deux copies de la duplication dans le génome test (flèche orange foncée) peut être découpée lors de son alignement sur le génome de référence, avec chaque moitié alignée à une extrémité de la copie unique (split-read). (ii) Une paire de lecture chevauchant les deux copies du génome test (flèches bleues) s'aligne aux deux extrémités, dans le sens opposé (everted read pairs) et avec un insert particulièrement long sur le génome de référence. (iii) Les deux copies du génome test vont s'aligner au même endroit sur le génome de référence, engendrant une augmentation de la profondeur de séquençage par rapport aux autres régions en copie unique.

Les variants structuraux peuvent ainsi, s'ils englobent des gènes, augmenter (dans le cas d'une duplication) ou diminuer (délétion) leurs expressions. On pourra alors s'intéresser au transcriptome de l'insecte, en comparant le nombre de transcrits produits en présence ou en absence de ce variant structural, afin de confirmer ou d'infirmer un impact du SV sur l'expression génique, et donc sur un potentiel mécanisme adaptatif. Le niveau d'expression d'un gène correspond ici au nombre de ses transcrits (ARNm) mesurés dans une cellule, avant que ces derniers ne soient traduits en protéines.

Les variants structuraux peuvent contribuer à l'adaptation, notamment *via* les inversions ou les duplications. Ces changements structuraux peuvent affecter la séquence codante ainsi que le niveau d'expression des gènes, jouant donc un rôle crucial dans l'adaptation des organismes à un nouvel environnement (North *et al.*, 2020). Il existe de nombreux exemples d'inversions impliquées dans l'adaptation, principalement en réduisant la recombinaison entre les loci inversés et non-inversés (Ayala *et al.*, 2014). Le potentiel rôle de cette inhibition de la recombinaison dans

l'évolution de supergènes sera abordé plus en détail ultérieurement (voir fin de la section 1.3.2). Quant aux duplications, elles peuvent avoir un impact sur le niveau d'expression des gènes par un effet gène-dosage (Kondrashov, 2012), mais également permettre à des variants alternatifs d'un gène d'être présents en tandem sur un même chromosome. Deux phénomènes peuvent alors se suivre : un évènement de duplication, puis une divergence entre les copies - l'une portant l'allèle de résistance, et l'autre l'allèle sensible. La deuxième copie de ce tandem hétérogène peut ainsi permettre de compenser un potentiel coût adaptatif de la première en l'absence d'insecticide. (Labbé *et al.*, 2007; Assogba *et al.*, 2015).

Les analyses des SVs à l'échelle du génome d'une population sont rares, et l'étendue de leur impact sur l'évolution donc encore sous-estimée (Lucas *et al.*, 2019). Cependant, plusieurs études ont associés des SVs à la résistance aux insecticides chez les insectes. Chez *Drosophila*, l'insertion d'éléments transposables en amont du gène *Cyp6g1* ainsi que l'augmentation du nombre de copies de ce gène semblent responsables de sa sur-expression, conférant une résistance au DDT (Schmidt *et al.*, 2010). Chez le légionnaire d'automne *Spodoptera frugiperda*, un locus au nombre de copies variables a été identifié, incluant un groupe (cluster) de cytochrome P450s (Gimenez *et al.*, 2020). L'inhibition des P450s chez la souche résistante par l'ajout de PBO a alors permis de confirmer la forte implication de ces enzymes dans le phénotype de résistance à la deltaméthrine. Des études ont également été réalisées chez de nombreuses espèces de moustiques, et ont permis de mettre en évidence des CNVs impliqués dans la résistance aux insecticides chez *Ae. aegypti* (Faucon *et al.*, 2015; Cattel *et al.*, 2020, 2021), *C. pipiens* (Weetman *et al.*, 2018b), *An. gambiae* et *An. coluzzii* (Lucas *et al.*, 2019).

Des analyses à l'échelle du génome peuvent donc permettre de détecter ces variants ponctuels ou structuraux en association avec différents statuts de résistance. Dans le cas des variations de nombre de copies, les études transcriptomiques peuvent mettre en évidence une variation de l'expression d'un gène (sur-expression dans le cadre d'une duplication, ou sous-expression pour une délétion). Néanmoins, une association n'implique pas nécessairement une causalité : une étape de validation fonctionnelle est nécessaire pour évaluer l'impact réel dudit évènement. Pour ce faire, le gène doit tout d'abord être désactivé, par interférence à ARN (RNAi) ou en retirant sa séquence du génome par CRISPR/Cas9 (Itokawa *et al.*, 2016). Le bon déroulement de cette modification expérimentale peut ensuite être confirmé par PCR quantitative, permettant de comparer la quantité d'ADN génomique relative pour chaque gène

entre plusieurs souches (Cattell *et al.*, 2020). Enfin, un test phénotypique permet de mettre en évidence le rétablissement d'une potentielle sensibilité à un insecticide, confirmant ainsi l'implication du gène silencié dans le phénotype de résistance initial.

Les variants alléliques, qu'ils ne soient longs que d'une paire de base ou de plusieurs dizaines, peuvent donc être associés au phénotype recherché. Néanmoins, leur (co-)localisation et la diversité alentours peuvent également donner des indices précieux quant à leurs histoires évolutives, et ainsi permettre de dérouler le fil d'une réponse adaptative à une pression de sélection.

1.3.2 Détection des signatures adaptatives

On peut détecter de la sélection positive en utilisant la signature qu'elle laisse dans les génomes. D'une part, une région soumise à forte sélection positive contiendra des polymorphismes à plus haute fréquence que dans une région neutre. On s'attend alors à une accélération du taux de substitution, pouvant être détectée en contrastant hétérozygotie et divergence (McDonald and Kreitman, 1991). D'autre part, une mutation soumise à une forte sélection positive montera en fréquence, en emmenant avec elle les mutations qui lui sont liées (proche sur le chromosome). Ce phénomène, appelé signature de sélection, correspond donc à une région du génome dans laquelle le polymorphisme observé diffère de celui attendu sous une évolution neutre.

En comparant les fréquences alléliques de populations divergentes, des loci sous sélection devraient donc pouvoir être identifiés (Lewontin and Krakauer, 1973). Ces loci peuvent être séparés en plusieurs catégories, en fonction de leur origine : une mutation *de novo*, une variation génétique existante, ou de l'introgession adaptative (Hawkins *et al.*, 2019).

Dans le cas d'une mutation *de novo* qui conférerait un avantage adaptatif, la sélection positive, ou sélection Darwinienne du nom de celui l'ayant formalisée (Darwin, 1866), engendrerait sa rapide montée en fréquence. En effet, un allèle avantageux qui augmenterait la probabilité de survie d'un organisme se répandrait rapidement au sein de la population (la progéniture portant cet allèle ayant à son tour de meilleures chances de survie et donc de se reproduire), à condition qu'il soit héritable. La mutation de cible *kdr L1014F* du moustique tigre *An. gambiae*, complètement absente des populations sensibles portant l'allèle dit "sauvage",

serait ainsi apparue par mutation aléatoire, et se serait répandue sous une pression de sélection positive dans les populations du Ghana en à peine 5 ans (Lynd *et al.*, 2010).

Une variation génétique pré-existante peut également devenir avantageuse à la suite d'un changement de l'environnement. Chez les moustiques *An. gambiae* et *An. coluzzii*, la mutation de cible G119S du gène *Ace-1* codant pour l'acétylcholine estérase confère une résistance aux insecticides organophosphorés et carbamates (Weill *et al.*, 2004). Ce variant, présent en très faible fréquence (<5%) *in natura*, serait par la suite monté en fréquence dans les populations exposées aux insecticides (Weetman *et al.*, 2015).

À proximité de ces mutations d'intérêt, des allèles neutres (soit sans effet avantageux direct pour la survie et la reproduction de l'individu) peuvent également monter en fréquence, par un effet appelé l'auto-stop génétique (Smith and Haigh, 1974). Sous l'influence de la sélection positive, la fixation de ces allèles neutres entraîne une diminution de la variabilité génétique dans la région autour de la mutation avantageuse, appelée balayage sélectif (Smith and Haigh, 1974). Ainsi, lorsque l'on observe une mutation rare *de novo* montant rapidement en fréquence et engendrant une forte baisse de la diversité génétique locale, on parle de balayage "dur" ou "hard sweep" (Smith and Haigh, 1974) (Figure 1.14).

A contrario, une mutation déjà existante chez plusieurs individus, associée à l'adaptation à un changement de l'environnement, et provoquant une plus faible diminution de la diversité génétique dans cette zone par rapport au reste du génome témoigne d'un balayage "léger", ou soft sweep (Hermisson and Pennings, 2005). En effet, le polymorphisme présent de manière ancestrale est lié à différentes mutations neutres en fonction des recombinants (individus) au sein de la population dans le cas d'un soft sweep.

Des transferts d'allèles de résistance peuvent également avoir lieu entre espèces proches. En effet, suite à une hybridation, on peut observer une introgression adaptative : lors de plusieurs rétrocroisements avec l'espèce cible, une sélection positive a lieu sur un gène en particulier de l'espèce donneuse (Hedrick, 2013). C'est le cas de la mutation *kdr L1014F* décrite ci-dessus. En effet, il existe deux espèces émergentes chez *An. gambiae s.s.*, déterminées en fonction de la présence de mutations du chromosome X et appelées formes M et S. La mutation *kdr L1014F* présente chez la forme M aurait ainsi été acquise *via* un évènement d'introgression à partir de la forme S (Lynd *et al.*, 2010).

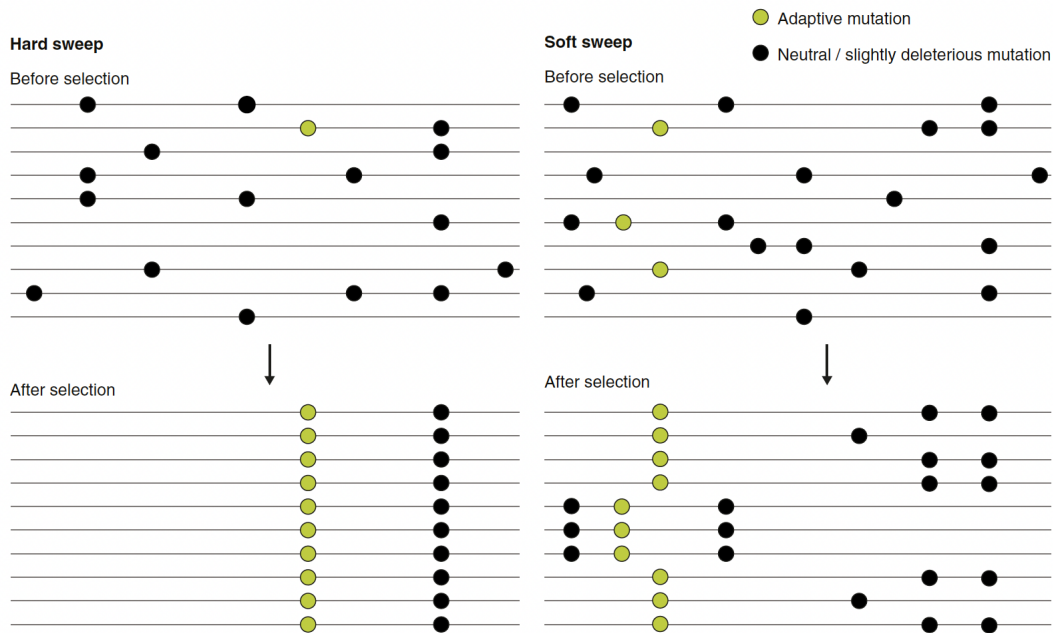


Figure 1.14 – Impact des balayages sélectifs sur la variation génétique (McCoy and Akey, 2017). Chaque ligne représente un génome haploïde, et chaque rond une mutation (adaptative en vert, neutre en noir). Un balayage sélectif dur (à gauche) est représenté par une mutation adaptative unique qui apparaît de novo et dont la fréquence augmente ensuite, entraînant avec elle les mutations neutres auxquelles elle est liée. En revanche, un balayage sélectif doux (à droite) correspond à une sélection sur une variation pré-existante, de sorte que les allèles adaptatifs sont présents sur plusieurs individus distincts qui vont tous monter en fréquence, ce qui n'entraîne qu'une faible réduction de la diversité génétique à proximité.

Néanmoins, l'introgression peut être limitée par la recombinaison, qui sépare les combinaisons d'allèles (Yeaman and Whitlock, 2011). Toutefois, des haplotypes appelés "supergènes" peuvent se différencier et être sélectionnés, régions où l'absence de recombinaison permet de transmettre collectivement un ensemble de loci (Schwander *et al.*, 2014). Un supergène social régulant le nombre de reines reproductrices contenues dans une colonie est ainsi partagé par de nombreuses espèces de fourmis, et se serait répandu par introgression (Kay *et al.*, 2022).

Mais les supergènes peuvent également évoluer au sein d'une même espèce. Chez le papillon *Heliconius numata*, un supergène contrôle ainsi le motif des ailes, lui permettant de mimer d'autres papillons peu appétissants (Le Poul *et al.*, 2014). Les souris sylvestres *Peromyscus maniculatus* possèdent quant à elles une inversion chromosomique contenant différents allèles des gènes codant pour la taille de la queue et la couleur de la fourrure (Hager *et al.*, 2022). Cette inversion est présente à 90% chez l'écotype de forêt, possédant une longue queue et un pelage noir, et absente chez l'écotype de prairie, à la queue courte et au pelage clair. Les inversions

peuvent en effet conduire à une recombinaison réduite, et sont connues pour leur rôle dans l'émergence de supergènes (Schwander *et al.*, 2014).

Pour détecter les traces de l'adaptation, on peut donc observer le taux de recombinaison, mais également la diversité génétique, en mesurant sa variabilité tout au long du génome et/ou entre différents génomes. Différents indices peuvent être calculés, tels que la diversité nucléotidique (π) et la différenciation des populations à partir du polymorphisme génétique (F_{ST} , défini ici en fonction des fréquences alléliques des différentes populations). En fonction de ces indices, mais également de la localisation ou encore de l'état (de référence ou alternatif) des variants détectés, il est ainsi possible de reconstruire l'histoire évolutive d'une espèce, en prédisant quels variants ont permis la survie d'une population face à une nouvelle pression de sélection.

1.4 Objectifs et organisation de la thèse

“Un projet de thèse, ça ne se passe jamais comme prévu” . Bien sûr que l’on a tou.te.s déjà entendu cette phrase, et que l’on essaye de s’y préparer. Pourtant, difficile de prévoir la paralysie générale et les confinements successifs entraînés par le Covid-19, et de résister aux sirènes de la bio-informatique dans ce cadre.

1.4.1 Projet initial

Le projet de thèse initial impliquait la mise en place d’élevages de souches de terrain issues de quatre villes françaises - Paris, Lyon, Bordeaux et Marseille - sélectionnées pour leur statut de métropole ayant une gare et un aéroport. L’entreprise avec laquelle la Cifre est réalisée étant un centre de formation délivrant entre autre le Certibiocide¹¹, elle est en contact avec de nombreuses entreprises qui réalisent des désinsectisations sur le terrain. L’objectif était donc d’échantillonner plusieurs populations de punaises de lit dans chaque ville grâce à ces contacts.

Une fois échantillonnées, ces populations devaient être phénotypées pour la résistance aux pyréthriinoïdes en réalisant des bioessais sur une fraction d’insectes (Figure 1.15).

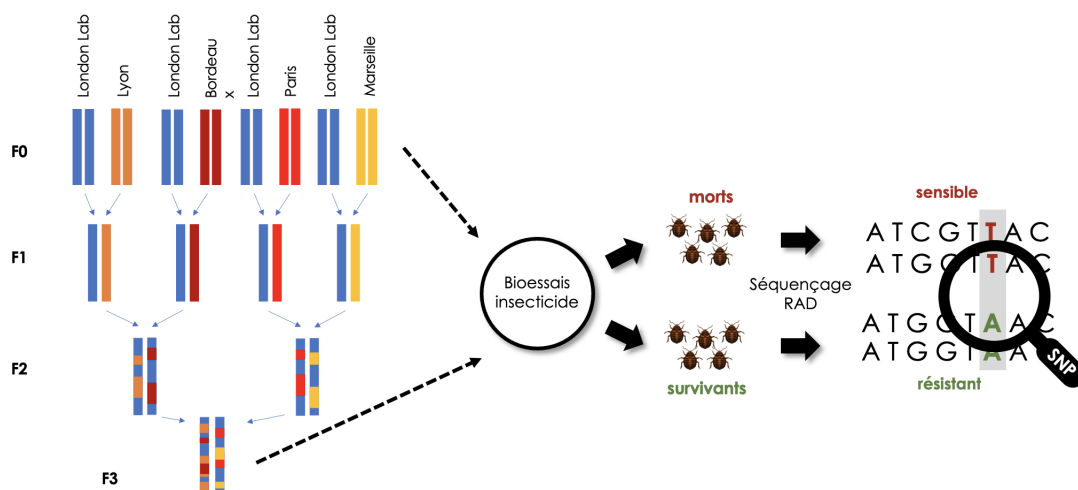


Figure 1.15 – *Projet initial : Design des croisements et protocole de détection des SNPs par RAD-seq.*

La population la plus résistante de chacune des quatre villes aurait ensuite été

11. Certibiocide : diplôme du Ministère en charge de l’environnement, obligatoire pour quiconque veut utiliser, vendre ou acheter certains types de produits biocides, dont les insecticides.

maintenue en laboratoire, puis croisée avec une souche sensible aux insecticides (London Lab, Cimexstore). Pour un gène impliqué dans le phénotype de résistance, l'hypothèse était alors que les individus de la génération parentale F0 sont homozygotes pour l'allèle R chez les punaises résistantes, et homozygotes S chez les punaises sensibles. En partant de ce postulat et suivant les lois de Mendel, 100% des individus de la génération F1 issus du croisement entre les souches parentales F0-sensible et F0-résistante devraient être hétérozygotes RS. Les lignées F1 et F2 devaient ensuite être successivement croisées entre elles, afin de produire une population F3 "super-recombinante", où tous les individus correspondent à des combinaisons alléliques différentes.

Des bioessais devaient ensuite être réalisés avec des individus issus de chaque lignée parentale F0 et de la lignée F3, avant d'en extraire l'ADN et de réaliser un séquençage RAD (Restriction-site Associated DNA). Les loci auxquels un polymorphisme d'un seul nucléotide (SNP) était détecté auraient ensuite été associés au phénotype de résistance observé *via* les bioessais. À proximité de ces loci de traits quantitatifs (QTL) se trouvent les gènes candidats potentiellement responsables de la résistance. Ces données auraient été complétées grâce à l'analyse des séquençages du génome complet de plusieurs individus en parallèle, afin d'identifier plus précisément les mécanismes (SNPs non-synonymes dans la séquence codante, variants structuraux..) présents sur les séquences sous-jacentes à ces variations.

Pour chaque ville, plusieurs enveloppes prépayées contenant chacune un pot de prélèvement ont été envoyées. Sur les 26 contacts ayant accepté de participer à la campagne d'échantillonnage, seuls 8 ont renvoyé le pot de prélèvement, avec entre 2-5 punaises de lit au lieu des 20 adultes minimum demandés. Accompagnée par l'entreprise Action 4D, j'ai alors réalisé un échantillonnage dans un appartement de Vénissieux (69) présentant une très forte infestation, au printemps 2021. Malgré les souvenirs impérissables créés - une couche de mues au sol de 2 cm d'épaisseur, des portions entières de papier peint recouvertes de déjections de punaises de lit, quelques insomnies par la suite (Figure 1.16) - le demi millier de punaises de lit récoltées ce jour là n'a pas suffi à lancer un élevage d'insectes de terrain viable. En parallèle, la qualité des élevages mis en place la première année s'était en effet grandement détériorée, avec une mortalité suspecte pouvant atteindre 50% par semaine à la suite d'un gorgement. Le projet a donc été réorienté en septembre 2021, à cause de ces difficultés persistantes.



Figure 1.16 – *Extraits d'une vidéo prise sur le terrain lors d'un échantillonnage de punaises de lit. A. Échantillonnage à la pince de punaises de lit sur le sommier. B. Matelas recouvert de déjections, mues (en bas à gauche par exemple), et insectes adultes. C. Regroupement de déjections de punaises de lit sur le papier peint. D. Sol de l'appartement (initialement beige), recouvert de mues et d'insectes morts.*

1.4.2 Réorientation du sujet

En amont de la thèse, des punaises de lit appartenant à quatre souches distinctes ont été achetées sur Cimexstore (<https://cimexstore.co.uk>) : deux souches de laboratoire sensibles aux insecticides, London Lab et German Lab, et deux souches de terrain résistantes aux insecticides, London Field et Sweden Field. Les génomes complets de plusieurs individus pour chaque souche ont ensuite été séquencés (pool-seq).

Ces séquences devaient servir de jeu de données test, afin de prendre en main différents outils bio-informatiques. Néanmoins, dans le contexte lié au Covid-19, et suite aux problématiques décrites ci-dessus, l'étude de ces génomes a finalement été faite de manière exhaustive, constituant le socle de cette thèse.

1.4.3 Plan de la thèse

Mon travail pendant cette thèse s'est attaché à réaliser une étude à l'échelle du génome des mécanismes impliqués dans la résistance aux insecticides, en se focalisant en particulier sur les mutations ponctuelles, les variants structuraux, et leur impact potentiel sur l'expression des gènes. Ces résultats sont présentés sous la forme d'articles rédigés en anglais puisqu'ils sont destinés à être - ou sont déjà - publiés dans des revues scientifiques. Le premier article est donc une exploration du génome de quatre souches de punaises de lit, résistantes ou non aux insecticides, en analysant le séquençage de pools de 30 individus (Chapitre 2). Il a vocation à décrypter les bases génomiques de la résistance aux insecticides par une approche sans *a priori*, en étudiant (i) les SNPs et (ii) les variants structuraux les plus différenciés entre des souches de punaises de lit aux statuts de résistance aux pyréthriinoïdes contrastés.

Le second manuscrit vise à compléter cette analyse génomique par une approche transcriptomique de l'expression des gènes de résistance aux insecticides (Chapitre 3). Cette analyse a pour objectif (i) d'identifier des amplifications constitutives entre les souches sensibles et résistantes, (ii) d'identifier des amplifications induites différenciellement entre ces souches et les conditions d'exposition aux insecticides, et (iii) de valider l'implication de potentiels variants structuraux candidats identifiés préalablement.

Le dernier chapitre est plus exploratoire, et consiste à analyser les micro-organismes constituant le cortège de la punaise de lit pour identifier de potentiels pathogènes (Chapitre 4). En effet, une mortalité spécifique traduite par un phénotype dit "oxydé" a été observée dans les élevages. On soupçonne un agent pathogène d'être à l'origine de cette mortalité. Des données méta-génomiques ont donc été extraites d'individus sains, mais également de punaises de lit présentant ce phénotype anormal, afin de proposer des candidats potentiellement responsables de la mortalité des punaises de lit observée.

2

Génomique de la résistance aux insecticides chez la punaise de lit

ARTICLE 1 : Un superlocus majeur de 6 Mb est impliqué dans la résistance aux pyréthriinoïdes chez la punaise de lit <i>Cimex lectularius</i>	78
--	----

Avant-propos

L'objectif de ce premier chapitre était d'identifier les variants génomiques associées à la résistance aux pyréthrinoïdes chez *C. lectularius*. Au cours de la section introductive 1.2.4, nous avons vu que de nombreuses études ont été réalisées afin de tester l'implication de certains mécanismes spécifiques dans le phénotype de résistance chez la punaise de lit. Néanmoins, la littérature manquait cruellement d'analyse à plus large échelle et sans *a priori*. Une seule étude à notre connaissance a rempli ces critères, et réalisé un croisement entre une souche sensible et une souche résistante au pyréthrinoïde deltaméthrine suivie d'un séquençage RAD, afin de détecter les loci associés à la résistance à cet insecticide. Un QTL majeur a alors été détecté, contenant plusieurs gènes considérés comme des candidats intéressants pour la résistance aux pyréthrinoïdes. Cependant, cette étude étant uniquement basée sur des marqueurs RAD, elle ne comprenait pas les séquences complètes de ces variants, ce qui ne permettait pas d'identifier les mécanismes sous-jacents.

Dans un premier temps, nous avons donc réalisé des expositions à l'insecticide deltaméthrine des quatre souches, afin de déterminer leurs statuts de résistance respectifs. Deux souches (German Lab et London Lab) ont été détectées comme sensibles, avec un niveau de résistance comparable à celui de la souche du génome de référence Harlan, tandis que deux autres (London Field et Sweden Field) sont résistantes à ce pyréthrinoïde, avec des ratios de résistance variant entre 17 et 199 X.

Nous avons ensuite détecté les polymorphismes d'un seul site (SNPs) entre ces quatre souches de punaises de lit, et calculé les indices de différenciation (F_{ST}) associés. Cette analyse a révélé une proximité génétique forte entre les souches London Lab et London Field, malgré la forte différenciation phénotypique observée préalablement. Les SNPs identifiés ont ensuite été filtrés en fonction de leur différentiel de fréquence corrélativement au phénotype de résistance grâce à plusieurs algorithmes et critères (indice de Bayes et statistique de contraste par BayPass, fréquence de l'allèle alternatif et F_{ST}), mettant en évidence un demi-millier de SNPs "outliers". Un des chromosomes putatifs (LG 11, LG pour groupe de liaison) de la punaise de lit a alors été détecté comme particulièrement enrichi en SNPs outliers, notamment au sein d'une région composée de trois scaffolds consécutifs. Cette région de 6 Mb, appelée "superlocus", a également été identifiée comme portant un nombre significativement plus élevé de gènes connus comme étant impliqués dans la résistance aux insecticides, et en particulier des enzymes de détoxification de type P450 et GST. Néanmoins, aucun SNP outlier n'a été

déecté au sein de ces gènes, bien que la présence de nombreux SNPs non-synonymes ayant une fréquence supérieure chez London Field ne soit relevée.

Dans un troisième temps, le génome a donc été scanné pour la présence de variants structuraux (SV), connus pour impacter potentiellement l'expression des gènes qu'ils englobent. La répartition de ces SVs le long du génome, tout comme celle des SNPs outliers, n'a alors pas semblé aléatoire, et plusieurs scaffolds ont été détectés comme enrichis en SV. C'est le cas de l'un des scaffolds du superlocus, pour lequel un enrichissement tout particulièrement en inversions a été identifié. Néanmoins, aucun des SVs les plus différenciés entre London Lab et London Field n'était localisé dans le superlocus, traduisant une potentielle absence de co-occurrence entre ces mécanismes.

Nous avons finalement émis l'hypothèse que le superlocus identifié était un supergène, où les allèles impliqués dans la résistance aux pyréthriinoïdes seraient regroupés, et qui serait maintenu par des inversions, réduisant la recombinaison au sein de cette région.

Le manuscrit suivant a été proposé pour publication à *Evolutionary Applications* début janvier 2023, où il a été accepté avec révision mineure.

A major 6 Mb superlocus is involved in pyrethroid resistance in the common bed bug *Cimex lectularius*

Chloé Haberkorn^{1,2,*} | Jean-Philippe David³ | Hélène Henri¹ | Jean-Marie Delpuech¹ | Romain Lasseur² | Fabrice Vavre¹ | Julien Varaldi¹

¹Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France

²IZInnovation, 13 Rue des Émeraudes, 69006 Lyon, France

³Laboratoire d'Écologie Alpine, UMR UGA-USMB-CNRS 5553 Université Grenoble Alpes CS 40700. 38058 Grenoble cedex 9, France

Correspondence

Chloé Haberkorn
Email: chloehbk@gmail.com

Funding information

This work was supported by a CIFRE grant N° 2019/0800, funded by Association Nationale de la Recherche et de la Technologie, from the collaboration between the Laboratoire de Biométrie et Biologie Évolutive and IZInnovation SAS ; and the Scientific Breakthrough Project Micro-be-have (Microbial impact on insect behavior) of Université de Lyon, within the programme 'Investissements d'Avenir' (ANR-11-IDEX-0007; ANR-16-IDEX-0005)

In the last few years, the bed bug *Cimex lectularius* has been an increasing problem world-wide, mainly due to the development of insecticide resistance to pyrethroids. The characterization of resistance alleles is a prerequisite to improve surveillance and resistance management. To identify genomic variants associated with pyrethroid resistance in *Cimex lectularius*, we compared the genetic composition of two recent and resistant populations with that of two ancient-susceptible strains using a genome-wide pool-seq design. We identified a large 6 Mb "superlocus" showing particularly high genetic differentiation and association with the resistance phenotype. This superlocus contained several clustered resistance genes, and was also characterized by a high density of structural variants (inversions, duplications). The possibility that this superlocus constitutes a resistance "supergene" that evolved after the clustering of alleles adapted to insecticide and after reduction in recombination is discussed.

KEYWORDS

Cimex lectularius; bed bug; pool-seq; insecticide resistance; SNP; duplication; inversion; superlocus

1 | INTRODUCTION

The common bed bug *Cimex lectularius* (Hemiptera) is a blood-sucking human parasite that can cause both physical and psychological disorders, such as allergic reactions (Alexander, 1984), sleep deprivation (Susser et al., 2012) or

even paranoia (Goddard and Deshazo, 2009). Although the massive use of insecticides in the 20th century helped to regulate their populations, a phenomenon of resurgence of the species has been observed since the late 1990s (Potter, 2011). This population boom in bed bugs is thought to be due to the increase in international travel, the growing popularity of second-hand market (Doggett et al., 2004), and the spread of insecticide resistance (Davies et al., 2012). Pyrethroid insecticides are widely used against bed bugs, thereby exercising a strong selection pressure. Hence, the current rise of bed bugs is believed to be mainly imputable to the emergence of pyrethroid resistance (Romero et al., 2007).

Insecticide resistance mechanisms have long been studied in a variety of insects, and are divided in two types: behavioral resistance, and physiological resistance (Lockwood et al., 1984). Behavioral resistance is defined as the ability to avoid or reduce lethal exposure to an insecticide. Physiological resistance consists in decreasing insecticide intake by preventing its entry, with decreased cuticular penetration (Balabanidou et al., 2018), or limiting its effect, through increased detoxification metabolism and decreased target site sensitivity (Feyereisen, 1995).

To date, no convincing evidence of behavioral resistance has been found in *C. lectularius* (see review Dang et al. 2017), although some behaviors that may limit the effect of the insecticide have been observed in both resistant and susceptible strains. Pyrethroid insecticide could indeed have an excitatory-repulsive effect on bed bugs, with increased movement whenever in contact, or the choice of an untreated place to hide (Romero et al., 2009a).

Cuticular resistance can be the consequence of a thickening or change in cuticular structure that decreases the rate of insecticide penetration. A morphological study confirmed that cuticle was thicker in a pyrethroid resistant strain of bed bugs collected in Australia, compared to a susceptible one (Lilly et al., 2016b). An increased cuticle thickening was correlated with a longer time-to-knockdown after insecticide exposure. Further studies based on RT-qPCR assays showed that cuticular protein genes, such as chitin synthase (CHS) or larval/pupal cuticle protein (L/PCP), had higher transcript levels in resistant bed bug strains (Mamidala et al., 2012; Koganemaru et al., 2013).

Following penetration of the cuticular barrier, detoxification metabolism may be mobilized to degrade, sequester and enhance the excretion of the insecticide. In resistant populations of bed bugs, an increased activity of several detoxification enzymes such as cytochromes P450s (Romero et al., 2009b) and esterases (Lilly et al., 2016a) has been observed. Higher expressions of glutathione-S-transferases (GSTs) have also been detected in resistant bed bug immature stages, but this mechanism is yet to be identified in adults (Mamidala et al., 2011). Some studies also used specific enzymatic inhibitors to determine the involvement of particular detoxification enzyme families in the resistant phenotype. Piperonyl butoxide (PBO), an inhibitor of P450s, significantly decreased resistance against the pyrethrinoid deltamethrin (Romero et al., 2009b). Similarly, an inhibitor of esterases increased bed bug mortality to the pyrethroid lambda-cyhalothrin (Lilly et al., 2016a).

Pyrethroid being the main chemical insecticide used today against bed bugs, its target protein site is also particularly studied (Dang et al., 2014, 2015; Akhouni et al., 2015; Balvín and Booth, 2018). Pyrethroids target the voltage-gated sodium channels (VGSC) in the nervous system of the insect. When pyrethroids bind to the VGSC, they induce paralysis of the insects by maintaining the channels in an open state (knock-down effect). A single nucleotide non-synonymous mutation can alter VGSC conformation, and therefore hinder pyrethroid binding, thus conferring knock-down resistance (*kdr* mutations).

Three *kdr* mutations have been associated with pyrethroid resistance in *C. lectularius*: V419L (valine in position 419 to leucine), L925I (leucine in position 925 to isoleucine), and I936F (isoleucine at position 936 to phenylalanine). V419L and L925I *kdr* mutations were detected in Eurasia and USA in early 2000s, while the I936F mutation was detected in Australia during the same period (Zhu et al., 2010, 2013; Dang et al., 2015). According to Durand et al. (2012), all bed bugs sampled in Paris (n=124) during 2011 were homozygous for the L925I mutation.

In bed bugs, many reports showed resistance through phenotypic assays to various insecticides (Romero and

Anderson, 2016; Lilly et al., 2018; Cáceres et al., 2019). Using knowledge obtained on other insect systems, the association between resistance phenotype and genes has been explored for a few candidate genes, such as VGSC (Zhu et al., 2010). Other wider approaches identified genes potentially involved in resistance by comparing the transcriptomic profile of resistant and susceptible lines (Adelman et al., 2011; Mamidala et al., 2012).

In an attempt to further identify the underlying genetic factors of resistance, Fountain et al. (2016) went through a RAD-tag based QTL mapping of resistance to deltamethrin by crossing a susceptible strain with a pyrethroid resistant one, both originating from London. Using 334 SNP markers, they found a major QTL associated with insecticide resistance on a Linkage Group (LG 12). Interestingly, this 10 Mb region contains several genes considered as valuable candidates for pyrethroid resistance, such as the VGSC, a carboxylesterase and several P450s. However, identified genes alone failed to fully explain variations in resistance phenotype, with the major QTL detected explaining only 64.2% of phenotypic variation. Moreover, since this study was based on RAD-tags only, it did not provide the full sequences of these putative resistance variants. Nor did this analysis provide any information on structural variants.

Indeed, structural variants such as duplications and inversions have been shown to be potentially involved in insecticide resistance, as in the mosquito *Anopheles gambiae* (Brooke et al., 2002; Assogba et al., 2015). Inversions can inhibit recombination between standard and inverted regions in the heterozygote state, thus having a potential protective effect on a large number of genes (Ayala et al., 2014; Mérot et al., 2020). Other structural variants, as gene copy number, could also lead to insecticide resistance. Changes in gene copy number (Copy Number Variation, or CNV), and specifically duplications, can lead to overexpression of genes through a "gene-dosage" effect (Kondrashov, 2012). Amplification of genes encoding detoxification enzymes can enhance insecticide metabolism in many insects, and most notably via three gene families: esterases, glutathione S-transferases (GSTs), and cytochrome P450 monooxygenases (P450s) (Bass and Field, 2011). However, the role of structural variants in the capacity of bed bugs to resist insecticides has not yet been investigated.

In this study, we aim to provide a genome-wide analysis of the genetic factors underlying deltamethrin resistance in bed bugs. We performed DNA pool-sequencing on four strains of *C. lectularius* with different insecticide resistance status. Based on both single nucleotide polymorphisms (SNPs) and structural variants (SVs) detection, we used population genomics to identify allelic variants and genomic regions associated with the resistance phenotype.

2 | MATERIALS AND METHODS

2.1 | Insects

The four strains used in this study were provided by CimexStore Ltd (Chepstow, United Kingdom). Two of these strains were susceptible to pyrethroids (S), as they were collected before their massive use and have been maintained under laboratory condition without insecticide exposure for more than 40 years : German Lab (GL, collected in Monheim, Germany) and London Lab (LL, collected in London, Great Britain). The other two resistant (R) populations were London Field (LF, collected in 2008 in London) moderately resistant to pyrethroids, and Sweden Field (SF, collected in 2015 in Malmö, Sweden), with a moderate-to-high resistance level, according to CimexStore statement. These status were confirmed by our bioassays (see below).

Insects were kept isolated before imaginal moult in 24-wells Petri dishes containing accordion-folded blotting papers, serving as harborage. Bed bugs were maintained at 25°C, 40% relative humidity (RH), and a photoperiod of 12:12h. Since males perform the so-called traumatic insemination to copulate, which can be costly for their partners (Stutt and Siva-Jothy, 2001), bioassays were conducted only on 7-day-old unfed virgin females. Insecticide resistance is thought to be autosomal in bed bugs (Feroz, 1969; Fountain et al., 2016).

2.2 | Topical Assay

The insecticide resistance level of the "field" strains and the susceptibility of the "lab" strains was first confirmed by determining the resistance ratio between these strains. Topical insecticide bioassays were carried out with deltamethrin (98% purity, Cluzeau, Sainte-Foy-La-Grande, France), a pyrethroid, as it remains the most used insecticide family recently for bed bug control. Three replicates of 10 insects were used both for control and per insecticide dose, ranging from 0.01 ng to 120 ng per μL . Bed bugs were previously immobilized by placing them in a Petri dish on ice for 5 minutes. Topical applications were then made onto the ventral surface of the thorax, between the coxae, with a 50- μL glass syringe attached to a repeating dispenser (Hamilton Co., Reno, NV). Treated insects were exposed to 1 μL of insecticide diluted in acetone, whereas control insects received 1 μL of acetone only. Mortality was assessed after 24 h by flipping each insect on dorsal side with a featherweight forceps, to see if it was able to reverse on the ventral side (alive) or if its movements were not coordinated enough to do so (moribund, considered as dead).

Dose–mortality data were analyzed by using binomial Generalized Linear Models with a probit transformation using the R/ecotox package v1.4.4 (Hlina et al., 2021). Statistical tests and resistance ratios were computed on lethal doses for 95% of individuals (LD_{95}). Resistant ratios were calculated on pairs as follows: highest LD_{95} (i.e. of the more resistant population of the pair) divided by the lowest LD_{95} (more susceptible population).

2.3 | DNA extraction and sequencing

For each strain, genomic DNA was extracted from 30 individual females (except for London Lab which had only 28) using NucleoSpin 96 Tissue Kit (Macherey Nagel, Hoerd, France) and eluted in 100 μL of BE buffer. DNA concentration of these samples was measured using Quant-iT PicoGreen Kit (ThermoFisher, Waltham MASS, USA) according to manufacturer's instructions.

Samples were then gathered with an equal DNA quantity into pools. DNA purification was performed for each pool with 1.8 times the sample volume in AMPure XP beads (Beckman Coulter, Fullerton CA, USA). Purified DNA were retrieved in 100 μL of ultrapure water. Pool concentrations were measured with Qubit using DNA HS Kit (Agilent, Santa Clara CA, USA). Final pool concentrations were as follow: 38.5 ng/ μL for London Lab, 41.6 ng/ μL for London Field, 40.3 ng/ μL for German Lab and 38 ng/ μL for Sweden Field.

Sequencing was performed using TruSeq Nano Kit (Illumina, San Diego CA, USA) to produce paired-end read of 2×150 bp length and a coverage of 25 X for London Lab, 32 X for London Field, 39.5 X for German Lab and 25.4 X for Sweden Field by Genotoul (Castanet-Tolosan, France).

2.4 | Pool-seq data processing

The whole pipeline with the detail of parameters used is available on GitHub (<https://github.com/chaberko-lbbe/clec-poolseq>). Quality control analysis of reads obtained from each line was performed using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). The raw data have been submitted to the Sequence Read Archive (SRA) database of NCBI under BioProject PRJNA826750. Sequencing reads were filtered using Trimmomatic software v0.39 (Bolger et al., 2014), which removes adaptors. FastUniq v1.1 was then used to remove PCR duplicates (Xu et al., 2012). Reads were mapped on the *C. lectularius* reference genome (Clec_2.1 assembly, Harlan strain) performed as part of the i5K project (Poelchau et al., 2015), with an estimated size of 510.83 Mb. Mapping was performed using BWA mem v0.7.4 (Li and Durbin, 2009). Sam files were converted to bam format using samtools v1.9, and cleaned of unmapped reads (Li et al., 2009). The 1573 nuclear scaffolds were kept in this analysis, while the

mitochondrial scaffold was not considered.

2.5 | Detecting Single Nucleotide Polymorphisms

2.5.1 | Overall SNPs analysis

Bam files corresponding to the four populations were converted into mpileup format with samtools v1.9. The mpileup file was then converted to sync format by PoPoolation2 version 1201 (Kofler et al., 2011). 8.03 million (M) SNPs were detected on this sync file using R/poolfstat package v2.0.0 (Hivert et al., 2018) and the following parameters: coverage per pool between 10 and 50. In parallel, nucleotidic diversity (π) was computed using PoPoolation2 Variance-sliding.pl (default parameters).

Principal Component Analysis (PCA) was performed using R/pcadapt package v4.3.3 and "pool" type parameter (Privé et al., 2020). Fixation indexes (F_{ST}) were computed with R/poolfstat for each pairwise population comparison of each SNP.

Global SNP pool was then trimmed on minor allele frequency (MAF) of 0.2 (computed as $MAF = 0.5 - |p - 0.5|$, with p being the average frequency across all four populations). This relatively high MAF value was chosen in order to remove loci for which we have very limited power to detect any association with the resistance phenotype in the BayPass analysis. BayPass v2.3 (Olazcuaga et al., 2020) was performed with the standard covariate model, considering covariate data such as strain resistance status. The final dataset was thus reduced to 2.92M SNPs located on 990 scaffolds.

2.5.2 | Selecting outlier SNPs

Several filters were combined in order to detect SNPs potentially associated with insecticide resistance, hereafter called outlier SNPs. We first used R/poolfstat to identify the most differentiated alleles between the two most closely related populations, namely London Lab and London Field. Then, BayPass was used, as it takes into account the covariance of allelic frequencies between populations, arising because of common demographic history, and enabled to identify SNPs potentially under selection (i.e. SNPs showing extreme genetic differentiation). Thirdly, we selected alleles in a derived state compared to the allele carried by the reference Harlan strain, and in higher frequency in London Field than in London Lab. These steps are described more in detail below.

At first, to identify mutations potentially involved in resistance, we used fixation index (F_{ST}) calculated with R/poolfstat for each pairwise population comparison for each SNP. SNPs with F_{ST} values belonging to the top 5% of the whole genome F_{ST} distribution computed between London Lab and London Field strains were considered as "outliers". This comparison between London strains was chosen because of their high genetic proximity (see Results).

In a second step, we analyzed the whole dataset (the four strains) using the package BayPass. This package aims at identifying genetic markers associated with covariates (in our case, the resistance phenotype), while taking into account the shared history of populations. The contrast statistic C2 was computed to compare allele frequencies between resistant (London Field and Sweden Field) and susceptible strains (London Lab and German Lab). We corrected the p-values associated with C2 using R/qvalue version 2.26.0 to compute a local false discovery rate (lfdr) for each p-value. SNPs were first selected when having a lfdr below 0.2.

For each SNP across the four populations, the likelihood of the data was then estimated under a first model without selection, as well as under a second alternative model invoking selection. Bayes factor (BF) was computed to represent the ratio of these two likelihoods. A ratio equal to 2 indicates that the data favor the second model (with

selection) twice as much as the first model (without selection). According to Kass and Raftery (1995), we considered SNPs having a $\log_{10}(\text{Bayes factor}) > 0.5$ as valuable candidates. In deciban (dB) units, used in BayPass (ie $10 \cdot \log_{10}(\text{BF})$), this corresponds to a threshold of 5.

Finally, we are assuming that the reference genome strain (Harlan strain), which is insecticide susceptible, carries the ancestral alleles for genes involved in insecticide resistance. Indeed, the Harlan strain was collected back in 1973 in the field, and thus has never been in contact with modern insecticides (Dang et al., 2017). From this perspective, we expected the frequency of alternative alleles (i.e. different from the reference genome allele at a position) to have increased over time under insecticide selective pressure. Therefore, only derived alleles with higher frequency in London Field were kept in the outlier list. SNPs that passed all combined filters described above were further considered as valid outliers, hence as associated with the resistance phenotype. A sliding window was computed for the number of outlier SNPs in 100 kb with a step of 10 kb, using `sw` function from the R package `pegas` v.1.1 (Paradis, 2010).

To locate SNPs in genes, we used the GFF3 file from NCBI (*C. lectularius* RefSeq genome, assembly accession GCF_000648675.2). It includes the type of feature and the coordinates of each genomic region (promoter, 3' or 5'UTRs, intronic or exonic regions, and intergenic). Variants and their impact on protein sequences were analyzed using R/VariantAnnotation v1.34.0 (Obenchain et al., 2014) and R/GenomicFeatures v1.40.1 (Lawrence et al., 2013).

Among the 2.92M SNPs, 1.52M (53%) were located within genes, with a huge majority inside intronic regions (1.45M, i.e. 51%). We identified 576 outlier SNPs (0.02%), 369 of which were located within genes (0.01% of all SNPs).

2.6 | Mapping scaffolds on linkage groups

Based on a QTL RAD-seq analysis, Fountain et al. (2016) constructed a genetic map for *C. lectularius* assembly *Clec_1.0*. Nevertheless, the genome assembly of *C. lectularius* has been updated in the meantime (v1.0 to 2.1). In order to improve the genetic map produced, we decided to use Fountain's RAD-seq data and R scripts to generate a genetic map based on the latest genome assembly. In order to assemble scaffolds into putative autosomes, hereafter called linkage groups (LG), we re-aligned Fountain's RAD sequencing tags (available here <http://dx.doi.org/10.5061/dryad.d4r50>) with GSNAP version 2020-06-01 on the RefSeq genome version 2.1, rather than version 1.0 used by Fountain. By following the same pipeline, we were able to keep 338 out of 441 RAD markers, distributed over 14 linkage groups, as in Fountain analysis. Finally, out of 1573 nuclear scaffolds, 151 were kept, as containing uniquely mapped markers. The 14 putative autosomes obtained had an overall similar gene content as in Fountain et al. (2016) (Supp. fig. 4). While Fountain managed to map >65% of the genome on linkage groups, we only obtained 46% (i.e. 234.7 Mb out of 510 Mb). One explanation could reside in the higher fragmentation of the new genome version. Indeed, the latest version of the genome was split into more scaffolds (*Clec_1.0*: 1402 scaffolds, *Clec_2.1*: 1574 scaffolds - see https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Cimex_lectularius/101/). RAD markers therefore potentially fall into smaller scaffolds, which can lead to this loss of information.

An analysis of enrichment in outlier SNPs was performed on LG using R/stats default package (binomial test with Bonferroni correction for multiple testing). Then, to further identify smaller genomic regions of interest, enrichment tests were done at the scaffold level.

2.7 | Functional analysis

In order to test whether some relevant biological functions were enriched in a set of candidate genes, we defined ten categories of genes potentially involved in resistance, similarly to what was done by [Faucon et al. \(2015\)](#), and benefiting from the extensive literature on the topic in several insects. Six of them corresponded to metabolic resistance : "Binding/Sequestration" (6 genes) , "GST" (14), "CCE" (39), "UDPGT" (39), "ABC transporter/MRP" (55), and "P450" (56). The category "Other detox" (64) includes transcription factors, which could modulate detoxification enzyme expressions ([Amezian et al., 2021](#)). While "Cuticle" (113 genes) encompasses cuticular resistance genes, "Insecticide target and nervous system" (31) covers all the targets of the different insecticide families, as well as genes that could contribute to maintain neurological activity. Finally, bed bugs being hematophageous insects, they are subject to oxydative stress after a blood meal, and therefore adapted to deal with this stress. This feature could be advantageous for their defense against oxidative stress caused by pyrethroid exposure ([Wang et al., 2016](#)). We thus defined a "Redox homeostasis" category (14 genes). Based on gene annotation extracted from the GFF3 file (GCF_000648675.2), 431 out of the 13,208 genes (without *tRNA* and mitochondrial scaffold) were assigned to one of these categories (Supplementary table 1). Analysis of enrichment of candidate genes in each resistance category were performed by comparing their proportions in genomic regions of interest with whole genome, using R/stats default package (proportion test with FDR correction for multiple comparisons).

2.8 | Structural variants

Emphasis was placed on two types of structural variants (SVs), duplications and inversions. London Field population was compared to the susceptible reference genome, and to the susceptible London Lab population, in order to detect SVs that could have been selected and underlie the resistance phenotype.

2.8.1 | Combining read depth and pairs orientation to detect SVs

The presence of SVs was inferred based on abnormal read pair orientation and/or distance (insert size), and read depth variation. We used poolCNVcomp ([Schridder et al., 2013, 2016](#); [North et al., 2020](#)), a series of scripts designed for the detection of tandem duplications in pool-seq data, and adapted them to detect inverted duplications and simple inversions as well. The first step identifies read pairs having abnormal orientation and extreme insert size (top 1% distribution) in each population. In a tandem duplication, read pairs that are falling between the two copies will both map on the reference genome in an reversed way (i.e., in opposite direction). For inverted duplication, one of the two paired reads will be in the wrong orientation. The 2nd step groups them together if they fall inside the same genomic region, using parameter `insertSizeDiffCutoff` (standard deviation of insert sizes distribution respectively to each scaffold, multiplied by four). At the end of these two first steps, candidate structural variants have been delineated, in comparison to the reference genome. The 3rd step then matched corresponding structural variants between population pools, using the parameter `distancecutoff` (redefined as the sum of `insertSizeDiffCutoff` from populations A and B), to determine if the coordinates are close enough to represent the same event in resistant and susceptible populations.

The 4th step aims at computing read depth in each pool. During this process, repeated regions were masked. To obtain those repeated regions, we performed a blast of transposable elements detected in *C.lectularius* by [Petersen et al. \(2019\)](#) on the latest genome assembly, i.e. *Clec_2.1*. All hits with more than 80% nucleotidic identity were considered as repeated regions and masked.

Next, in North et al. (2020) 4pt5 step, we used the coverage data calculated for each pool to test whether London Field had more copies compared to London Lab in putative duplications. Under this hypothesis, the coverage ratio London Field/London Lab on a locus of interest should be disproportionately high compared to a null expectation. To determine a threshold, we estimated empirically the read depth ratio London Field/London Lab under the null hypothesis of no amplification between London Field and London Lab strains. This expectation was computed for a set of windows of different sizes (500 bp to 2.5 Mb, $n=10\,000$), randomly sampled in the genomes. We then compared observed values with the corresponding null distribution (for a given size class), and considered that events with ratios falling above the 75% upper quantile were valid candidates for high copy number amplifications.

We finally filtered out events smaller than the size class of 3.5 kb (half the median size of genes, being 7 kb), leading to our global set of events. We used the same script to identify inversions not associated with duplications. The single difference with the previous pipeline was that the coverage ratio had to be between the 25% and 75% quantiles.

The presence of transposable elements may interfere with the calling of SVs due to multiple mapping. We therefore removed structural variants if their coordinates on extremities were overlapping for more than 100 bp with TE (on 150 pb).

2.8.2 | Identifying highly differentiated SVs

A rough estimate of the allelic frequency of each event was done by computing the coverage in the 150 bp inward flanking regions of each event for each strain, using samtools v1.9. For each pool, we then estimated the frequency of reads supporting the SV by dividing the number of reads with peculiar orientation by the mean value of the coverage in the flanking regions. Events showing a higher frequency in London Field than in London Lab were selected.

We then calculated the F_{ST} index of differentiation based on these estimated frequencies. F_{ST} falling in the top 10% were considered as valid candidates ($q_{90} = 0.040$ for inversions, 0.040 for inverted duplications and 0.036 for tandem duplications).

We tested whether scaffolds were enriched with structural variants or not, by comparing the observed number of SV per scaffold with the expected one under the hypothesis of homogeneous distribution among the scaffolds (expected number being the total number of SV divided by the total size of scaffolds, multiplied by each scaffold size). Scaffolds having less than 5 expected events were then removed. We then computed a χ^2 test for each scaffold (1 d.f.), and corrected the associated p-values using Bonferroni method. The same method was used to test whether scaffolds were enriched in a specific type of structural variant (tandem duplications, inverted duplications, or inversions).

Finally, we detected for each category events that were overlapping with genes. To do so, we extracted all genes falling into windows encompassing an event and 1 kb further upstream and downstream. We also looked at SVs affecting candidate resistance genes.

3 | RESULTS

3.1 | Resistance phenotype of the four strains

We tested the resistance status of the four *C. lectularius* strains to the pyrethroid insecticide deltamethrin. As expected, the two recently collected field strains showed higher deltamethrin resistance as compared to the two ancient lab strains (Figure 1.A).

German Lab and London Lab had the lowest LD₉₅ values, respectively 0.4 ng and 1.3 ng of deltamethrin per bed

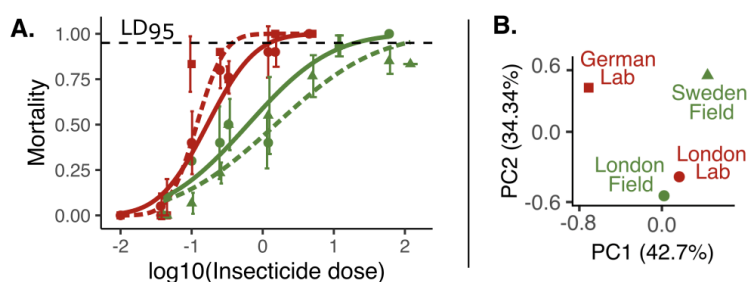


FIGURE 1 Analysis of the four bed bug strains used (German Lab: red square, London Lab: red round, London Field: green round, Sweden Field: green triangle). A. Dose-response curves to the insecticide deltamethrin exposure in bed bug populations German Lab (dashed red line), London Lab (red line), London Field (green line) and Sweden Field (dashed green line). Average mortality percentages are shown for each \log_{10} -transformed insecticide dose (in $\text{ng}/\mu\text{L}$) with standard deviation across replicates. Lines correspond to probit transformations of the data. B. Projection of four bed bug populations on the top two principal components using PCA on 8.03 million SNPs.

bug. On the contrary, London Field and Sweden Field had higher LD_{95} , respectively 21.0 ng and 69.2 ng (Supplementary Table 2). Resistance ratios to this dose (RR_{95}) between field and lab strains were all $>10 \text{ X}$ (Table 1), which is usually considered as high resistance ratio (Mazzarri and Georghiou, 1995).

Population comparison	Statistical pgof test	Standard error	Resistance ratio
London Lab - London Field	2.95 (*)	0.414	17 X
German Lab - London Field	4.34 (**)	0.411	60 X
London Lab - Sweden Field	1.48 (-)	1.17	55 X
German Lab - Sweden Field	1.96 (*)	1.17	199 X
German Lab - London Lab	6.81 (**)	0.082	4 X
London Field - Sweden Field	0.42 (-)	1.24	3 X

TABLE 1 Pearson's chi square goodness-of-fit (pgof) statistical test on insecticide deltamethrin LD_{95} for *Cimex lectularius* population comparison, with associated standard errors and resistance ratios. Significance levels for statistical tests are indicated, based on p-values (not significant (-), <0.05 (*), <0.001 (**)).

Comparisons revealed that all strains had significantly different phenotypes in terms of insecticide resistance, exception made of Sweden Field versus London Field and London Lab. Overall, these results confirmed the resistance status of the two field strains (R) compared to the two susceptible lab strains (S).

3.2 | Genome-wide genetic differentiation

8.03 million (M) SNPs, distributed over 1,077 scaffolds, were detected. The genetic diversity, measured as π , was very similar across the four populations (Table 2). After PCA, the four populations were projected onto two principal components, PC1 and PC2, explaining respectively 42.71% and 34.34% of variance (Figure 1.B). The four strains clustered according to their geographic origin, rather than their age or their deltamethrin resistance status. The historical and modern strains collected in London were noticed to be very close to each other on the 2D plot, whereas strains collected in different countries were distant (respectively Germany, Sweden and United Kingdom). This observation suggests that London Lab and London Field shared a common genetic background comparatively to the other strains.

Average fixation indices (F_{ST}) were then computed to provide a measure of pairwise population genetic differen-

tiation. As expected from the PCA analysis (Figure 1.B), the mean F_{ST} calculated between London Field and London Lab was the lowest ($F_{ST} = 0.018$, i.e. 1.8%), compared to the other comparisons which revealed much higher levels of differentiation (between 4.8% and 8.9%, Table 2). It therefore confirmed the weak genetic differentiation between the two strains from London. Both PCA and F_{ST} analysis thus indicate that the comparison between London Lab and London Field is particularly relevant to identify candidates potentially involved in insecticide resistance.

Strains	German Lab	London Lab	London Field	Sweden Field
German Lab	0.0048	0.059	0.065	0.089
London Lab	-	0.0045	0.018	0.061
London Field	-	-	0.0045	0.048
Sweden Field	-	-	-	0.0050

TABLE 2 Mean F_{ST} values for pairwise comparisons of *Cimex lectularius* strains, and within-strain nucleotide diversity (π , in bold).

3.3 | Overview of SNPs

Within the 2.92M SNPs that passed the QC filters, 576 outlier SNPs were identified (using top 5% $F_{ST} LL-LF$, BayPass metrics and $F_{alt} LF > F_{alt} LL$, see Methods for detail). Among them, 369 were detected within genes. With only 7 outliers in putative resistance genes, there was no significant overall enrichment. Only one non-synonymous outlier SNP was found out of 10,318 non-synonymous SNPs. This non-synonymous mutation was detected in a homeobox protein one cut (LOC106666929) located in a scaffold not included in any linkage groups and not particularly enriched for outlier SNPs (NW_019392707.1, Supplementary figure 3).

3.3.1 | Distribution of outlier SNPs

Among the 576 outlier SNPs, 261 were located within linkage groups (LGs). To identify LG of particular interest, we first tested whether some were enriched with outlier SNPs, by comparing the proportion of outlier SNPs for each LG to the average value computed for the entire genome.

LG 8 and 11 were significantly enriched with outlier SNPs (indicated by a star next to LG number in fig. 2). In LG 8, despite a global 2.18 X enrichment, no scaffold was significantly enriched for outlier SNPs, nor were any outliers detected in putative resistance genes.

Still, a large number of outliers were found on the second scaffold, namely NW_019393011.1, mainly inside two uncharacterized genes (106663904 and 106663882) and a small conductance calcium-activated potassium channel protein (106663893). This last gene was not included in our pre-defined resistance genes categories, but can still be involved in insecticide resistance through its potential role in synaptic functioning (Vergara et al., 1998).

By far, the highest enrichment was observed for LG 11 with a 4.86 X enrichment (p.value = 1.66e-34, Supplementary figure 1). Interestingly, most of the outlier SNPs clustered within three consecutive scaffolds within this LG (NW_019392721.1, NW_019392763.1 and NW_019942502.1.1). These three scaffolds were also significantly enriched for outlier SNPs (underlined in red in fig. 2). Within this approximately 6 Mb region, we found two outlier SNPs in the intronic regions of two putative insecticide resistance genes, both on scaffold NW_019392721.1: a choline O-acetyltransferase (106664013), an enzyme that catalyzes the biosynthesis of acetylcholine, and an acid sphingomyelinase-like phosphodiesterase 3a (SMPDL3a, 106664028), which may be involved in insecticide detox-

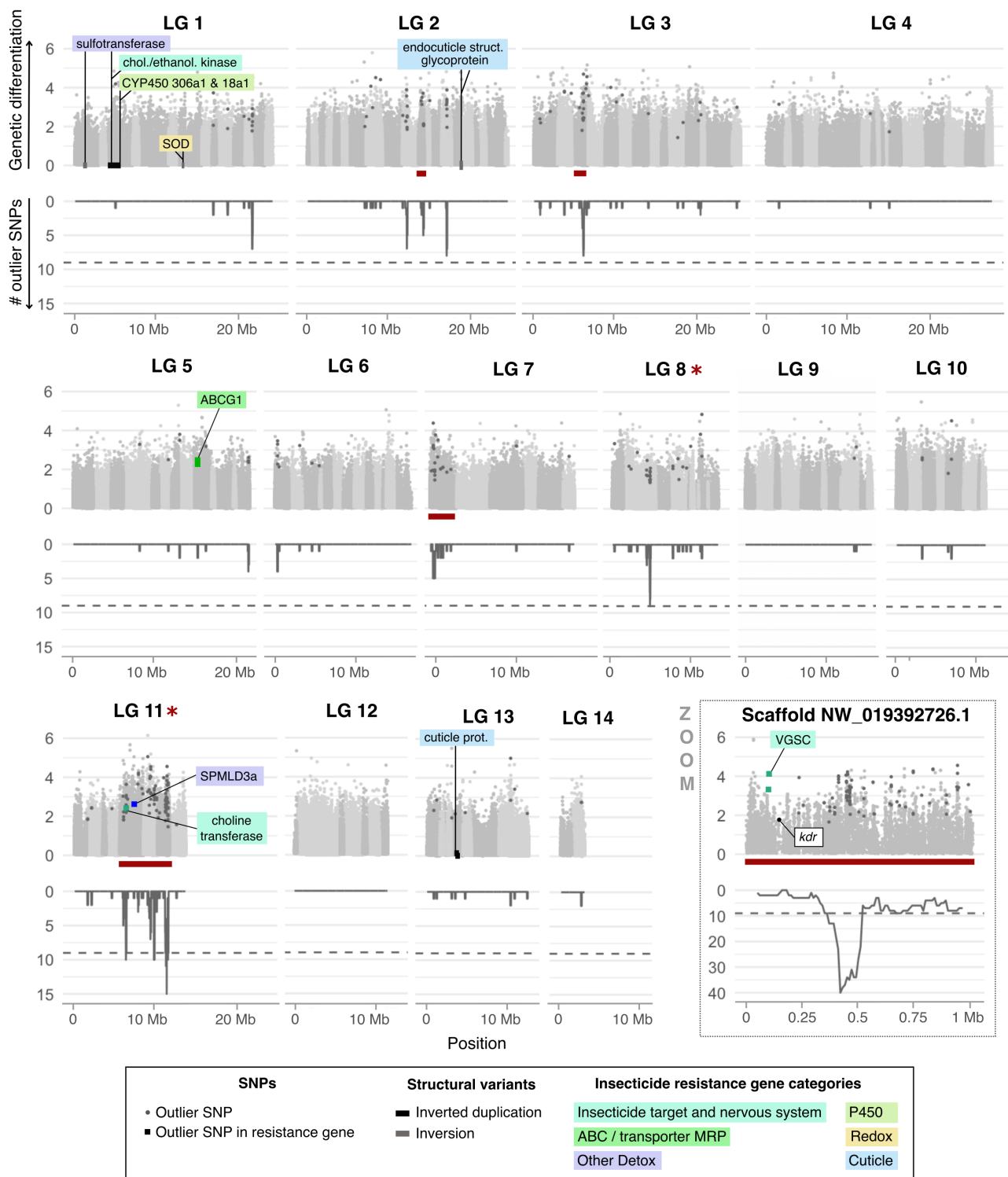


FIGURE 2 Outlier SNPs in *Cimex lectularius* linkage groups. Each plot represents a linkage group (LG), exception made of scaffold NW_019392726.1 (unplaced, zoomed on). Within a LG, scaffolds are delimited by different shades of grey. In the upper part of each plot, genetic differentiation is represented with empirical p-values, computed on F_{ST} index between London Lab and London Field populations for each SNP in $-\log_{10}$ scale. Outlier SNPs (see Materials & methods) are distinguished by a darker grey. Square-shaped points are the outlier SNPs that fall within candidate resistance genes, and are colored according to those genes' categories. Their protein products are labeled. Although not a outlier SNP, the location of the non-synonymous *kdr* L925I mutation was given as an indication. In the lower part, the number of outlier SNPs counted in each 100 kb sliding window with a step of 10 kb is represented. Dashed line stands for a threshold of 2 outlier SNPs, which delineates quantile 99%. The positions of structural variants (inverted duplications in black and inversions in grey) overlapping with resistance genes and with top 10% F_{ST} values are displayed on the x-axis ($y=0$). Annotations of overlapping resistance genes are labeled, and are coloured according to the resistance category. 88

ification (Zhang et al., 2012). However, many other candidate resistance genes were found inside this region, thus being further explored below (see Section 3.3.2).

Other scaffolds in LG detected as enriched were rather dispersed, without close proximity to putative resistance genes. Finally, two outlier SNPs were detected in an ABC transporter G, located on LG 5 which is not particularly enriched in outliers, neither at the scale of the whole LG nor at the scale of its scaffolds. ABC transporter G may be involved in the excretion of insecticides out of the cells (Dermauw and Van Leeuwen, 2014).

Eight scaffolds not included in the 14 linkage groups were significantly enriched for outlier SNPs (Supplementary figure 3). Two of these scaffolds carried outlier SNPs in putative resistance genes. A first outlier SNP was detected in an intronic region of a beta-14-galactosyltransferase 4-like on scaffold NW_019392678.1. Galactosyltransferase is a type of glycosyltransferase, an enzyme capable of contributing to xenobiotic detoxification through the conjugation of an electrophilic group to xenobiotics or to their metabolites (Nagare et al., 2021).

Two intronic outlier SNPs were also detected in the voltage-gated sodium channel gene (VGSC, 106667833), on scaffold NW_019392726.1, in close proximity with the *kdr* mutation. The *kdr* non-synonymous SNP of the VGSC L925I (leucine to isoleucine in position 925, i.e. position 148,668 on NW_019392726.1), was indeed further identified. Frequencies of the derived allele were as follows: 0% in London Lab, 11.1% in German Lab, 44.4% in London Field, and 50% in Sweden Field. This locus was part of the top 1.6% F_{ST} in the London Lab versus London Field comparison. The *kdr* mutation has been extensively studied in insects, including bed bugs, and is associated with pyrethroid resistance (see review Dang et al., 2017). Although our analysis did not consider *kdr* mutation as an outlier *per se*, *kdr* mutation still belongs to a peculiar 1 Mb genomic region (i.e. the scaffold size) enriched for outlier SNPs.

3.3.2 | Detection of a superlocus enriched in both outlier SNPs and resistance genes

Three consecutive scaffolds enriched for outliers were detected on the LG 11 (NW_019392721.1, NW_019392763.1 and NW_019942502.1, Figure 3). In the following, this 6 Mb genomic region is referred to as a "superlocus". Among the 138 genes located within this superlocus, nine were part of a resistance gene category, which represents a significant enrichment for resistance genes when compared to the entire genome (prop.test, p-value = 0.027). More precisely, two categories of resistance genes were significantly enriched: GSTs (corrected p-value = 3.0e-9, 3 genes in the superlocus out of 14 in the whole genome), and P450s (corrected p-value = 0.029, 3 genes in the superlocus out of 56 in the whole genome). The same trend was observed for the category 'Insecticide target and nervous system', although this was only marginally significant (corrected p-value = 0.063, 2 genes out of 31). This result suggests a clustering of genes potentially involved in insecticide resistance in this superlocus.

Indeed, the three P450s detected in the superlocus were located next to each other, all belonging to the sub-family CYP6B (106663981, 106663982 and 106663983). Members of the CYP6B P450 sub-family have been shown to metabolize a wide range of natural and synthetic xenobiotics in insects, including pyrethroids, as in *Helicoverpa armigera* (Shi et al., 2021). The same pattern was observed for the three GST detected, clustered on scaffold NW_019392763.1 (106666926, 106664023 and 106663963). GST have also been frequently associated with pyrethroid resistance, as for *Anopheles sinensis* (Tao et al., 2022).

Interestingly, despite none of these genes carried outlier SNP, they showed 23 non-synonymous SNPs with an alternative allele frequency higher in London Field than in London Lab (Supplementary table 3). Four of these non-synonymous SNPs were also part of the top 5% F_{ST} , with two for each P450s gene 106663983 and 106663982.

Because the abundance of non-synonymous mutations may be the consequence of adaptive evolution, we calculated the ratio $\frac{\pi_n}{\pi_s}$ for all the genes of the genome. The $\frac{\pi_n}{\pi_s}$ for the three CYP6B genes were in the top 5% of the distribution (quantile 95% = 0.88, mean = 0.19, versus 1.07, 1.25 and 1.50 respectively for genes 106663981,

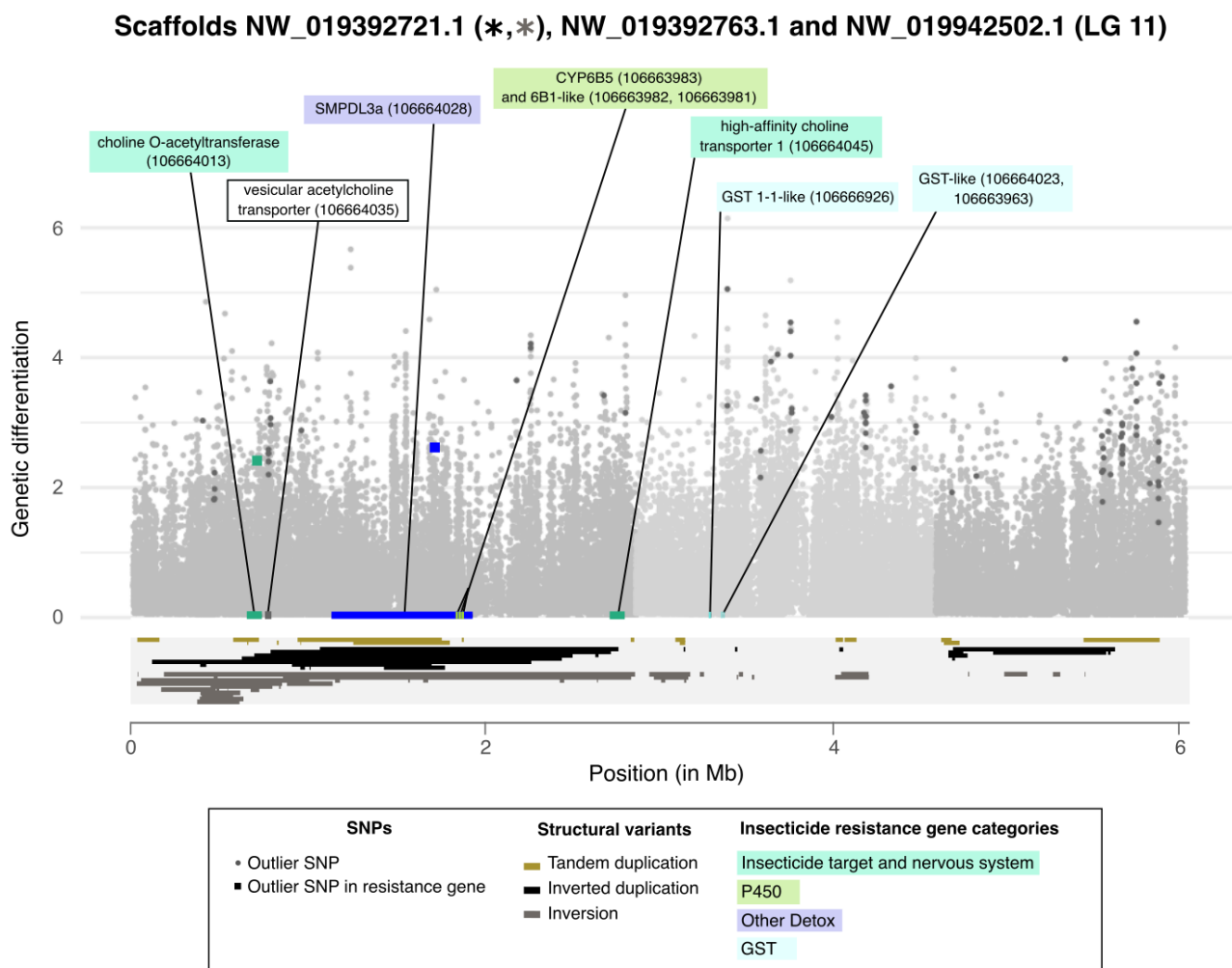


FIGURE 3 Architecture of *Cimex lectularius* superlocus on LG 11. Within this superlocus, scaffolds NW_019392721.1, NW_019392763.1 and NW_019942502.1.1 are delimited by different shades of light grey. Genetic differentiation is represented with empirical p-values, computed on F_{ST} index between London Lab and London Field populations for each SNP in $-\log_{10}$ scale, depending on their position within each scaffold. Outlier SNPs (see Materials & methods) are distinguished by a darker grey. Square-shaped points are the outlier SNPs that fall within resistance genes, and are colored according to those genes' categories. Resistance gene positions are indicated on the x-axis and colored accordingly to their categories. Protein product of resistance genes are labeled, along with the NCBI accession number of the genes between brackets. One vesicular acetylcholine transporter, carrying outlier SNPs and in close proximity to choline-O-acetyltransferase, was also labelled likewise. The positions of structural variants were represented inside a light grey box under the x-axis. Inverted duplications are in black, inversions in grey, and tandem duplications in gold. A star of the same color was added next to the scaffold name if it was enriched in a type of event.

106663982 and 106663983). This result supports that adaptive evolution has shaped the polymorphism for these P450s.

3.4 | Structural variants associated with resistance

Structural variants (SVs) may play an important role in insecticide resistance (Assogba et al., 2015). We detected 2,497 SVs with higher frequency in the resistant strain London Field compared to the susceptible London Lab, encompassing 1,338 inversions, 650 inverted duplications, and 509 tandem duplications.

Of the 2,497 SVs detected, 143 were overlapping with 130 putative resistance genes (out of 431 in the whole genome), which is more than what is expected by chance (p -value=3.53e-11). This indicates an enrichment of SVs in genes possibly related to insecticide resistance. Among SVs overlapping with candidate resistance genes, 80 genes were found within inversions, 50 in inverted duplications and 30 in tandem duplications.

Seven scaffolds were significantly enriched in SVs compared to what was expected based on their respective lengths, if one assumes that events were equally distributed among scaffolds: NW_019392673.1 (LG 1), NW_019392754.1 (LG 7), NW_019392785.1 (LG 10), NW_019392721.1 (LG 11), NW_019392782.1 (LG 12), NW_019392665.1 (LG 13) and NW_019392786.1 (unplaced) (details in Supplementary figure 2). Among them, the only scaffold that was significantly enriched both for outlier SNPs and SVs was the scaffold NW_019392721.1 (Supplementary figure 2), located within the superlocus of LG 11 and encoding the three *CYP6B* with unusual high $\frac{\pi_n}{\pi_s}$ ratios (Figure 3).

For each SV category (inversion, inverted duplication, tandem duplication), events falling above the 90% quantile of F_{ST} were retained, leading to 227 SVs overlapping with 281 different genes. Among them, ten SVs overlapped with 18 resistance genes. Seven cuticular genes were identified (in three inversions and one inverted duplication), showing an enrichment in this resistance category (p -value = 0.01). The largest event detected was a 1.5 Mb inverted duplication located on LG 1 (NW_019392679.1). This event overlapped with four putative resistance genes : a sulfotransferase, a choline kinase, and two P450s (Supplementary table 4). The five other events overlapped genes possibly involved in redox homeostasis (SOD and NADPH-dependent oxidoreductase), in metabolic resistance (UDPGT on scaffold NW_019392656.1), and in other detox (sulfotransferase).

4 | DISCUSSION

Insecticide resistance evolution seems to be the key factor in the re-emergence of the common bed bug *C. lectularius* (Davies et al., 2012). By using targeted sequencing (Zhu et al., 2010), transcriptomic analysis (Zhu et al., 2012; Mamidala et al., 2012; Koganemaru et al., 2013), or QTL mapping (Fountain et al., 2016), several loci have been associated with insecticide resistance (reviewed in Dang et al. 2017). In this paper, we provide a genome-wide analysis to identify underlying sequences and additional regions of the genome possibly involved in insecticide resistance. By comparing the whole genomes of two susceptible and two resistant lines using pool-seq data, we identified outlier SNPs and structural variants (inversion, tandem duplications, and inverted duplications) highly differentiated in the resistance phenotype, in close proximity with candidate resistance genes. In addition to providing novel insecticide resistance data, the genomic characterization of CimexStore strains will hopefully benefit future customer-researchers that regularly use these strains for their research.

4.1 | Relatively low pyrethroid resistance level

Using topical application assays, we quantified the deltamethrin resistance level of the four strains. As expected from the collection dates, the two ancient strains (LL and GL) were highly susceptible to deltamethrin. In most studies, the Harlan strain is used as a insecticide susceptible reference strain (Dang et al., 2017). The susceptibility of the LL and GL strains was comparable to that of the Harlan strain (0.34 ng and 0.15 ng per insect respectively for London Lab and German Lab, compared to 0.15 ng for the Harlan strain Cáceres et al. (2019)).

This high sensitivity is expected since these two "lab" strains were collected before the re-emergence of bed bugs, as was the Harlan strain. Conversely, the two relatively-recent strains (LF and SF), respectively collected in 2008 and 2015 were much more resistant to deltamethrin. However, their resistance ratios were not particularly high (17 to 199 fold based on LD₉₅), as compared to Cáceres et al. (2019), where field strains from Barracas and Pergamino were respectively 6,945 and 10,178 times higher than their susceptible reference. Resistance ratio should be compared with caution, since the bioassays or analytical methods can be widely different and have an impact. However, very high RR phenomena have frequently been observed for recently collected field strains (see review Dang et al. 2017), but often decline after rearing in the laboratory without insecticide exposure (Polanco et al., 2011). The existence of an evolutionary cost of resistance in the absence of insecticide exposure has been frequently described in insects, and may explain the decrease in frequency of resistance alleles since the establishment of the strains (Kliot and Ghanim, 2012). This phenomenon may explain the relatively low level of resistance for the two CimexStore "field" strains purchased, for which no information was provided regarding insecticide selection pressure during the breeding in laboratory conditions.

4.2 | Non-synonymous outlier SNP in a transcription factor

Adaptation of organisms to their environment may rely on protein evolution thanks to non-synonymous mutations, and/or thanks to expression changes involving mutations in regulatory sequences (Wang et al., 2005). Among the 576 outlier SNPs, only one was associated with an amino-acid change. The encoded protein, annotated as a homeobox protein onecut (106666929 on Supplementary fig. 3), is composed of a onecut domain and one homeodomain which together confer DNA binding properties (Bürglin and Affolter, 2016). Homeobox proteins onecut are transcription factors involved in many pathways including nervous system differentiation (Nguyen et al., 2000). Several transcription factors have been associated with resistance to insecticide, as in *Plutella xylostella* (Zhao et al., 2018). Yet, how this particular transcription factor could be involved in insecticide resistance remains unclear.

4.3 | A superlocus involved in insecticide resistance

The most striking result obtained from this genome-wide analysis is the identification of a 6 Mb genomic region, showing an unusual high genetic differentiation and associated with the resistance phenotype of the four strains (BayPass approach). Several phenomena may explain such local genetic differentiation. Large genetic differentiation islands may evolve because of a local reduction in recombination rate due to the effect of linked selection (Burri et al., 2015). However, if this factor was at play, then we would expect a higher genetic differentiation in this genomic region for all comparisons, including between the two "field" strains or between the two "lab" strains. This was not the case, since this genomic region was particularly differentiated in the London "Field" versus London "Lab" comparison, but not in other comparisons (Suppl. fig. 5). High local genetic differentiation may thus be the consequence of differential selection. In this case, the selective factor is most likely the exposition to insecticides that recent field

strains experienced, as compared to older "lab" ones. This region was part of LG 11 and was composed of three scaffolds. Those scaffolds were enriched with outlier SNPs, and one of them was also enriched with structural variants.

The detection of this superlocus is consistent with a previous study from [Fountain et al. \(2016\)](#) who constructed a linkage map and performed a QTL mapping for pyrethroid resistance in *C. lectularius*. Their QTL analysis based on RAD tag sequencing identified a single peak associated with resistance. Due to the relatively low number of markers ($n=334$), to the relative small number of recombinants screened ($n=71$) and to the small number of generations allowing recombination (F2), the resolution achieved was relatively low and led to the identification of a large 10.8 Mb region. This region was located on their LG 12 and contains the highly expected insecticide resistance candidate VGSC, as well as many other genes, including putative resistance genes (P450 and GSTs). This large locus explained most of the variance in insecticide resistance in their experimental display. Comparing the two LGs belonging to the two linkage maps indicated that the majority of the genes found on our-LG 11 was shared with Fountain-LG 12 (Supplementary figure 4). However, the scaffold carrying the VGSC gene which was part of Fountain-LG 12 was not included in our-LG 11. Instead, the scaffold was left unplaced in any LG. The explanation for this discrepancy is simple. In the previous genome assembly Clec_1.0, the VGSC gene belonged to a 21 Mb scaffold (NW_014465016.1) containing 6 RAD tags (See Supplementary table S2 of [Fountain et al. 2016](#)), that allowed [Fountain et al. \(2016\)](#) to incorporate VGSC gene into their LG 12. This scaffold was then split in smaller pieces in Clec_2.1 assembly, leaving VGSC in a 1 Mb-scaffold not carrying any RAD marker anymore (NW_019392726.1). This difference in assemblies thus prevented its integration into the new genetic map.

The second version of the assembly was obtained by using Redundans ([Thomas et al., 2020](#)) on top of the first assembly. Redundans ([Pryszcz and Gabaldón, 2016](#)) is designed for solving problems linked to the presence of heterozygosity in genome sequencing projects. Based on read depth and paired-end information, it identifies putative "heterozygous contigs" which correspond to different allelic versions of a single locus. They are then merged in order to produce as far as possible a homozygous version of the genome. Although the reason explaining the split of the initial 21 Mb scaffold (NW_014465016.1) into smaller pieces is unclear, we may hypothesize that some of the contigs that composed this scaffold were detected as "heterozygous contigs" ultimately leading to the observed split. Future investigations should certainly clarify this point and more generally clarify the structure of the haplotypes in this genomic region.

We note however that in the 21 Mb scaffold NW_014465016.1 of Fountain-LG 12, VGSC was flanked by two genes that are both included in our LG 11 (a carboxyl terminal-hydrolase on scaffold NW_019942502.1, and a glutathione S-transferase on scaffold NW_019392763.1). The position of the scaffold containing VGSC gene (NW_019392726.1) may thus possibly be between the two scaffolds previously mentioned, in the middle of the superlocus. If this is the case, it would further reinforce the strong signal of genetic differentiation in this genomic area.

Independently of its exact genomic localization, we expected the mutant sites of the VGSC gene (with their so-called knockdown resistance *-kdr-* alleles) to be part of our outlier list. This was expected since *kdr* mutations are responsible for or are associated with resistance to pyrethroids in various arthropod pests ([Dong et al., 2014](#)), including bed bugs ([Dang et al., 2017](#)). Additionally, CimexStore website indicates the presence of the common *kdr* mutant L925I in the two field strains. Accordingly, we detected this mutation in our dataset ([Dang et al., 2017](#)). However, this particular site was not considered as an outlier SNP in our analysis. This site fulfilled only two out of three criteria used since (i) the F_{ST} associated with this site was part of the top 5% in the London strains comparison and (iii) the *kdr* mutant was derived compared to the reference genome, but (ii) it was filtered out by the BayPass analysis, indicating an overall weak association with resistance phenotype in the four strains. Several scenarios may explain this unexpected result. First of all, we can hypothesize that this mutation is not as crucial for insecticide resistance in *C. lectularius* as previously thought. Under this scenario, other mechanisms such as detoxification, cuticle thickening

or behavioural avoidance would be more central in bed bug insecticide resistance. Importantly however, since these strains have been reared in the lab since years by the private company CimexStore, we do not exactly know under which conditions they have been maintained. One crucial point to keep in mind is that *kdr* mutants are highly counter-selected in the absence of insecticide (Freeman et al., 2021; Wang et al., 2021) and may thus decrease in frequency if not regularly exposed to insecticides. We can therefore question whether the frequency estimated here reflects the true initial frequency of that mutation in the different strains. If present, this effect may thus reduce the genetic differentiation observed on this locus (and surrounding region) between field "resistant" and ancient lab "susceptible" strains. Despite the fact that our analysis did not consider *kdr* mutant as part of the outlier list, we found two outlier SNPs in an intron of the VGSC gene. Additionally, VGSC was located next to a region containing the highest density of outlier SNPs, with up to 40 outlier SNPs per 100 kb window (Figure 2). Altogether, these last observations indicate that VGSC belongs, at least, to a peculiar 1 Mb region in relation to insecticide resistance.

4.4 | A peculiar cytochrome P450 cluster

A cluster of three P450 genes (two *CYP6B1*-like and one *CYP6B5*-like) was detected in scaffold NW_019392721.1, in the middle of the LG11 superlocus. No outlier SNP were detected in any of these loci. However, four SNPs were part of the top 5% F_{ST} distribution for the London strains comparison. In addition, this cluster was characterized by an unexpectedly high level of non-synonymous mutations ($n=19$). Accordingly, the $\frac{\pi_n}{\pi_s}$ ratio for all three genes was in the top 5% or the genome-wide distribution, suggesting that selection may have acted on these loci in a recent past. Similar increased rate of evolution has been observed for *CYP6B56* of the navel orangeworm *Amyelois transitella* (Calla et al., 2021). In this species, *CYP6B56* is likely involved in the pyrethroid bifenthrin resistance, since elevated expressions of *CYP6B56* were observed after insecticide exposure in both resistant and susceptible populations (Calla et al., 2021).

4.5 | Structural variants might contribute to shape resistance phenotype

Structural variants may contribute to adaptation, notably through inversions or duplications. These structural changes can affect the coding sequences as well as the expression levels of the genes, thus playing a crucial role in evolution and adaptation (North et al., 2020). CNVs in particular may impact the level of expression of the genes through a dosage mechanism (Kondrashov, 2012). In addition, CNVs can allow alternative variants of a gene to appear in tandem, through gene duplication. This may be selected for, especially when the allele encoding resistance also carry a fitness cost in the absence of insecticide. In that case, the second nearby copy may compensate for this cost and "heterogeneous" tandem duplications may be selected for (Labbé et al., 2007; Assogba et al., 2015). Despite their potential importance, SVs have received less attention than SNPs, likely because they are harder to identify. Population-level genome-wide analyses of SVs are thus rare, and the extent of their impact on evolution remain overlooked (Lucas et al., 2019). However, some SVs have been associated with insecticide resistance in insects, such as in *Drosophila* (Schmidt et al., 2010), in several mosquito species (Faucon et al., 2015; Weetman et al., 2018; Lucas et al., 2019; Cattel et al., 2020, 2021), or in the fall armyworm with the identification of a CYP cluster (Gimenez et al., 2020). To the best of our knowledge, the present study is the first to investigate SVs in an agnostic genome-wide fashion in bed bugs. We found that SVs are enriched in putative resistance genes. A similar test of enrichment focusing on copy number variants (CNVs) has been conducted on the mosquito *Anopheles gambiae*, using the Ag1000G Phase 2 dataset (1142 whole genome sequences of wild-caught mosquitoes collected from 16 populations) (Lucas et al., 2019). The authors found an enrichment for CNVs in gene-containing regions. More specifically, they found an enrichment in

putative metabolic resistance gene families. Similar result has been observed in *Drosophila* (Schrider et al., 2016) and *Aedes aegypti* (Faucon et al., 2015; Cattel et al., 2020, 2021). Additionally, several independent amplification events were frequently found as covering the same genes, possibly segregating within the same populations (Lucas et al., 2019). We reached the same conclusion, with some genomic locations highly enriched with overlapping events. This was in particular the case for scaffold NW_019392721.1 in LG 11 which contained several candidate genes, including a cluster of three P450, and that was enriched both with outlier SNPs and SVs. Furthermore, we found a few set of events overlapping resistance genes with high F_{ST} values (top 10%, max : top 1%, see Supp. table 4). Among them, a significant enrichment for cuticle genes was detected. Although the functional importance of these SVs in the resistance phenotype remains to be further investigated, these results support their potential role in adaptation to insecticides in bed bugs.

4.6 | A supergene involved in insecticide resistance?

A significant enrichment for genes potentially involved in insecticide resistance was observed in the 6 Mb superlocus. Colocalization of genes involved in detoxification in clusters is frequently observed in insects, including the mosquito *Anopheles gambiae* (Ranson et al., 2002). This clustering is probably the consequence of local duplications or of translocation of genes into the cluster. The discovery of a single large (at least 6 Mb long) locus probably involved in insecticide resistance is uncommon in the literature. In most cases reported to date, the loci identified by population genomics analysis are relatively small (<200 kb in Weedall et al. 2019) and numerous (Pélissié et al., 2022), and no evidence of genomic clustering among those loci is provided. A few exceptions exist however. In the navel orangeworm (*Amyelois transitella*), an exceptionally large region with low polymorphism, extending to up to 1.3 Mb was detected in a DDT and bifenthrin resistant population (Calla et al., 2021). This region contained several genes likely involved in insecticide resistance, such as VGSC and a cytochrome P450 gene cluster. The genomic co-localization of genes that are selected for in the same environment (i.e. under insecticide exposition) may theoretically lead to the emergence of so-called "supergenes" if a subsequent reduction of recombination between beneficial alleles occurs (Schwander et al., 2014). In a spatially or temporally heterogeneous environment (in that case treated versus untreated households), co-transmission of alleles may be selected for. Several fascinating examples of such evolution have been reported, from the classical case of Butterfly mimicry in the genus *Heliconius* (Le Poul et al., 2014) to the evolution of prairie or forest ecotypes of deer mice *Peromyscus maniculatus* (Hager et al., 2022). Although our analysis does not provide any evidence of a reduction in recombination in this superlocus (Supp. fig. 6), it is intriguing to observe such enrichment for putative resistance genes in a highly differentiated and large genomic island. Additionally, our dataset suggests an enrichment for structural variants, especially inversions, in one of the three scaffolds composing this superlocus (Fig. 3 & Supplementary fig. 2). Since inversions may lead to reduced recombination and are known to be instrumental in the emergence of supergenes (Schwander et al., 2014), it is tempting to speculate that a supergene combining several beneficial alleles in relation to insecticide exposure segregate in *C. lectularius* populations. This hypothesis should be further tested in the future.

5 | LIMITS OF THE PRESENT STUDY AND PERSPECTIVE

In this study, we focused our attention on selected variants that had particularly high frequency in both LF and SF strains compared to LL and GL strains, while taking into account their common history. In species with large (recent) effective population size, the very same mutant may be repeatedly selected in different environments under similar

selective pressure, through convergent evolution (Karasov et al., 2010). However, species with lower effective present-day population size are likely limited by the mutation rate, and such convergence is unlikely. It is unclear whether bed bug populations are limited by mutations or not. If they are, then we do not expect similar mutants to be selected in different environments, unless migration from one population brings the favorable allele in the second population. Additionally, since multiple pathways to pyrethroid resistance may evolve in different populations (Djouaka et al., 2008), we acknowledge that this analysis may miss insecticide resistance mutations that would be specific to one or another resistant population.

Another limit of our study relies on the fact that we considered "derived" alleles as opposed to "ancestral", carried by the Harlan strain from which the reference genome was sequenced. Harlan strain was collected in Fort Dix (New Jersey, USA), back in 1973. We thus tried to identify mutations involved in so-called hard selective sweeps. Many examples of hard sweeps have been described as being involved in insecticide resistance (Hawkins et al., 2019; Cohen et al., 1985; Calla et al., 2021). However, adaptation may also occur through soft sweeps involving standing genetic variation (Hawkins et al., 2019). In that case, the mutants segregated during some time in the population before the selective pressure started. If adaptation to pyrethroids involves standing genetic variation, then our pipeline may have filtered out some valid candidates. However, standing alleles that later increase in frequency due to positive selection are usually at low initial frequency (before the environmental shift). In the three-spined stickleback, where adaptation from standing genetic variation has been extensively studied (DeFaveri and Merilä, 2014; Heckwolf et al., 2020), marine populations adapted repeatedly through standing genetic variation to freshwater habitat. Most adaptive haplotypes in fresh water environment can be found at frequency < 10% in marine environment (Bassham et al., 2018). Since the reference genome has a small probability of having a low frequency allele, then our pipeline still has a high probability of considering it as "derived". In addition, even in case the Harlan strain carried a later on beneficial allele, our pipeline should at least identify linked "derived" (possibly neutral) variants, initially associated with the beneficial mutant through linkage disequilibrium. In conclusion, we acknowledge that our pipeline may have missed some adaptive alleles from standing genetic variation, although the effect may be limited.

Whether they arose through soft or hard selective sweeps, and whether they involve protein changes or not, it is crucial to identify insecticide resistance mutations in bed bug populations. As our study was restricted to the analysis of intra-scaffold SVs, thus preventing the detection of events spanning the entire superlocus, future efforts should be made to clarify the genomic structure and polymorphism of LG 11. This can be achieved using long read sequencing and/or Hi-C data. This could clarify the structure of the haplotypes and may help to identify the genetic determinants underlying the resistance evolution. From an applied perspective, knowing the genetic basis of such mutations can be used to design diagnostic tests, and thus to track insecticide resistance in the field (Van Leeuwen et al., 2020).

acknowledgements

This work was performed using the computing facilities of the CC LBBE/PRABI.

conflict of interest

The authors declare no conflicts of interest.

data availability statement

The data that support the findings of this study are openly available on SRA under BioProject PRJNA826750.

references

- Adelman, Z. N., Kilcullen, K. A., Koganemaru, R., Anderson, M. A., Anderson, T. D. and Miller, D. M. (2011) Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a single population. *PLoS ONE*, **6**, 1–9.
- Akhoundi, M., Kengne, P., Cannet, A., Brengues, C., Berenger, J. M., Izri, A., Marty, P., Simard, F., Fontenille, D. and Delaunay, P. (2015) Spatial genetic structure and restricted gene flow in bed bugs (*Cimex lectularius*) populations in France. *Infection, Genetics and Evolution*, **34**, 236–243. URL: <http://dx.doi.org/10.1016/j.meegid.2015.06.028>.
- Alexander, J. O. (1984) Infestation by Hemiptera. In *Arthropods and Human Skin*, 57–74. London: Springer London. URL: https://doi.org/10.1007/978-1-4471-1356-0_6.
- Amezian, D., Nauen, R. and Le Goff, G. (2021) Transcriptional regulation of xenobiotic detoxification genes in insects - An overview. *Pesticide Biochemistry and Physiology*, **174**, 104822. URL: <https://www.sciencedirect.com/science/article/pii/S0048357521000535>.
- Assogba, B. S., Djogbénu, L. S., Milesi, P., Berthomieu, A., Perez, J., Ayala, D., Chandre, F., Makoutodé, M., Labbé, P. and Weill, M. (2015) An ace-1 gene duplication resorbs the fitness cost associated with resistance in *Anopheles gambiae*, the main malaria mosquito. *Scientific Reports*, **5**, 19–21.
- Ayala, D., Ullastres, A. and González, J. (2014) Adaptation through chromosomal inversions in *Anopheles*. *Frontiers in Genetics*, **5**, 1–10.
- Balabanidou, V., Grigoraki, L. and Vontas, J. (2018) Insect cuticle: a critical determinant of insecticide resistance. *Current Opinion in Insect Science*, **27**, 68–74. URL: <https://www.sciencedirect.com/science/article/pii/S2214574517301190>.
- Balvín, O. and Booth, W. (2018) Distribution and frequency of pyrethroid resistance-associated mutations in host lineages of the bed bug (Hemiptera: Cimicidae) across Europe. *Journal of Medical Entomology*, **55**, 923–928.
- Bass, C. and Field, L. M. (2011) Gene amplification and insecticide resistance. *Pest Management Science*, **67**, 886–890.
- Bassham, S., Catchen, J., Lescak, E., von Hippel, F. A. and Cresko, W. A. (2018) Repeated Selection of Alternatively Adapted Haplotypes Creates Sweeping Genomic Remodeling in Stickleback. *Genetics*, **209**, 921–939. URL: <https://doi.org/10.1534/genetics.117.300610>.
- Bolger, A. M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
- Brooke, B. D., Hunt, R. H., Chandre, F., Carnevale, P. and Coetzee, M. (2002) Stable Chromosomal Inversion Polymorphisms and Insecticide Resistance in the Malaria Vector Mosquito *Anopheles gambiae* (Diptera: Culicidae). *Journal of Medical Entomology*, **39**, 568–573. URL: <https://doi.org/10.1603/0022-2585-39.4.568>.
- Bürglin, T. R. and Affolter, M. (2016) Homeodomain proteins: an update. *Chromosoma*, **125**, 497–521. URL: <https://doi.org/10.1007/s00412-015-0543-8>.
- Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S. A., Sætre, G. P., Török, J. and Ellegren, H. (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, **25**, 1656–1665.
- Cáceres, M., Santo-Orihuela, P. L. and Vassena, C. V. (2019) Evaluation of Resistance to Different Insecticides and Metabolic Detoxification Mechanism by Use of Synergist in the Common Bed Bug (Heteroptera: Cimicidae). *Journal of medical entomology*, **56**, 1324–1330.
- Calla, B., Demkovich, M., Siegel, J. P., Viana, J. P. G., Walden, K. K. O., Robertson, H. M. and Berenbaum, M. R. (2021) Selective Sweeps in a Nutshell: The Genomic Footprint of Rapid Insecticide Resistance Evolution in the Almond Agroecosystem. *Genome Biology and Evolution*, **13**, evaa234. URL: <https://doi.org/10.1093/gbe/evaa234>.

- Cattel, J., Faucon, F., Le Péron, B., Sherpa, S., Monchal, M., Grillet, L., Gaude, T., Laporte, F., Dusfour, I., Reynaud, S. and David, J. P. (2020) Combining genetic crosses and pool targeted DNA-seq for untangling genomic variations associated with resistance to multiple insecticides in the mosquito *Aedes aegypti*. *Evolutionary Applications*, **13**, 303–317.
- Cattel, J., Haberhorn, C., Laporte, F., Gaude, T., Cumer, T., Renaud, J., Sutherland, I. W., Hertz, J. C., Bonneville, J. M., Arnaud, V., Fustec, B., Boyer, S., Marcombe, S. and David, J. P. (2021) A genomic amplification affecting a carboxylesterase gene cluster confers organophosphate resistance in the mosquito *Aedes aegypti*: From genomic characterization to high-throughput field detection. *Evolutionary Applications*, **14**, 1009–1022.
- Cohen, J., Newman, C. and Briand, F. (1985) A stochastic theory of community food webs II. Individual webs.
- Dang, K., Doggett, S. L., Veera Singham, G. and Lee, C. Y. (2017) Insecticide resistance and resistance mechanisms in bed bugs, *Cimex* spp. (Hemiptera: Cimicidae). *Parasites and Vectors*, **10**, 1–31.
- Dang, K., Toi, C. S., Lilly, D. G., Bu, W. and Doggett, S. L. (2015) Detection of knockdown resistance mutations in the common bed bug, *Cimex lectularius* (Hemiptera: Cimicidae), in Australia. *Pest Management Science*, **71**, 914–922.
- Dang, K. A. I., Toi, C. S., Lilly, D. G., Lee, C.-y., Naylor, R., Tawatsin, A., Thavara, U., Bu, W. and Doggett, S. L. (2014) Detection of knockdown resistance (kdr) in *Cimex lectularius* and *Cimex hemipterus* (Hemiptera: Cimicidae). *The Eighth International Conference on Urban Pests*, **71**, 914–922.
- Davies, T. G., Field, L. M. and Williamson, M. S. (2012) The re-emergence of the bed bug as a nuisance pest: Implications of resistance to the pyrethroid insecticides. *Medical and Veterinary Entomology*, **26**, 241–254.
- DeFaveri, J. and Merilä, J. (2014) Local adaptation to salinity in the three-spined stickleback? *Journal of Evolutionary Biology*, **27**, 290–302. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jeb.12289>.
- Dermauw, W. and Van Leeuwen, T. (2014) The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. *Insect Biochemistry and Molecular Biology*, **45**, 89–110.
- Djouaka, R. F., Bakare, A. A., Coulibaly, O. N., Akogbeto, M. C., Ranson, H., Hemingway, J. and Strode, C. (2008) Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of *Anopheles gambiae* s.s. from Southern Benin and Nigeria. *BMC Genomics*, **9**, 538. URL: <https://doi.org/10.1186/1471-2164-9-538>.
- Doggett, S. L., Geary, M. J. and Russell, R. C. (2004) The resurgence of bed bugs in Australia: with notes on their ecology and control. *Environmental health*, **4**, 30–38.
- Dong, K., Du, Y., Rinkevich, F., Nomura, Y., Xu, P., Wang, L., Silver, K. and Zhorov, B. S. (2014) Molecular biology of insect sodium channels and pyrethroid resistance. *Insect Biochemistry and Molecular Biology*, **50**, 1–17. URL: <https://www.sciencedirect.com/science/article/pii/S0965174814000551>.
- Durand, R., Cannet, A., Berdjane, Z., Bruel, C., Haouchine, D., Delaunay, P. and Izri, A. (2012) Infestation by pyrethroids resistant bed bugs in the suburb of Paris, France. *Parasite (Paris, France)*, **19**, 381–387.
- Faucon, F., Dusfour, I., Gaude, T., Navratil, V., Boyer, F., Chandre, F., Sirisopa, P., Thanispong, K., Juntarajumong, W., Poupardin, R., Chareonviriyaphap, T., Girod, R., Corbel, V., Reynaud, S. and David, J.-p. (2015) Unravelling genomic changes associated with insecticide resistance in the dengue mosquito *Aedes aegypti* by deep targeted sequencing. *Genome Research*, 1–13.
- Feroz, M. (1969) Toxicological and genetical studies of organophosphorus-resistance in *Cimex lectularius* L. *Bulletin of Entomological Research*, **59**, 377–382.
- Feyereisen, R. (1995) Molecular biology of insecticide resistance. *Toxicology Letters*, **82-83**, 83–90. URL: <https://www.sciencedirect.com/science/article/pii/0378427495034706>.
- Fountain, T., Ravinet, M., Naylor, R., Reinhardt, K. and Butlin, R. K. (2016) A linkage map and QTL analysis for pyrethroid resistance in the bed bug *cimex lectularius*. *G3: Genes, Genomes, Genetics*, **6**, 4059–4066.

- Freeman, J. C., San Miguel, K. and Scott, J. G. (2021) All resistance alleles are not equal: the high fitness cost of super-kdr in the absence of insecticide. *Pest Management Science*, **77**, 3693–3697. URL: <https://doi.org/10.1002/ps.6115>.
- Gimenez, S., Abdelgaffar, H., Goff, G. L., Hilliou, F., Blanco, C. A., Hänniger, S., Bretaudeau, A., Legeai, F., Nègre, N., Jurat-Fuentes, J. L., d'Alençon, E. and Nam, K. (2020) Adaptation by copy number variation increases insecticide resistance in the fall armyworm. *Communications Biology*, **3**, 664. URL: <https://doi.org/10.1038/s42003-020-01382-6>.
- Goddard, J. and Deshazo, R. (2009) Bed Bugs (*Cimex lectularius*) and Clinical Consequences of Their Bites. *JAMA : the journal of the American Medical Association*, **301**, 1358–1366.
- Hager, E. R., Harringmeyer, O. S., Wooldridge, T. B., Theingi, S., Gable, J. T., McFadden, S., Neugeboren, B., Turner, K. M., Jensen, J. D. and Hoekstra, H. E. (2022) A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science*, **377**, 399–405. URL: <https://doi.org/10.1126/science.abg0718>.
- Hawkins, N. J., Bass, C., Dixon, A. and Neve, P. (2019) The evolutionary origins of pesticide resistance. *Biological Reviews*, **94**, 135–155.
- Heckwolf, M. J., Meyer, B. S., Häslér, R., Höppner, M. P., Eizaguirre, C. and Reusch, T. B. (2020) Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. *Science Advances*, **6**.
- Hivert, V., Leblois, R., Petit, E. J., Gautier, M. and Vitalis, R. (2018) Measuring genetic differentiation from pool-seq data. *Genetics*, **210**, 315–330.
- Hlina, B. L., Birceanu, O., Robinson, C. S., Dhiyebi, H. and Wilkie, M. P. (2021) The relationship between thermal physiology and lampricide sensitivity in larval sea lamprey (*Petromyzon marinus*). *Journal of Great Lakes Research*, **47**, S272–S284. URL: <https://doi.org/10.1016/j.jglr.2021.10.002>.
- Karasov, T., Messer, P. W. and Petrov, D. A. (2010) Evidence that Adaptation in *Drosophila* Is Not Limited by Mutation at Single Sites. *PLOS Genetics*, **6**, 1–10. URL: <https://doi.org/10.1371/journal.pgen.1000924>.
- Kass, R. E. and Raftery, A. E. (1995) Bayes Factors. *Journal of the American Statistical Association*, **90**, 773–795.
- Kliot, A. and Ghanim, M. (2012) Fitness costs associated with insecticide resistance. *Pest Management Science*, **68**, 1431–1437.
- Kofler, R., Pandey, R. V. and Schlötterer, C. (2011) PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, **27**, 3435–3436.
- Koganemaru, R., Miller, D. M. and Adelman, Z. N. (2013) Robust cuticular penetration resistance in the common bed bug (*cimex lectularius* L.) correlates with increased steady-state transcript levels of CPR-type cuticle protein genes. *Pesticide Biochemistry and Physiology*, **106**, 190–197. URL: <http://dx.doi.org/10.1016/j.pestbp.2013.01.001>.
- Kondrashov, F. A. (2012) Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 5048–5057. URL: <http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.1108>.
- Labbé, P., Berthomieu, A., Berticat, C., Alout, H., Raymond, M., Lenormand, T. and Weill, M. (2007) Independent duplications of the acetylcholinesterase gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Molecular Biology and Evolution*, **24**, 1056–1067.
- Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T. and Carey, V. J. (2013) Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, **9**, 1–10.
- Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V. and Joron, M. (2014) Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature Communications*, **5**, 5644. URL: <https://doi.org/10.1038/ncomms6644>.
- Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

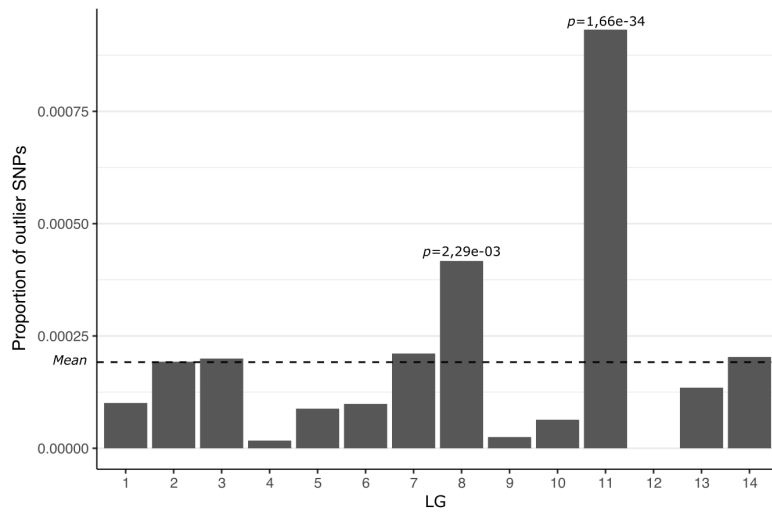
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- Lilly, D. G., Dang, K., Webb, C. E. and Doggett, S. L. (2016a) Evidence for metabolic pyrethroid resistance in the common bed bug (Hemiptera: Cimicidae). *Journal of Economic Entomology*, **109**, 1364–1368.
- (2018) Are Australian field-collected strains of *Cimex lectularius* and *Cimex hemipterus* (Hemiptera: Cimicidae) resistant to deltamethrin and imidacloprid as revealed by topical assay? *Austral Entomology*, **57**, 77–84.
- Lilly, D. G., Latham, S. L., Webb, C. E. and Doggett, S. L. (2016b) Cuticle thickening in a pyrethroid-resistant strain of the common bed bug, *cimex lectularius* L. (hemiptera: Cimicidae). *PLoS ONE*, **11**, 6–16.
- Lockwood, J. A., Sparks, T. C. and Story, R. N. (1984) Evolution of Insect Resistance to Insecticides: A Reevaluation of the Roles of Physiology and Behavior. *Bulletin of the Entomological Society of America*, **30**, 41–51. URL: <https://doi.org/10.1093/besa/30.4.41>.
- Lucas, E. R., Miles, A., Harding, N. J., Clarkson, C. S., Lawniczak, M. K., Kwiatkowski, D. P., Weetman, D. and Donnelly, M. J. (2019) Whole-genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes. *Genome Research*, **29**, 1250–1261.
- Mamidala, P., Jones, S. C. and Mittapalli, O. (2011) Metabolic resistance in bed bugs. *Insects*, **2**, 36–48.
- Mamidala, P., Wijeratne, A. J., Wijeratne, S., Kornacker, K., Sudhamalla, B., Rivera-Vega, L. J., Hoelmer, A., Meulia, T., Jones, S. C. and Mittapalli, O. (2012) RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug. *BMC Genomics*, **13**.
- Mazzarri, M. B. and Georghiou, G. P. (1995) Characterization of resistance to organophosphate, carbamate, and pyrethroid insecticides in field populations of *Aedes aegypti* from Venezuela. *Journal of the American Mosquito Control Association*, **11**, 315–322.
- Mérot, C., Oomen, R. A., Tigano, A. and Wellenreuther, M. (2020) A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. *Trends in Ecology & Evolution*, **35**, 561–572. URL: <https://doi.org/10.1016/j.tree.2020.03.002>.
- Nagare, M., Ayachit, M., Agnihotri, A., Schwab, W. and Joshi, R. (2021) Glycosyltransferases: the multifaceted enzymatic regulator in insects. *Insect Molecular Biology*, **30**, 123–137.
- Nguyen, D. N. T., Rohrbaugh, M. and Lai, Z.-C. (2000) The *Drosophila* homolog of *Onecut* homeodomain proteins is a neural-specific transcriptional activator with a potential role in regulating neural differentiation. *Mechanisms of Development*, **97**, 57–72. URL: <https://www.sciencedirect.com/science/article/pii/S0925477300004317>.
- North, H. L., Caminade, P., Severac, D., Belkhir, K. and Smadja, C. M. (2020) The role of copy-number variation in the reinforcement of sexual isolation between the two European subspecies of the house mouse: CNV and reinforcement in the house mouse. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **375**.
- Obenchain, V., Lawrence, M., Carey, V., Gogarten, S., Shannon, P. and Morgan, M. (2014) VariantAnnotation : a Bioconductor package for exploration and annotation of genetic variants . *Bioinformatics*, **30**, 2076–2078. URL: <https://doi.org/10.1093/bioinformatics/btu168>.
- Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., Diepenbrock, L. M., Kenis, M., Zhang, J., Chen, X., Borowiec, N., Facon, B., Vogt, H., Price, D. K., Vogel, H., Prud'homme, B., Estoup, A. and Gautier, M. (2020) A Whole-Genome Scan for Association with Invasion Success in the Fruit Fly *Drosophila suzukii* Using Contrasts of Allele Frequencies Corrected for Population Structure. *Molecular Biology and Evolution*, **37**, 2369–2385. URL: <https://doi.org/10.1093/molbev/msaa098>.
- Paradis, E. (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, **26**, 419–420. URL: <https://doi.org/10.1093/bioinformatics/btp696>.

- Pélissié, B., Chen, Y. H., Cohen, Z. P., Crossley, M. S., Hawthorne, D. J., Izzo, V. and Schoville, S. D. (2022) Genome Resequencing Reveals Rapid, Repeated Evolution in the Colorado Potato Beetle. *Molecular Biology and Evolution*, **39**. URL: <https://doi.org/10.1093/molbev/msac016>.
- Petersen, M., Armisen, D., Gibbs, R. A., Hering, L., Khila, A., Mayer, G., Richards, S., Niehuis, O. and Misof, B. (2019) Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Ecology and Evolution*, **21**, 1–15.
- Poelchau, M., Childers, C., Moore, G., Tsavatapalli, V., Evans, J., Lee, C. Y., Lin, H., Lin, J. W. and Hackett, K. (2015) The i5k Workspace@NAL-enabling genomic data access, visualization and curation of arthropod genomes. *Nucleic Acids Research*, **43**, D714–D719.
- Polanco, A. M., Brewster, C. C. and Miller, D. M. (2011) Population growth potential of the bed bug, *Cimex lectularius* L.: A life table analysis. *Insects*, **2**, 173–185.
- Potter, M. F. (2011) The history of bed bug management- With lessons from the past. *American Entomologist*, **57**, 14–25.
- Privé, F., Luu, K., Vilhjálmsson, B. J. and Blum, M. G. B. (2020) Performing Highly Efficient Genome Scans for Local Adaptation with R Package pcadapt Version 4. *Molecular Biology and Evolution*, **37**, 2153–2154. URL: <https://doi.org/10.1093/molbev/msaa053>.
- Pryszcz, L. P. and Gabaldón, T. (2016) Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Research*, **44**, e113–e113. URL: <https://doi.org/10.1093/nar/gkw294>.
- Ranson, H., Claudianos, C., Ortell, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H. and Feyereisen, R. (2002) Evolution of Supergene Families Associated with Insecticide Resistance. *Science*, **298**, 179–181. URL: <https://doi.org/10.1126/science.1076781>.
- Romero, A. and Anderson, T. D. (2016) High levels of resistance in the common bed bug, *cimex lectularius* (hemiptera: Cimicidae), to neonicotinoid insecticides. *Journal of Medical Entomology*, **53**, 727–731.
- Romero, A., Potter, M. F. and Haynes, K. F. (2009a) Behavioral responses of the bed bug to insecticide residues. *Journal of Medical Entomology*, **46**, 51–57.
- (2009b) Evaluation of Piperonyl Butoxide as a Deltamethrin Synergist for Pyrethroid-Resistant Bed Bugs. *Journal of Economic Entomology*, **102**, 2310–2315.
- Romero, A., Potter, M. F., Potter, D. A. and Haynes, K. F. (2007) Insecticide Resistance in the Bed Bug: A Factor in the Pest's Sudden Resurgence? *Journal of Medical Entomology*, **44**, 175–178.
- Schmidt, J. M., Good, R. T., Appleton, B., Sherrard, J., Raymant, G. C., Bogwitz, M. R., Martin, J., Daborn, P. J., Goddard, M. E., Batterham, P. and Robin, C. (2010) Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. *PLoS Genetics*, **6**, 1–11.
- Schrider, D. R., Begun, D. J. and Hahn, M. W. (2013) Detecting highly differentiated copy-number variants from pooled population sequencing. *Pacific Symposium on Biocomputing.*, 344–355.
- Schrider, D. R., Hahn, M. W. and Begun, D. J. (2016) Parallel Evolution of Copy-Number Variation across Continents in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **33**, 1308–1316.
- Schwander, T., Libbrecht, R. and Keller, L. (2014) Supergenes and Complex Phenotypes. *Current Biology*, **24**, R288–R294. URL: <https://doi.org/10.1016/j.cub.2014.01.056>.
- Shi, Y., Jiang, Q., Yang, Y., Feyereisen, R. and Wu, Y. (2021) Pyrethroid metabolism by eleven *Helicoverpa armigera* P450s from the CYP6B and CYP9A subfamilies. *Insect Biochemistry and Molecular Biology*, **135**, 103597. URL: <https://www.sciencedirect.com/science/article/pii/S0965174821000801>.

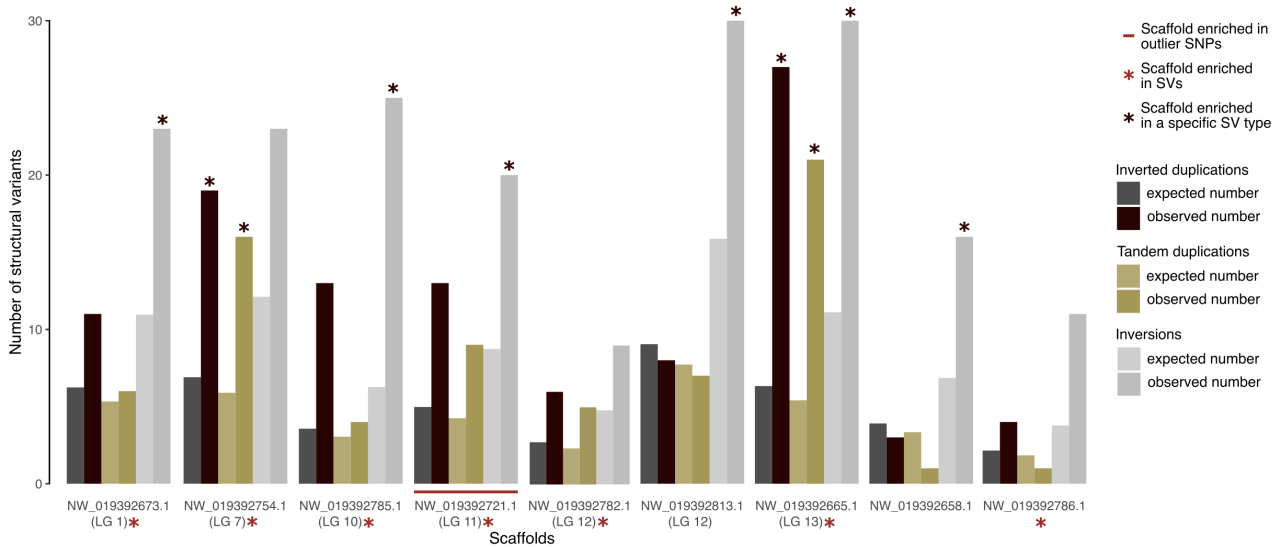
- Stutt, A. D. and Siva-Jothy, M. T. (2001) Traumatic insemination and sexual conflict in the bed bug *Cimex lectularius*. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5683–5687.
- Susser, S. R., Perron, S., Fournier, M., Jacques, L., Denis, G., Tessier, F. and Roberge, P. (2012) Mental health effects from urban bed bug infestation (*Cimex lectularius* L.): a cross-sectional study. *BMJ Open*, **2**. URL: <https://bmjopen.bmj.com/content/2/5/e000838>.
- Tao, F., Si, F.-L., Hong, R., He, X., Li, X.-Y., Qiao, L., He, Z.-B., Yan, Z.-T., He, S.-L. and Chen, B. (2022) Glutathione S-transferase (GST) genes and their function associated with pyrethroid resistance in the malaria vector *Anopheles sinensis*. *Pest Management Science*, **78**, 4127–4139. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ps.7031>.
- Thomas, G. W., Dohmen, E., Hughes, D. S., Murali, S. C., Poelchau, M., Glastad, K., Anstead, C. A., Ayoub, N. A., Batterham, P., Bellair, M., Binford, G. J., Chao, H., Chen, Y. H., Childers, C., Dinh, H., Doddapaneni, H. V., Duan, J. J., Dugan, S., Esposito, L. A., Friedrich, M., Garb, J., Gasser, R. B., Goodisman, M. A., Gundersen-Rindal, D. E., Han, Y., Handler, A. M., Hatakeyama, M., Hering, L., Hunter, W. B., Ioannidis, P., Jayaseelan, J. C., Kalra, D., Khila, A., Korhonen, P. K., Lee, C. E., Lee, S. L., Li, Y., Lindsey, A. R., Mayer, G., McGregor, A. P., McKenna, D. D., Misof, B., Munidasa, M., Munoz-Torres, M., Muzny, D. M., Niehuis, O., Osuji-Lacy, N., Palli, S. R., Panfilio, K. A., Pechmann, M., Perry, T., Peters, R. S., Poynton, H. C., Prpic, N. M., Qu, J., Rotenberg, D., Schal, C., Schoville, S. D., Scully, E. D., Skinner, E., Sloan, D. B., Stouthamer, R., Strand, M. R., Szucsich, N. U., Wijeratne, A., Young, N. D., Zattara, E. E., Benoit, J. B., Zdobnov, E. M., Pfrender, M. E., Hackett, K. J., Werren, J. H., Worley, K. C., Gibbs, R. A., Chipman, A. D., Waterhouse, R. M., Bornberg-Bauer, E., Hahn, M. W. and Richards, S. (2020) Gene content evolution in the arthropods. *Genome Biology*, **21**, 1–14.
- Van Leeuwen, T., Dermauw, W., Mavridis, K. and Vontas, J. (2020) Significance and interpretation of molecular diagnostics for insecticide resistance management of agricultural pests. *Current opinion in insect science*, **39**, 69–76.
- Vergara, C., Latorre, R., Marrion, N. V. and Adelman, J. P. (1998) Calcium-activated potassium channels. *Current Opinion in Neurobiology*, **8**, 321–329. URL: <https://www.sciencedirect.com/science/article/pii/S0959438898800561>.
- Wang, D., Johnson, A. D., Papp, A. C., Kroetz, D. L. and Sadée, W. (2005) Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C>T affects mRNA stability. *Pharmacogenetics and genomics*, **15**, 693–704.
- Wang, K., Zhao, J. N., Bai, J. Y., Shang, Y. Z., Zhang, S. Q., Hou, Y. F., Chen, M. H. and Han, Z. J. (2021) Pyrethroid Resistance and Fitness Cost Conferred by the super-kdr Mutation M918L in *Rhopalosiphum padi* (Hemiptera: Aphididae). *Journal of Economic Entomology*, **114**, 1789–1795. URL: <https://doi.org/10.1093/jee/toab117>.
- Wang, X., Martínez, M.-a., Dai, M., Chen, D., Ares, I., Romero, A., Castellano, V., Martínez, M., Luis, J., Martínez-larrañaga, M.-r., Anadón, A. and Yuan, Z. (2016) Permethrin-induced oxidative stress and toxicity and metabolism. A review. *Environmental Research*, **149**, 86–104. URL: <http://dx.doi.org/10.1016/j.envres.2016.05.003>.
- Weedall, G. D., Mugenzi, L. M. J., Menze, B. D., Tchouakui, M., Ibrahim, S. S., Amvongo-Adjia, N., Irving, H., Wondji, M. J., Tchoupo, M., Djouaka, R., Riveron, J. M. and Wondji, C. S. (2019) A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. *Science translational medicine*, **11**.
- Weetman, D., Djogbenou, L. S. and Lucas, E. (2018) Copy number variation (CNV) and insecticide resistance in mosquitoes: evolving knowledge or an evolving problem? *Current Opinion in Insect Science*, **27**, 82–88. URL: <https://doi.org/10.1016/j.cois.2018.04.005>.
- Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J. and Chen, S. (2012) FastUniq: A Fast De Novo Duplicates Removal Tool for Paired Short Reads. *PLoS ONE*, **7**, 1–6.
- Zhang, Y., Jiang, R., Wu, H., Liu, P., Xie, J., He, Y. and Pang, H. (2012) Next-generation sequencing-based transcriptome analysis of *Cryptolaemus montrouzieri* under insecticide stress reveals resistance-relevant genes in ladybirds. *Genomics*, **100**, 35–41. URL: <http://dx.doi.org/10.1016/j.ygeno.2012.05.002>.
- Zhao, Q., Ma, D., Huang, Y., He, W., Li, Y., Vasseur, L. and You, M. (2018) Genome-wide investigation of transcription factors provides insights into transcriptional regulation in *Plutella xylostella*. *Molecular Genetics and Genomics*, **293**, 435–449. URL: <http://dx.doi.org/10.1007/s00438-017-1389-y>.

- Zhu, F., Gujar, H., Gordon, J. R., Haynes, K. F., Potter, M. F. and Palli, S. R. (2013) Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Scientific Reports*, **3**, 1–8.
- Zhu, F., Sams, S., Mural, T., Haynes, K. F., Potter, M. F. and Palli, S. R. (2012) RNA interference of NADPH-cytochrome P450 reductase results in reduced insecticide resistance in the bed bug, *Cimex lectularius*. *PLoS ONE*, **7**.
- Zhu, F., Wigginton, J., Romero, A., Moore, A., Ferguson, K., Palli, R., Potter, M. F., Haynes, K. F. and Palli, S. R. (2010) Widespread distribution of knockdown resistance mutations in the bed bug, *Cimex lectularius* (Hemiptera: Cimicidae), populations in the United States. *Archives of Insect Biochemistry and Physiology*, **73**, 245–257.

6 | SUPPLEMENTARY



SUPPLEMENTARY FIGURE 1 Proportions of outlier SNPs among SNPs in linkage groups (LG). The dashed line represents average proportion ($\alpha = 0.00019$). Binomial test were performed between observed proportion and average proportion in each LG, and significant corrected p-values were detailed above corresponding LG.



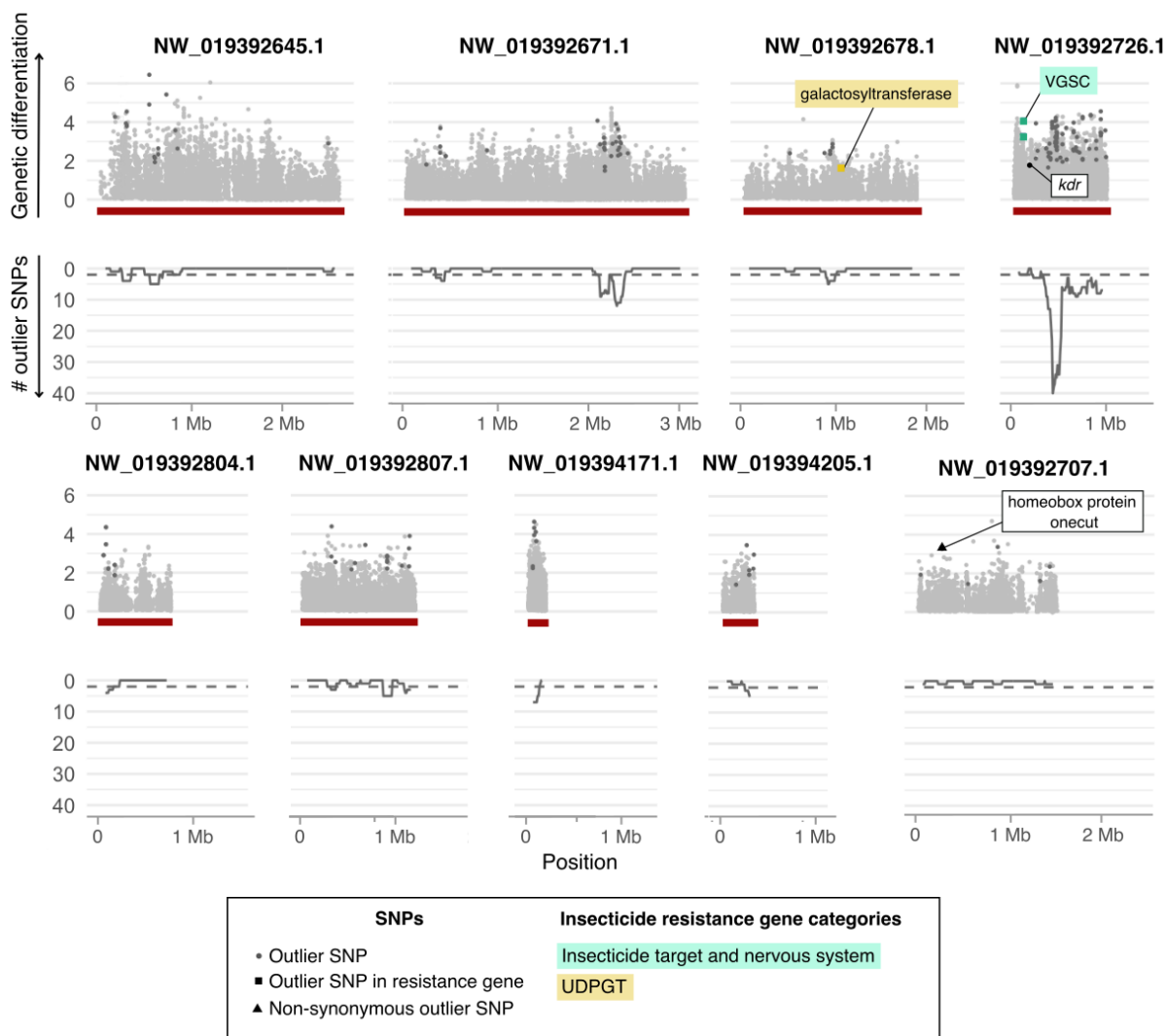
SUPPLEMENTARY FIGURE 2 Expected and observed numbers of structural variants (SVs) for scaffolds enriched with structural variants in general (delineated with a red star below) or enriched with a specific type of structural variant (namely tandem duplications, inverted duplications and inversions, delineated with a black star above corresponding observed number). The scaffold also enriched with outlier SNPs is underlined in red.

	<i>n</i>	slope	<i>df</i>	χ^2	LD ₉₅	95% mortality CI95
London Lab	280	1.840	27	178.10	1.3	91.2-98.8%
German Lab	200	3.123	19	172.97	0.4	89.8-100%
London Field	170	1.030	17	75.76	21.0	89.7-100%
Sweden Field	210	0.941	21	97.04	69.2	90.5-99.5%

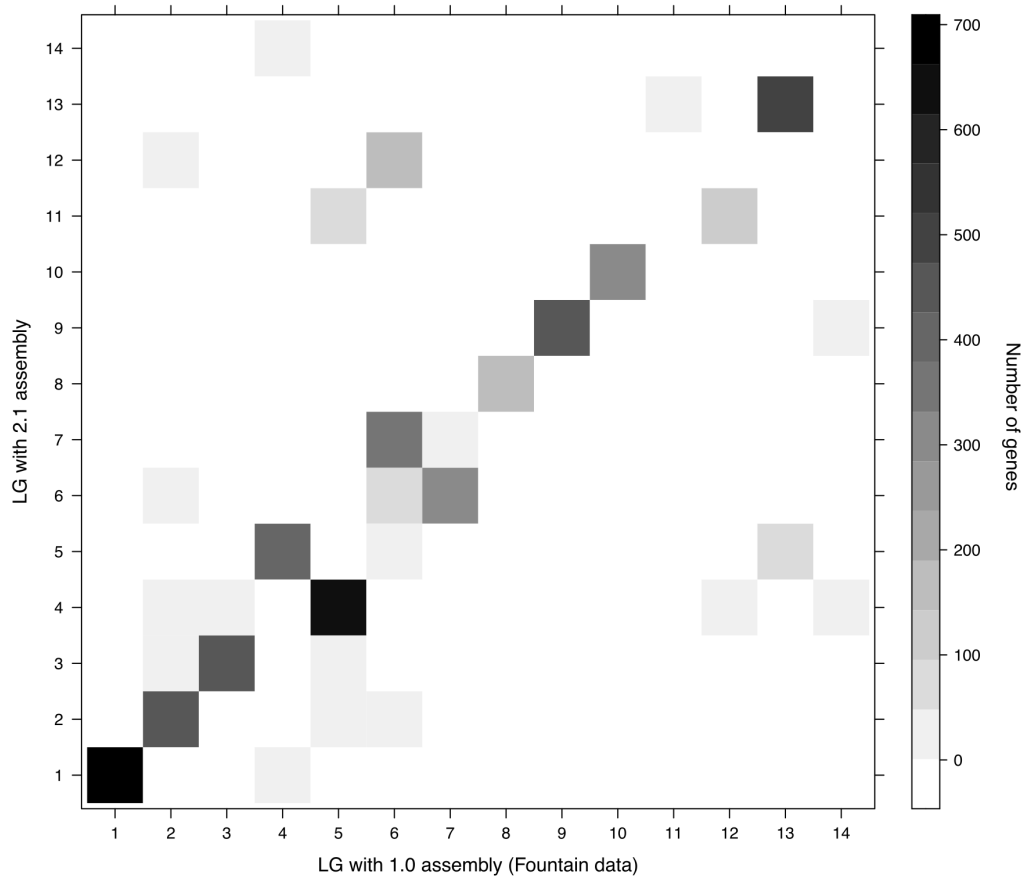
SUPPLEMENTARY TABLE 2 Additional information for dose-mortality analysis performed on *Cimex lectularius* strains. Number of insects (*n*), slope, degree of freedom (*df*) and chi-square (χ^2 , computed as null deviance minus residual deviance), lethal dose (in ng) for 95% of insects in the population (LD₉₅) were given for each strains. The confidence interval at 95% (CI95) for 95% mortality was also computed for each strain using the predict function from the R/ecotox package.

Scaff.	Gene ID	Position	Annotation	Location	F_{alt} LL	F_{alt} LF	F_{ST} (pval)
2763	106666926	419008	GST1-1-like	exon NS	0.20	0.38	-0.01 (0.38)
	106666926	419242	GST1-1-like	exon NS	0.08	0.27	0.05 (0.25)
	106666926	419369	GST1-1-like	exon NS	0.06	0.24	0.03 (0.28)
	106663963	503362	GST-like	exon NS	0.10	0.46	0.16 (0.10)
2721	106663983	1833078	CYP6B5-like	exon NS	0.50	0.93	0.29 (0.03*)
	106663983	1833420	CYP6B5-like	exon NS	0.41	0.79	0.17 (0.09)
	106663983	1833792	CYP6B5-like	exon NS	0.46	0.78	0.11 (0.15)
	106663983	1833793	CYP6B5-like	exon NS	0.46	0.76	0.09 (0.18)
	106663983	1833810	CYP6B5-like	exon NS	0.50	0.77	0.07 (0.20)
	106663983	1833837	CYP6B5-like	exon NS	0.56	0.75	-0.02 (0.40)
	106663983	1836843	CYP6B5-like	exon NS	0.58	1.00	0.33 (0.02*)
	106663983	1838158	CYP6B5-like	exon NS	0.55	0.85	0.13 (0.12)
	106663982	1849007	CYP6B1-like	exon NS	0.27	0.72	0.28 (0.04*)
	106663982	1849044	CYP6B1-like	exon NS	0.29	0.75	0.2 (0.03*)
	106663982	1849067	CYP6B1-like	exon NS	0.29	0.63	0.15 (0.11)
	106663982	1849919	CYP6B1-like	exon NS	0.59	0.84	0.07 (0.21)
	106663981	1873544	CYP6B1-like	exon NS	0.33	0.50	0.00 (0.35)
	106663981	1873558	CYP6B1-like	exon NS	0.33	0.53	0.02 (0.30)
	106663981	1874026	CYP6B1-like	exon NS	0.59	0.66	-0.05 (0.50)
	106663981	1874033	CYP6B1-like	exon NS	0.54	0.57	-0.06 (0.53)
	106663981	1875827	CYP6B1-like	exon NS	0.07	0.25	0.01 (0.32)
	106663981	1875904	CYP6B1-like	exon NS	0.23	0.35	-0.04 (0.46)
	106663981	1875976	CYP6B1-like	exon NS	0.72	0.75	-0.09 (0.69)

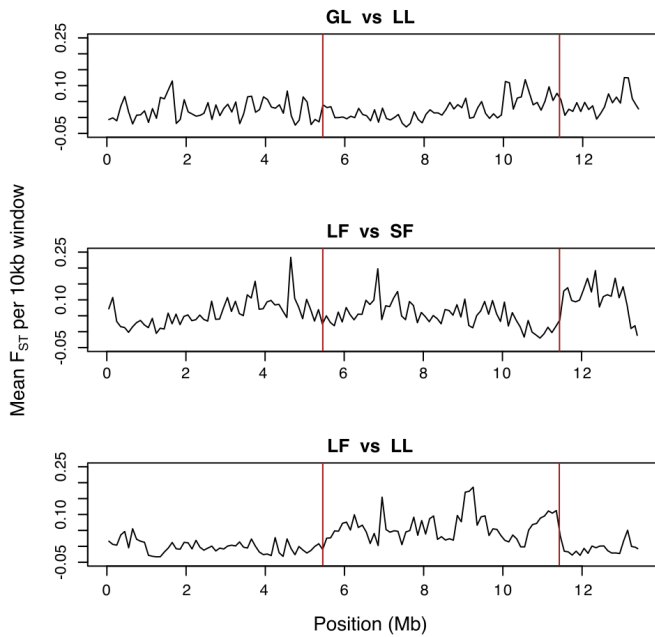
SUPPLEMENTARY TABLE 3 Non-synonymous (NS) SNPs falling within putative resistant genes inside the superlocus detected, with frequency for derived allele from ancestral susceptible population (F_{alt}) superior in London Field (London Field) than in London Lab (London Lab). Gene annotation GST stands for glutathione-S-transferase, CYP for cytochromes P450. To obtain scaffold name, one should add "NW_01939" at the beginning, and ".1" at the end



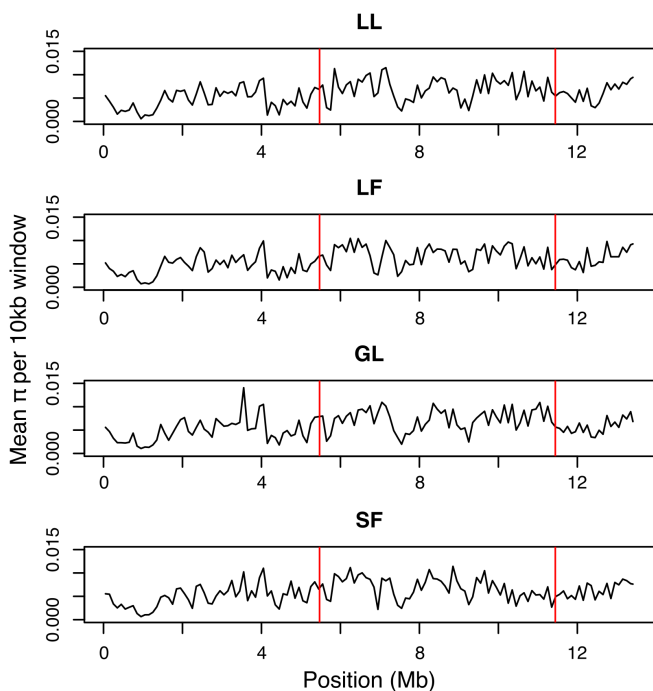
SUPPLEMENTARY FIGURE 3 Several supplementary scaffolds for *Cimex lectularius* genome not located on linkage groups. Scaffolds detected as significantly enriched in outlier SNPs are underlined in red. The scaffold carrying the unique outlier non-synonymous SNP is also represented (NW_019392707.1). In the upper part of a plot, the empirical p-value (p) of genetic differentiation index (F_{ST}) between London Lab and London Field populations in $-\log_{10}$ scale is plotted. The x-axis indicates the position within each scaffold. Outlier SNPs (see Materials & methods) are distinguished by a darker grey. Square-shaped points are the outlier SNPs that fall within candidate resistance genes, and are colored according to those genes' categories. Their protein products are labeled. Non-synonymous outlier mutation was represented with a black triangle. Although not a outlier SNP, the location of the non-synonymous *kdr* L925I mutation was given as an indication. Positions of genes with an expected amplified copy number are displayed with a dashed black vertical line and labelled likewise. In the lower part, "# outlier SNPs" represents the number of outlier SNPs counted in each 100 kb sliding window with a step of 10 kb. Dashed line stands for a threshold of 2 outlier SNPs, which delineates quantile 99%.



SUPPLEMENTARY FIGURE 4 Joint-distribution of genes in Linkage Groups (LG) in [Fountain et al. \(2016\)](#) paper and the present paper. Genes present on the first diagonal are located on the same LG in both cases.



SUPPLEMENTARY FIGURE 5 Differentiation between strains over LG 11. Mean F_{ST} values were computed in 10 kb windows with 10 kb step for three comparisons: between the two susceptible strains (GL = German Lab and LL = London Lab), the two resistant strains (LF = London Field and SF = Sweden Field), and between Londonian strains (LF and LL). The position of a "superlocus" (NW_019392721.1, NW_019392763.1 and NW_019942502.1) is delineated with red lines.



SUPPLEMENTARY FIGURE 6 Nucleotidic diversity within strains over LG 11. Mean π values were computed in 10 kb windows with 10 kb step for the four strains: two susceptible (GL = German Lab and LL = London Lab) and two resistant (LF = London Field and SF = Sweden Field). The position of a "superlocus" (NW_019392721.1, NW_019392763.1 and NW_019942502.1) is delineated with red lines.

LG	Scaff.	Event	Size	Gene ID	Annotation	F_{LL}	F_{LF}	F_{ST}	RDR (q)
1	2640	Inv	324 kb	106672175	sulfotransferase	0	0.09	0.05 (7%)	1.23 (1.21-1.24)
1	2679	InvDup	1.494 Mb	106661337	sulfotransferase	0.06	0.24	0.07 (4%)	1.23 (1.23)
				106661366	choline kinase				
				106661634	CYP450 306a1				
				106661588	CYP450 18a1				
1	2641	Inv	262 kb	106669934	SOD [Cu-Zn]	0	0.10	0.05 (6%)	1.22 (1.21-1.24)
				106669936	SOD [Cu-Zn] 2				
2	2979	Inv	7 kb	106665265	endocuticle str. prot.	0	0.12	0.06 (4%)	1.15 (1.15-1.29)
				106665194	endocuticle str. prot.				
13	2685	Inv	76 kb	106662993	cuticle protein	0.16	0.41	0.08 (2%)	1.23 (1.20-1.25)
		Inv	72 kb	106662993	cuticle protein	0	0.09	0.05 (7%)	1.24 (1.20-1.25)
-	2647	Inv	484 kb	106668191	NADPH oxidored.	0.06	0.26	0.08 (3%)	1.23 (1.23)
-	2656	InvDup	381 kb	106669691	glucosyltransf.	0.22	0.47	0.06 (4%)	1.24 (1.23)
-		Inv	221 kb	112126170	galactosyltransf.	0	0.24	0.13 (1%)	1.22 (1.21-1.24)
				106669680	glucuronosyltransf.				
-	2881	InvDup	27 kb	106665141	cuticle protein	0.02	0.16	0.06 (5%)	1.29 (1.26)
				106664994	cuticle protein				
				106664995	cuticle protein				
				106664992	cuticle protein				

SUPPLEMENTARY TABLE 4 Resistance genes overlapping with structural variants with extreme F_{ST} value (top 10%, specified in brackets). SVs events are delineated as tandem duplications (TandDup), inverted duplications (InvDup), and inversions (Inv). Read Depth Ratio (RDR) between London Field and London Lab was given for each event, along with the threshold quantile (q). Since duplications were all detected in higher frequency in London Field, we only showed the q75 quantile used. For inversions, we showed the q25-q75 interval. "SOD" stands for superoxide dismutase. Gene annotations have been shortened, but full informations can be found using gene ID. To obtain scaffold name, one should add "NW_01939" at the beginning, and ".1" at the end.

SUPPLEMENTARY TABLE 1 Putative resistance genes in bed bug genome (Clec_2.1), by resistance category.

Gene ID	Scaffold	Start	End	Product	Category
LOC10666204	NW_019392863.1	894581	903819	ABC transporter F family member 4-like	ABC transporter / MRP
LOC106663218	NW_019392637.1	1401594	1413659	ABC transporter G family member 1	ABC transporter / MRP
LOC106671981	NW_019392689.1	1926150	2117072	ABC transporter G family member 14	ABC transporter / MRP
LOC106661187	NW_019392653.1	1452064	1462393	ABC transporter G family member 20-like	ABC transporter / MRP
LOC106668456	NW_019392700.1	694634	737446	ABC transporter G family member 20-like	ABC transporter / MRP
LOC106668230	NW_019394219.1	520600	672495	ABC transporter G family member 20	ABC transporter / MRP
LOC106660996	NW_019392653.1	1404658	1420128	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106661276	NW_019392653.1	1427987	1437993	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106663888	NW_019392709.1	60749	80839	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106664452	NW_019392717.1	459642	469182	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106665422	NW_019392709.1	12650	27132	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106664659	NW_019394148.1	209174	216953	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106673832	NW_019392754.1	449700	459353	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106664468	NW_019392717.1	482976	504713	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106666528	NW_019394148.1	169442	190754	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106673850	NW_019392754.1	385063	445710	ABC transporter G family member 23-like	ABC transporter / MRP
LOC106662535	NW_019392661.1	809918	847628	ABC transporter G family member 23	ABC transporter / MRP
LOC106666490	NW_019392718.1	446063	464729	ABC transporter G family member 37-like	ABC transporter / MRP
LOC106668775	NW_019392633.1	694387	715869	ATP-binding cassette sub-family A member 1-like	ABC transporter / MRP
LOC106672589	NW_019392632.1	633502	648482	ATP-binding cassette sub-family A member 2-like	ABC transporter / MRP
LOC106662935	NW_019392812.1	337138	354135	ATP-binding cassette sub-family A member 3-like	ABC transporter / MRP
LOC106663172	NW_019392756.1	255818	268201	ATP-binding cassette sub-family A member 3-like	ABC transporter / MRP
LOC106671149	NW_019392653.1	916833	934568	ATP-binding cassette sub-family A member 3-like	ABC transporter / MRP
LOC106672550	NW_019392832.1	648491	650627	ATP-binding cassette sub-family A member 7-like	ABC transporter / MRP
LOC106672494	NW_019392702.1	1490242	1495435	ATP-binding cassette sub-family B member 10 mitochondrial	ABC transporter / MRP
LOC106668646	NW_019392681.1	1185942	1197855	ATP-binding cassette sub-family B member 6 mitochondrial	ABC transporter / MRP
LOC106670188	NW_019392747.1	1096932	1106296	ATP-binding cassette sub-family B member 7 mitochondrial	ABC transporter / MRP
LOC106666896	NW_019392707.1	528060	543937	ATP-binding cassette sub-family B member 8 mitochondrial	ABC transporter / MRP
LOC106661463	NW_019392641.1	1985786	2060192	ATP-binding cassette sub-family C member Sur	ABC transporter / MRP
LOC106663905	NW_019392657.1	2409303	2427417	ATP-binding cassette sub-family D member 2	ABC transporter / MRP
LOC106671881	NW_019392637.1	2072411	2083220	ATP-binding cassette sub-family D member 3	ABC transporter / MRP
LOC106670179	NW_019392747.1	1302489	1306793	ATP-binding cassette sub-family E member 1	ABC transporter / MRP
LOC106673881	NW_019392677.1	3040630	3046011	ATP-binding cassette sub-family F member 1	ABC transporter / MRP
LOC106671934	NW_019392637.1	2226141	2231152	ATP-binding cassette sub-family F member 2	ABC transporter / MRP
LOC106672894	NW_019392654.1	538104	550426	ATP-binding cassette sub-family F member 3	ABC transporter / MRP
LOC106663239	NW_019394173.1	395611	433682	ATP-binding cassette sub-family G member 1-like	ABC transporter / MRP
LOC106670261	NW_019392668.1	644977	687162	ATP-binding cassette sub-family G member 1-like	ABC transporter / MRP
LOC106672532	NW_019392755.1	989292	1008389	ATP-binding cassette sub-family G member 1-like	ABC transporter / MRP
LOC106669255	NW_019392801.1	192409	271461	ATP-binding cassette sub-family G member 1-like	ABC transporter / MRP
LOC106665537	NW_019392682.1	716378	959324	ATP-binding cassette sub-family G member 1-like	ABC transporter / MRP
LOC106670073	NW_019392644.1	1346254	1453427	ATP-binding cassette sub-family G member 4-like	ABC transporter / MRP
LOC106670072	NW_019392644.1	647119	719373	ATP-binding cassette sub-family G member 4	ABC transporter / MRP
LOC106665518	NW_019392682.1	973424	1009568	ATP-binding cassette sub-family G member 4	ABC transporter / MRP
LOC106670075	NW_019392644.1	749946	794199	ATP-binding cassette sub-family G member 4	ABC transporter / MRP
LOC106670850	NW_019392778.1	339196	345679	ATP-binding cassette sub-family G member 5	ABC transporter / MRP
LOC106662553	NW_019392661.1	104201	110581	multidrug resistance protein 1-like	ABC transporter / MRP
LOC106662607	NW_019392661.1	112026	119542	multidrug resistance protein 1-like	ABC transporter / MRP
LOC106667059	NW_019394211.1	389439	435066	multidrug resistance-associated protein 1-like	ABC transporter / MRP
LOC106667036	NW_019394211.1	318923	358929	multidrug resistance-associated protein 1	ABC transporter / MRP
LOC106667966	NW_019392873.1	700021	715769	multidrug resistance-associated protein 4-like	ABC transporter / MRP
LOC106672163	NW_019392640.1	2029711	2031885	multidrug resistance-associated protein 4-like	ABC transporter / MRP
LOC106672929	NW_019392708.1	1515815	1544980	multidrug resistance-associated protein 5-like	ABC transporter / MRP
LOC106670079	NW_019392644.1	954176	969607	multidrug resistance-associated protein 7	ABC transporter / MRP
LOC106672210	NW_019392640.1	2015525	2029701	probable multidrug resistance-associated protein lethal(2)03659	ABC transporter / MRP
LOC106672227	NW_019392640.1	2056017	2115401	probable multidrug resistance-associated protein lethal(2)03659	ABC transporter / MRP
LOC106662258	NW_019392664.1	1919075	1922947	hexamerin-like	Binding / Sequestration
LOC106673360	NW_019392769.1	90822	94449	hexamerin-like	Binding / Sequestration
LOC106671104	NW_019392805.1	633052	639523	protein NPC2 homolog	Binding / Sequestration
LOC106671106	NW_019392805.1	644144	652415	protein NPC2 homolog	Binding / Sequestration
LOC112126357	NW_019392670.1	1200496	1210727	protein NPC2 homolog	Binding / Sequestration
LOC106670274	NW_019392865.1	275530	277710	SAP30-binding protein	Binding / Sequestration
LOC106662425	NW_019392683.1	1552886	1558158	acyl-coenzyme A thioesterase 9 mitochondrial-like	CCE
LOC106663204	NW_019392685.1	1041829	1062754	acyl-protein thioesterase 1	CCE
LOC106670324	NW_019392863.1	636891	640544	arginine esterase-like	CCE
LOC112126927	NW_019392734.1	50860	51759	arginine esterase-like	CCE
LOC106669634	NW_019392656.1	2048057	2055481	carboxylesterase 5A	CCE
LOC106670290	NW_019392668.1	523854	536542	carboxylesterase 5A-like	CCE
LOC106670743	NW_019392778.1	459396	460624	cholinesterase 1-like	CCE
LOC106670868	NW_019392778.1	435834	439198	cholinesterase 1-like	CCE
LOC106670882	NW_019392778.1	449281	451226	cholinesterase 1-like	CCE
LOC106670883	NW_019392778.1	444562	446480	cholinesterase 1-like	CCE
LOC106670903	NW_019392777.1	44383	46471	cholinesterase 1-like	CCE
LOC106671954	NW_019392637.1	2755058	2757370	cholinesterase 1-like	CCE
LOC106671958	NW_019392637.1	2763409	2766041	cholinesterase 1-like	CCE
LOC106663242	NW_019392637.1	1099206	1101586	cholinesterase 1-like	CCE
LOC106666937	NW_019392707.1	86601	128502	cholinesterase 2	CCE
LOC106670890	NW_019392778.1	455161	457278	cholinesterase-like	CCE
LOC106661624	NW_019392732.1	1875148	1881078	ester hydrolase C11orf54 homolog	CCE
LOC106670759	NW_019392777.1	81232	89923	esterase E4-like	CCE
LOC106670825	NW_019392777.1	39138	41232	esterase E4-like	CCE
LOC106663701	NW_019392743.1	443360	461837	esterase E4-like	CCE
LOC106660901	NW_019392666.1	234405	238450	esterase OVCA2	CCE
LOC106670891	NW_019392778.1	457348	459400	esterase SG1-like	CCE
LOC106670908	NW_019392778.1	440717	444255	esterase-5A-like	CCE
LOC106670923	NW_019392638.1	230610	234244	liver carboxylesterase-like	CCE
LOC106665627	NW_019392737.1	949827	956638	lysosomal thioesterase PPT2 homolog	CCE
LOC106667550	NW_019392795.1	137560	153207	neuropathy target esterase sws	CCE
LOC106663184	NW_019392685.1	118288	119658	palmitoyl-protein thioesterase 1	CCE
LOC106670761	NW_019392777.1	76628	78620	para-nitrobenzyl esterase-like	CCE
LOC106670478	NW_019392645.1	1193850	1213409	resistance to inhibitors of cholinesterase protein 3	CCE
LOC106661555	NW_019392641.1	1361158	1365409	ubiquitin thioesterase OTU1	CCE
LOC106672454	NW_019392811.1	383175	391075	ubiquitin thioesterase otubain-like	CCE
LOC106662178	NW_019392688.1	3779390	3792108	ubiquitin thioesterase trabd	CCE
LOC106663630	NW_019392663.1	632384	635691	venom carboxylesterase-6-like	CCE
LOC106667314	NW_019392665.1	984011	986758	venom carboxylesterase-6-like	CCE
LOC106670760	NW_019392777.1	89918	92232	venom carboxylesterase-6-like	CCE
LOC106670763	NW_019392777.1	71786	75089	venom carboxylesterase-6-like	CCE
LOC106670922	NW_019392638.1	223021	230259	venom carboxylesterase-6-like	CCE
LOC106667302	NW_019392665.1	652726	655958	venom carboxylesterase-6-like	CCE
LOC106663757	NW_019392793.1	1353071	1359102	venom carboxylesterase-6-like	CCE
LOC106662755	NW_019392634.1	1017093	1017933	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106663419	NW_019394230.1	1112100	1112505	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106668007	NW_019392660.1	978871	979333	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106668008	NW_019392660.1	995520	995988	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106669225	NW_019392801.1	329958	330580	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106669244	NW_019392801.1	317773	318220	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106670764	NW_019394221.1	291792	292643	adult-specific cuticular protein ACP-20-like	Cuticle

LOC106670985	NW_019392762.1	377090	377578	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106669243	NW_019392801.1	314545	321343	adult-specific cuticular protein ACP-20-like	Cuticle
LOC106668974	NW_019392829.1	111463	122731	adult-specific cuticular protein ACP-22-like	Cuticle
LOC106662330	NW_019392769.1	1412597	1419601	chitinase-3-like protein 2	Cuticle
LOC106664620	NW_019392699.1	1244558	1254915	chitinase-like protein 4	Cuticle
LOC106663144	NW_019392637.1	1377034	1384646	chitinase-like protein EN03	Cuticle
LOC106672198	NW_019392640.1	2860143	2862023	cuticle protein 1-like	Cuticle
LOC106662952	NW_019392685.1	763191	763978	cuticle protein 10.6-like	Cuticle
LOC106662042	NW_019392675.1	595496	597255	cuticle protein 12.5-like	Cuticle
LOC106662986	NW_019392685.1	756768	757614	cuticle protein 12.5-like	Cuticle
LOC106663252	NW_019392685.1	739604	742836	cuticle protein 12.5-like	Cuticle
LOC106662949	NW_019392685.1	644458	645845	cuticle protein 16.5-like	Cuticle
LOC106662993	NW_019392685.1	660245	660906	cuticle protein 16.5-like	Cuticle
LOC106663240	NW_019392685.1	771781	776163	cuticle protein 16.5-like	Cuticle
LOC106664538	NW_019393130.1	676107	677165	cuticle protein 18.6-like	Cuticle
LOC106669685	NW_019392851.1	67591	75016	cuticle protein 18.7-like	Cuticle
LOC106670323	NW_019392851.1	329822	331325	cuticle protein 18.7-like	Cuticle
LOC106670437	NW_019392851.1	54557	58081	cuticle protein 18.7-like	Cuticle
LOC106669157	NW_019392922.1	265243	287794	cuticle protein 19	Cuticle
LOC106662738	NW_019392959.1	66893	72056	cuticle protein 19-like	Cuticle
LOC106670164	NW_019392747.1	1087491	1092297	cuticle protein 19-like	Cuticle
LOC106672090	NW_019392669.1	961973	966399	cuticle protein 19-like	Cuticle
LOC106672091	NW_019392669.1	926720	944937	cuticle protein 19-like	Cuticle
LOC106665192	NW_019392861.1	209716	211152	cuticle protein 19-like	Cuticle
LOC106665132	NW_019392689.1	1133359	1135343	cuticle protein 2-like	Cuticle
LOC106666081	NW_019392863.1	1012356	1017859	cuticle protein 2-like	Cuticle
LOC106666880	NW_019392920.1	121808	126199	cuticle protein 21-like	Cuticle
LOC106667149	NW_019392920.1	117064	117615	cuticle protein 21-like	Cuticle
LOC106668084	NW_019392647.1	1807698	1811330	cuticle protein 21-like	Cuticle
LOC106672092	NW_019392669.1	975462	981311	cuticle protein 21-like	Cuticle
LOC106668796	NW_019392633.1	736551	743743	cuticle protein 21-like	Cuticle
LOC106668980	NW_019393016.1	61375	61860	cuticle protein 6.4-like	Cuticle
LOC106662967	NW_019392685.1	749624	750953	cuticle protein 64-like	Cuticle
LOC106661691	NW_019392675.1	571822	572607	cuticle protein 67-like	Cuticle
LOC106664977	NW_019392881.1	751992	752720	cuticle protein 7-like	Cuticle
LOC106664992	NW_019392881.1	847326	848130	cuticle protein 7-like	Cuticle
LOC106664993	NW_019392881.1	784330	785195	cuticle protein 7-like	Cuticle
LOC106664994	NW_019392881.1	829254	830461	cuticle protein 7-like	Cuticle
LOC106664995	NW_019392881.1	836386	836895	cuticle protein 7-like	Cuticle
LOC106665141	NW_019392881.1	825669	826488	cuticle protein 7-like	Cuticle
LOC106665144	NW_019392881.1	813522	814333	cuticle protein 7-like	Cuticle
LOC106665145	NW_019392881.1	818639	819577	cuticle protein 7-like	Cuticle
LOC106670088	NW_019392644.1	1098171	1100062	cuticle protein 7-like	Cuticle
LOC106664996	NW_019392881.1	854257	854768	cuticle protein 7-like	Cuticle
LOC106666064	NW_019392993.1	2456896	2458930	cuticle protein 7-like	Cuticle
LOC106663264	NW_019392685.1	827380	829182	cuticle protein 70 isoforms A and B-like	Cuticle
LOC106665209	NW_019392728.1	271279	280884	cuticle protein CP14.6-like	Cuticle
LOC106662043	NW_019392675.1	574693	581201	cuticle protein-like	Cuticle
LOC106666062	NW_019392993.1	2437772	2454768	cuticle protein-like	Cuticle
LOC106666098	NW_019392993.1	2462850	2468539	cuticle protein-like	Cuticle
LOC106664154	NW_019392770.1	576990	592635	endochitinase	Cuticle
LOC106665267	NW_019392979.1	2661	12567	endocuticle structural glycoprotein ABD-4-like	Cuticle
LOC106665312	NW_019392979.1	141880	145662	endocuticle structural glycoprotein ABD-4-like	Cuticle
LOC106673133	NW_019392773.1	29819	56248	endocuticle structural glycoprotein ABD-4-like	Cuticle
LOC106672558	NW_019392832.1	494386	495149	endocuticle structural glycoprotein ABD-5-like	Cuticle
LOC106672559	NW_019392832.1	499052	499735	endocuticle structural glycoprotein ABD-5-like	Cuticle
LOC106665182	NW_019392822.1	813529	825042	endocuticle structural glycoprotein SgAbd-1-like	Cuticle
LOC106665194	NW_019392979.1	33885	34615	endocuticle structural glycoprotein SgAbd-2-like	Cuticle
LOC106665265	NW_019392979.1	28428	29138	endocuticle structural glycoprotein SgAbd-2-like	Cuticle
LOC106665280	NW_019392822.1	723158	725314	endocuticle structural glycoprotein SgAbd-2-like	Cuticle
LOC106665281	NW_019392822.1	732013	734777	endocuticle structural glycoprotein SgAbd-2-like	Cuticle
LOC106665190	NW_019392979.1	103231	107539	endocuticle structural glycoprotein SgAbd-3-like	Cuticle
LOC106661226	NW_019392719.1	1430179	1435817	endocuticle structural glycoprotein SgAbd-8-like	Cuticle
LOC106665307	NW_019392979.1	162088	167574	endocuticle structural glycoprotein SgAbd-8-like	Cuticle
LOC106665184	NW_019392822.1	780550	784904	endocuticle structural glycoprotein SgAbd-9-like	Cuticle
LOC106672421	NW_019392811.1	827642	841408	larval cuticle protein 1-like	Cuticle
LOC106665305	NW_019392979.1	72345	74288	larval cuticle protein 16/17-like	Cuticle
LOC106664047	NW_019392920.1	77411	82689	larval cuticle protein A1A-like	Cuticle
LOC106662714	NW_019392854.1	414287	423622	larval cuticle protein A2B-like	Cuticle
LOC106663020	NW_019393130.1	646663	647367	larval cuticle protein A2B-like	Cuticle
LOC106663472	NW_019393130.1	664025	664742	larval cuticle protein A2B-like	Cuticle
LOC106664854	NW_019393130.1	625738	626515	larval cuticle protein A2B-like	Cuticle
LOC106664863	NW_019393130.1	630029	630899	larval cuticle protein A2B-like	Cuticle
LOC106664944	NW_019392920.1	33208	36857	larval cuticle protein A2B-like	Cuticle
LOC106664954	NW_019392920.1	18033	19719	larval cuticle protein A2B-like	Cuticle
LOC106665779	NW_019392920.1	134703	138892	larval cuticle protein A2B-like	Cuticle
LOC112128313	NW_019394221.1	284208	288241	larval cuticle protein A2B-like	Cuticle
LOC106664300	NW_019393130.1	673262	674032	larval cuticle protein A3A-like	Cuticle
LOC106664492	NW_019392649.1	1017882	1018792	larval cuticle protein A3A-like	Cuticle
LOC106664494	NW_019392649.1	1014291	1015173	larval cuticle protein A3A-like	Cuticle
LOC106664495	NW_019392649.1	1046176	1046953	larval cuticle protein A3A-like	Cuticle
LOC106664847	NW_019393130.1	622264	623309	larval cuticle protein A3A-like	Cuticle
LOC106668088	NW_019392647.1	1814837	1818873	larval cuticle protein A3A-like	Cuticle
LOC106663241	NW_019392685.1	780017	780782	larval cuticle protein F1-like	Cuticle
LOC106672543	NW_019392708.1	1008580	1011016	larval cuticle protein LCP-17-like	Cuticle
LOC106665309	NW_019392979.1	186959	198238	larval cuticle protein LCP-22-like	Cuticle
LOC106665203	NW_019392822.1	700405	706482	larval cuticle protein LCP-30-like	Cuticle
LOC106661705	NW_019392675.1	670277	676250	larval/pupal cuticle protein H1C-like	Cuticle
LOC106662953	NW_019392685.1	793130	801686	larval/pupal cuticle protein H1C-like	Cuticle
LOC106662415	NW_019392683.1	1725398	1739059	probable chitinase 10	Cuticle
LOC106663441	NW_019394219.1	2457055	2530941	probable chitinase 10	Cuticle
LOC106672194	NW_019392640.1	2848741	2856866	probable chitinase 2	Cuticle
LOC106661029	NW_019392719.1	1392866	1401379	pupal cuticle protein 20-like	Cuticle
LOC106665200	NW_019393114.1	1484	16460	pupal cuticle protein 27-like	Cuticle
LOC106665180	NW_019392822.1	517155	521561	pupal cuticle protein 36-like	Cuticle
LOC106665279	NW_019392822.1	570971	574856	pupal cuticle protein 36-like	Cuticle
LOC106662985	NW_019394176.1	18368	18922	pupal cuticle protein C1B-like	Cuticle
LOC106661690	NW_019392675.1	552737	572619	pupal cuticle protein C1B-like	Cuticle
LOC106664983	NW_019392689.1	1140222	1141153	pupal cuticle protein Edg-78E	Cuticle
LOC112127002	NW_019392747.1	513719	541601	pupal cuticle protein Edg-84A-like	Cuticle
LOC112126514	NW_019392685.1	768214	769391	pupal cuticle protein G1A-like	Cuticle
LOC10666844	NW_019392851.1	233508	234507	pupal cuticle protein-like	Cuticle
LOC106669676	NW_019392851.1	85988	88614	pupal cuticle protein-like	Cuticle
LOC106670525	NW_019392851.1	220047	220929	pupal cuticle protein-like	Cuticle
LOC10666744	NW_019392851.1	302740	306657	pupal cuticle protein-like	Cuticle
LOC106670614	NW_019392851.1	292169	295062	pupal cuticle protein-like	Cuticle
LOC106666926	NW_019392763.1	418843	424890	glutathione S-transferase 1-1-like	GST
LOC106672932	NW_019392708.1	1597703	1613296	glutathione S-transferase 1-like	GST

LOC106672874	NW_019392654.1	1227157	1235076	glutathione S-transferase C-terminal domain-containing protein homolog	GST
LOC106669350	NW_019392673.1	775187	776890	glutathione S-transferase theta-1-like	GST
LOC106674273	NW_019392879.1	108224	112470	glutathione S-transferase theta-1-like	GST
LOC106663963	NW_019392763.1	502888	504185	glutathione S-transferase-like	GST
LOC106664023	NW_019392763.1	493269	495040	glutathione S-transferase-like	GST
LOC106670087	NW_019392644.1	1037675	1040967	glutathione S-transferase-like	GST
LOC106674256	NW_019392929.1	126420	138992	glutathione S-transferase-like	GST
LOC106663686	NW_019392743.1	915216	922800	glutathione S-transferase-like	GST
LOC106674292	NW_019392705.1	92556	112962	glutathione S-transferase-like	GST
LOC106660806	NW_019392860.1	407581	412358	microsomal glutathione S-transferase 1	GST
LOC106660807	NW_019392860.1	401137	405986	microsomal glutathione S-transferase 1-like	GST
LOC106666464	NW_019393002.1	87719	92139	microsomal glutathione S-transferase 1-like	GST
LOC112127332	NW_019392803.1	7	886	acetylcholinesterase 1-like	Insecticide targets and nervous system
LOC106669437	NW_019392803.1	7080	7993	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106669386	NW_019394179.1	14474	95844	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106669436	NW_019394179.1	823	3373	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106671932	NW_019392637.1	3045820	3055263	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106664272	NW_019392716.1	1473387	1612281	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106670762	NW_019392777.1	96556	101099	acetylcholinesterase-like	Insecticide targets and nervous system
LOC106667833	NW_019392726.1	70489	160309	sodium channel protein para	Insecticide targets and nervous system
LOC106663917	NW_019392731.1	736886	841540	acetylcholine receptor subunit alpha-L1	Insecticide targets and nervous system
LOC106663856	NW_019392731.1	368096	640645	acetylcholine receptor subunit alpha-like 1	Insecticide targets and nervous system
LOC106667040	NW_019392752.1	515203	667159	acetylcholine receptor subunit alpha-like 1	Insecticide targets and nervous system
LOC106661487	NW_019392641.1	1376370	1733207	acetylcholine receptor subunit alpha-like	Insecticide targets and nervous system
LOC106663555	NW_019393021.1	19562	130418	acetylcholine receptor subunit alpha-like	Insecticide targets and nervous system
LOC106667659	NW_019392711.1	730818	735387	acetylcholine receptor subunit alpha	Insecticide targets and nervous system
LOC106670509	NW_019392645.1	983067	1006610	acetylcholine receptor subunit beta-like 1	Insecticide targets and nervous system
LOC106663897	NW_019392731.1	854544	901879	acetylcholine receptor subunit beta-like 2	Insecticide targets and nervous system
LOC106664183	NW_019392813.1	3010999	3019778	sodium channel protein Nach-like	Insecticide targets and nervous system
LOC112126565	NW_019392689.1	745189	755441	sodium channel protein Nach-like	Insecticide targets and nervous system
LOC106666198	NW_019392813.1	3020164	3026288	sodium channel protein Nach-like	Insecticide targets and nervous system
LOC106663131	NW_019392685.1	573164	576724	sodium channel protein Nach-like	Insecticide targets and nervous system
LOC106671947	NW_019392637.1	2736171	2742098	sodium channel protein Nach-like	Insecticide targets and nervous system
LOC106662446	NW_019392730.1	1347727	1365982	sodium channel protein Nach	Insecticide targets and nervous system
LOC106664013	NW_019392721.1	654275	728579	choline O-acetyltransferase	Insecticide targets and nervous system
LOC106663497	NW_019392994.1	248717	283175	choline transporter-like protein 1	Insecticide targets and nervous system
LOC106663199	NW_019392637.1	1641739	1647995	choline transporter-like protein 1	Insecticide targets and nervous system
LOC106664599	NW_019392699.1	1393361	1399325	choline-phosphate cytidyltransferase A-like	Insecticide targets and nervous system
LOC106661366	NW_019392679.1	581342	587489	choline/ethanolamine kinase	Insecticide targets and nervous system
LOC106670009	NW_019392767.1	34877	62279	cholinephosphotransferase 1	Insecticide targets and nervous system
LOC106664045	NW_019392721.1	2708877	2785986	high-affinity choline transporter 1	Insecticide targets and nervous system
LOC106665165	NW_019393001.1	1533	223145	muscarinic acetylcholine receptor DM1	Insecticide targets and nervous system
LOC106669459	NW_019392673.1	1193563	1544001	muscarinic acetylcholine receptor M3	Insecticide targets and nervous system
LOC106664028	NW_019392721.1	1129637	1924395	acid sphingomyelinase-like phosphodiesterase 3a	Other detox
LOC106673997	NW_019394227.1	1409312	1560623	acid sphingomyelinase-like phosphodiesterase 3b	Other detox
LOC106673346	NW_019392903.1	44063	57038	alcohol dehydrogenase [NADP(+)]-like	Other detox
LOC106665551	NW_019394228.1	1116919	1123701	alcohol dehydrogenase class-3	Other detox
LOC106663326	NW_019392658.1	1385714	1415831	aldehyde dehydrogenase family 3 member B1-like	Other detox
LOC106663925	NW_019392831.1	329070	335035	aldehyde dehydrogenase dimeric NADP-prefering-like	Other detox
LOC106662070	NW_019392772.1	792027	802270	aldehyde dehydrogenase mitochondrial-like	Other detox
LOC106662313	NW_019392772.1	777412	786641	aldehyde dehydrogenase mitochondrial-like	Other detox
LOC106673635	NW_019393245.1	875	2398	aldehyde dehydrogenase mitochondrial-like	Other detox
LOC106664902	NW_019394178.1	638588	657496	alpha-aminoadipic semialdehyde dehydrogenase	Other detox
LOC106667764	NW_019392779.1	768597	956181	aryl hydrocarbon receptor	Other detox
LOC106672777	NW_019392714.1	458134	463171	aryl hydrocarbon receptor nuclear translocator homolog	Other detox
LOC106661160	NW_019392643.1	1080074	1199232	aryl hydrocarbon receptor nuclear translocator homolog	Other detox
LOC112127883	NW_019392939.1	115521	218774	aryl hydrocarbon receptor nuclear translocator-like protein 1	Other detox
LOC106662809	NW_019392631.1	346739	535797	bifunctional heparan sulfate N-deacetylase/N-sulfotransferase	Other detox
LOC106666582	NW_019394202.1	1886643	1895153	carbohydrate sulfotransferase 11	Other detox
LOC106666671	NW_019392800.1	437313	536222	carbohydrate sulfotransferase 11	Other detox
LOC106668062	NW_019392680.1	1119696	1129356	carbohydrate sulfotransferase 11	Other detox
LOC106666364	NW_019392957.1	578767	616784	carbohydrate sulfotransferase 4-like	Other detox
LOC106666406	NW_019392957.1	692185	699202	cytochrome b5	Other detox
LOC106670348	NW_019392630.1	226656	227971	cytochrome b5 domain-containing protein 1	Other detox
LOC106665027	NW_019392796.1	62293	70934	cytochrome b5 reductase 4	Other detox
LOC106665775	NW_019392748.1	1618736	1624233	cytochrome b5-like	Other detox
LOC112128019	NW_019393008.1	209329	210083	cytochrome b5-related protein-like	Other detox
LOC106661542	NW_019392764.1	512150	523899	cytochrome b5-related protein-like	Other detox
LOC106668953	NW_019392675.1	2613630	2626277	epoxide hydrolase 4-like	Other detox
LOC106660930	NW_019392860.1	365781	370507	epoxide hydrolase 4-like	Other detox
LOC106670929	NW_019392638.1	109262	138520	estrogen sulfotransferase-like	Other detox
LOC106672175	NW_019392640.1	1462403	1467250	estrogen sulfotransferase-like	Other detox
LOC106666127	NW_019392993.1	3311565	3314020	galactosylceramide sulfotransferase-like	Other detox
LOC106673145	NW_019392850.1	508758	518319	heparan sulfate 2-O-sulfotransferase 1	Other detox
LOC106665955	NW_019392781.1	529446	602890	heparan sulfate 2-O-sulfotransferase pipe	Other detox
LOC106674312	NW_019392694.1	1504524	1552194	heparan sulfate glucosamine 3-O-sulfotransferase 1	Other detox
LOC106669891	NW_019392636.1	827366	854933	heparan sulfate glucosamine 3-O-sulfotransferase 6	Other detox
LOC106663693	NW_019392793.1	690375	702355	heparan-sulfate 6-O-sulfotransferase 1	Other detox
LOC106667598	NW_019392803.1	787421	811734	indole-3-acetaldehyde oxidase-like	Other detox
LOC106673826	NW_019392677.1	942107	962765	juvenile hormone epoxide hydrolase 1-like	Other detox
LOC112128252	NW_019394213.1	409405	415561	L-aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	Other detox
LOC106674102	NW_019392722.1	1450387	1498141	Niemann-Pick C1 protein	Other detox
LOC106670210	NW_019392936.1	98823	108522	Niemann-Pick type protein homolog 1B-like	Other detox
LOC106673939	NW_019392886.1	727838	731368	phenoloxidase 2	Other detox
LOC106667591	NW_019392795.1	316208	321043	probable methylmalonate-semialdehyde dehydrogenase [acylating] mitochondrial	Other detox
LOC106664801	NW_019392742.1	1238897	1243972	protein-tyrosine sulfotransferase	Other detox
LOC106670655	NW_019392696.1	575979	582581	repressor of RNA polymerase III transcription MAF1 homolog	Other detox
LOC106670919	NW_019392638.1	815190	988569	segmentation protein cap'n'collar	Other detox
LOC106671553	NW_019392663.1	462453	474509	succinate-semialdehyde dehydrogenase mitochondrial	Other detox
LOC106662801	NW_019392631.1	125680	145906	sulfotransferase 1 family member D1-like	Other detox
LOC106662802	NW_019392631.1	116762	165595	sulfotransferase 1C1-like	Other detox
LOC106663680	NW_019392765.1	14787	31692	sulfotransferase 1C4-like	Other detox
LOC106661337	NW_019392679.1	394571	398165	sulfotransferase 4A1-like	Other detox
LOC106661186	NW_019392686.1	1069185	1075697	sulfotransferase family cytosolic 1B member 1-like	Other detox
LOC106664404	NW_019392716.1	1930673	1932818	sulfotransferase family cytosolic 1B member 1-like	Other detox
LOC106666818	NW_019392742.1	1771166	1773352	transcription factor MafB	Other detox
LOC106669401	NW_019392709.1	1562469	1564464	transcription factor MafK	Other detox
LOC106670655	NW_019392696.1	575979	582581	repressor of RNA polymerase III transcription MAF1 homolog	Other detox
LOC106667952	NW_019393013.1	34961	111566	probable G-protein coupled receptor 158	Other detox
LOC112128143	NW_019393512.1	30471	45678	probable G-protein coupled receptor 83	Other detox
LOC106669197	NW_019392712.1	1892100	2330018	probable G-protein coupled receptor CG31760	Other detox
LOC106670401	NW_019392630.1	468551	528616	probable G-protein coupled receptor CG31760	Other detox
LOC106664904	NW_019392742.1	920092	1021672	probable G-protein coupled receptor Mth-like 1	Other detox
LOC106670699	NW_019392698.1	735577	743099	probable G-protein coupled receptor Mth-like 5	Other detox
LOC106661588	NW_019392679.1	1369475	1372931	cytochrome P450 18a1	P450
LOC106669063	NW_019392802.1	661471	669431	cytochrome P450 302a1 mitochondrial	P450
LOC106661634	NW_019392679.1	1331257	1363174	cytochrome P450 306a1	P450
LOC106668669	NW_019392731.1	4269610	4281806	cytochrome P450 307a1-like	P450

LOC106663742	NW_019392793.1	1013009	1020204	cytochrome P450 315a1 mitochondrial	P450
LOC106662345	NW_019392664.1	1182141	1185651	cytochrome P450 4C1-like	P450
LOC106665122	NW_019392689.1	1062167	1069390	cytochrome P450 4C1-like	P450
LOC106669268	NW_019392801.1	1599	16364	cytochrome P450 4C1-like	P450
LOC106673663	NW_019392702.1	1559979	1568275	cytochrome P450 4C1-like	P450
LOC106666332	NW_019392703.1	337054	346222	cytochrome P450 4C1-like	P450
LOC106666379	NW_019392703.1	324489	328820	cytochrome P450 4C1-like	P450
LOC106667061	NW_019392678.1	997347	1013904	cytochrome P450 4c3	P450
LOC106666149	NW_019393086.1	416683	430861	cytochrome P450 4g15-like	P450
LOC106666150	NW_019393086.1	394289	404788	cytochrome P450 4g15-like	P450
LOC106666181	NW_019392993.1	3657260	3663526	cytochrome P450 4g15-like	P450
LOC106666182	NW_019392993.1	3669602	3682596	cytochrome P450 4g15-like	P450
LOC106672025	NW_019392751.1	472315	475079	cytochrome P450 4V2-like	P450
LOC106663845	NW_019393011.1	1379924	1384041	cytochrome P450 6a2-like	P450
LOC106663846	NW_019393011.1	1347676	1361025	cytochrome P450 6a2-like	P450
LOC106664001	NW_019393218.1	2462	3424	cytochrome P450 6a2-like	P450
LOC106663982	NW_019392721.1	1845460	1853897	cytochrome P450 6B1-like	P450
LOC106667256	NW_019392665.1	616424	622095	cytochrome P450 6B1-like	P450
LOC106663981	NW_019392721.1	1865271	1876491	cytochrome P450 6B1-like	P450
LOC106663983	NW_019392721.1	1827917	1838338	cytochrome P450 6B5-like	P450
LOC106670428	NW_019392691.1	979767	988537	cytochrome P450 6j1-like	P450
LOC106670429	NW_019392691.1	905360	913829	cytochrome P450 6j1-like	P450
LOC106674131	NW_019392694.1	1193629	1200458	cytochrome P450 6j1-like	P450
LOC106662622	NW_019392661.1	361896	365391	cytochrome P450 6k1-like	P450
LOC106665507	NW_019392942.1	40597	49096	cytochrome P450 6k1-like	P450
LOC106670925	NW_019392638.1	198636	208352	cytochrome P450 9e2-like	P450
LOC106670912	NW_019392638.1	242129	253284	cytochrome P450 9e2-like	P450
LOC106668127	NW_019392647.1	1215470	1256385	ecdysone 20-monoxygenase	P450
LOC106666447	NW_019394148.1	129498	133257	heme oxygenase 1	P450
LOC106664686	NW_019392965.1	15481	124122	NADPH oxidase 5	P450
LOC106664146	NW_019392695.1	1420309	1432335	probable cytochrome P450 301a1 mitochondrial	P450
LOC106665205	NW_019392657.1	2128592	2140338	probable cytochrome P450 303a1	P450
LOC106662677	NW_019392939.1	266752	272174	probable cytochrome P450 305a1	P450
LOC106670089	NW_019392644.1	818922	821824	probable cytochrome P450 49a1	P450
LOC106672902	NW_019392935.1	221932	265202	probable cytochrome P450 49a1	P450
LOC106663722	NW_019392765.1	334283	347144	probable cytochrome P450 6a13	P450
LOC106670924	NW_019392638.1	236123	242221	probable cytochrome P450 6a13	P450
LOC106667291	NW_019392665.1	552269	559967	probable cytochrome P450 6a14	P450
LOC106670277	NW_019392668.1	1276458	1279945	probable cytochrome P450 6a14	P450
LOC106673710	NW_019392665.1	518591	524400	probable cytochrome P450 6a14	P450
LOC106667404	NW_019392665.1	575815	586206	probable cytochrome P450 6a14	P450
LOC106667287	NW_019392665.1	526060	538429	probable cytochrome P450 6a14	P450
LOC106667392	NW_019392665.1	591986	601172	probable cytochrome P450 6a14	P450
LOC106673506	NW_019393645.1	445	971	probable cytochrome P450 6a17	P450
LOC106673503	NW_019393105.1	4880	6517	probable cytochrome P450 6a18	P450
LOC106673525	NW_019392677.1	2953292	2955150	probable cytochrome P450 6a18	P450
LOC106662957	NW_019392756.1	244716	249453	probable cytochrome P450 6a20	P450
LOC106673495	NW_019392667.1	3231079	3232890	probable cytochrome P450 6d5	P450
LOC106673892	NW_019392667.1	3213924	3215662	probable cytochrome P450 6d5	P450
LOC112128126	NW_019393228.1	8	1146	probable cytochrome P450 9f2	P450
LOC112128127	NW_019393228.1	1798	2730	probable cytochrome P450 9f2	P450
LOC106672837	NW_019392880.1	116018	126694	probable cytochrome P450 CYP44	P450
LOC106668336	NW_019392673.1	1948364	1986143	NADPH-cytochrome P450 reductase	Red/Ox
LOC106668191	NW_019392647.1	1519656	1527622	NADPH-dependent diflavin oxidoreductase 1	Red/Ox
LOC10666428	NW_019394219.1	731832	738671	NADPH adrenodoxin oxidoreductase mitochondrial	Red/Ox
LOC106668215	NW_019392655.1	711818	716899	catalase-like	Red/Ox
LOC106669210	NW_019392973.1	11971	23778	catalase-like	Red/Ox
LOC106665783	NW_019392930.1	214657	218440	catalase-like	Red/Ox
LOC106663060	NW_019392939.1	431188	468198	catalase	Red/Ox
LOC106669936	NW_019392641.1	1026998	1054250	superoxide dismutase [Cu-Zn] 2-like	Red/Ox
LOC106669934	NW_019392641.1	990484	1009113	superoxide dismutase [Cu-Zn]-like	Red/Ox
LOC106672504	NW_019392877.1	204708	214564	superoxide dismutase [Cu-Zn]	Red/Ox
LOC106673167	NW_019392805.1	45449	47354	superoxide dismutase [Mn] mitochondrial-like	Red/Ox
LOC106672941	NW_019392708.1	1425364	1431085	superoxide dismutase [Mn] mitochondrial-like	Red/Ox
LOC106661134	NW_019392643.1	1893718	1906435	thioredoxin reductase 1 cytoplasmic-like	Red/Ox
LOC106672011	NW_019392751.1	195105	198796	thioredoxin reductase 1 cytoplasmic-like	Red/Ox
LOC106664858	NW_019392687.1	211896	221426	thioredoxin reductase 1 mitochondrial	Red/Ox
LOC106670865	NW_019392777.1	444448	450100	thioredoxin reductase SEP1-like	Red/Ox
LOC106665860	NW_019392823.1	561145	563013	thioredoxin reductase-like selenoprotein T homolog CG3887	Red/Ox
LOC112126170	NW_019392656.1	2844561	2845475	2-hydroxyacylphosphingosine 1-beta-galactosyltransferase-like	UDPGT
LOC106673665	NW_019392677.1	3587699	3601102	beta-13-galactosyltransferase 1-like	UDPGT
LOC106671051	NW_019393055.1	291110	292588	beta-13-galactosyltransferase 5	UDPGT
LOC106670458	NW_019392645.1	1253976	1266529	beta-13-galactosyltransferase 5-like	UDPGT
LOC106673839	NW_019392677.1	1481825	1589325	beta-13-galactosyltransferase 5	UDPGT
LOC106670905	NW_019392777.1	131183	132566	beta-13-galactosyltransferase 6-like	UDPGT
LOC106665695	NW_019392748.1	1422092	1423512	beta-13-galactosyltransferase bm	UDPGT
LOC106661293	NW_019392641.1	2358243	2389688	beta-13-galactosyltransferase	UDPGT
LOC112128463	NW_019392652.1	1861433	1863900	beta-14-galactosyltransferase 1-like	UDPGT
LOC10667087	NW_019392678.1	1046741	1089426	beta-14-galactosyltransferase 4-like	UDPGT
LOC106661089	NW_019392686.1	1083927	1089202	beta-14-galactosyltransferase 1-like	UDPGT
LOC106673052	NW_019392713.1	913	9339	beta-14-galactosyltransferase 1-like	UDPGT
LOC10667582	NW_019392795.1	1045338	1159812	beta-14-galactosyltransferase 1	UDPGT
LOC106673158	NW_019392864.1	954434	975800	beta-14-galactosyltransferase 1	UDPGT
LOC106664216	NW_019392897.1	221340	231007	ceramide glucosyltransferase	UDPGT
LOC106672624	NW_019392776.1	1378977	1388148	dolichyl pyrophosphate Man9GlcNAc2 alpha-13-glucosyltransferase	UDPGT
LOC106669303	NW_019392746.1	2402639	2409097	dolichyl-diphosphooligosaccharide-protein glycosyltransferase 48 kDa subunit	UDPGT
LOC106665665	NW_019392737.1	863688	865714	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1	UDPGT
LOC106660789	NW_019392666.1	1354383	1361119	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 2	UDPGT
LOC106669900	NW_019392666.1	1353854	1358483	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 4	UDPGT
LOC106665335	NW_019392733.1	1252922	1255661	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit DAD1	UDPGT
LOC106664591	NW_019392699.1	1401595	1407248	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit STT3A	UDPGT
LOC106669873	NW_019392760.1	776208	786231	dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit STT3B	UDPGT
LOC106672378	NW_019392817.1	637809	642900	dolichyl-phosphate beta-glucosyltransferase	UDPGT
LOC106661495	NW_019392679.1	1708165	1764544	fringe glycosyltransferase	UDPGT
LOC106661064	NW_019392734.1	464435	477826	galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 2-like	UDPGT
LOC106668840	NW_019392784.1	23115	31245	glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1-like	UDPGT
LOC106662171	NW_019392688.1	3728607	3743437	glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1-like	UDPGT
LOC106661464	NW_019392713.1	1372312	1498363	glycosyltransferase 25 family member	UDPGT
LOC106674402	NW_019392666.1	1290179	1295488	glycosyltransferase-like domain-containing protein 1	UDPGT
LOC106661801	NW_019392635.1	1935966	1940046	lactosylceramide 4-alpha-galactosyltransferase-like	UDPGT
LOC106668751	NW_019392662.1	350058	351439	lactosylceramide 4-alpha-galactosyltransferase-like	UDPGT
LOC106671962	NW_019392723.1	976837	986126	O-glucosyltransferase rumi homolog	UDPGT
LOC106664743	NW_019392807.1	367903	383963	probable dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-13-glucosyltransferase	UDPGT
LOC106666658	NW_019392706.1	1088646	1092399	putative Dol-P-Glc:Glc(2)Man(9)GlcNAc(2)-PP-Dol alpha-12-glucosyltransferase	UDPGT
LOC106669691	NW_019392656.1	1296114	1309595	UDP-glucose:glycoprotein glucosyltransferase	UDPGT
LOC106669680	NW_019392656.1	2850533	2871761	UDP-glucuronosyltransferase 2A3-like	UDPGT
LOC106667555	NW_019392795.1	116513	121174	UDP-glucuronosyltransferase 2B4-like	UDPGT
LOC106662520	NW_019392683.1	2480196	2484164	xylosylprotein 4-beta-galactosyltransferase	UDPGT

3

Réponse transcriptomique à l'exposition aux pyréthrinoïdes

ARTICLE 2 : Réponse transcriptomique au traitement pyréthrinoïdes chez des souches de punaises de lit résistantes et sensibles.	118
---	-----

Avant-propos

Dans le chapitre précédent, une analyse à l'échelle du génome entier de quatre souches de punaises de lit aux statuts de résistance contrastés a été menée. Cette analyse a permis d'identifier des variants ponctuels (SNPs) et structuraux (duplications, inversions) fortement différenciés entre les souches de punaises de lit résistantes et sensibles aux pyréthriinoïdes. Une région génomique a été détectée comme combinant une forte densité de SNPs et de variants structuraux corrélés au phénotype de résistance, ainsi qu'un enrichissement en gènes potentiellement impliqués dans cette résistance, notamment des enzymes de détoxification. Cette région, appelée superlocus, correspond à un QTL majeur impliqué dans la résistance aux pyréthriinoïdes, détecté par une étude précédente fondatrice de nos travaux.

Ce nouveau chapitre a donc vocation à creuser les conséquences potentielles de ces variants génomiques d'un point de vue transcriptomique, en analysant des données obtenues sur deux des souches précédemment étudiées - London Lab et London Field. Ces analyses transcriptomiques permettent la mise en évidence de potentiels différentiels d'expression, pouvant engendrer par exemple une détoxification accrue des insecticides, comme cela a été observé chez de nombreuses espèces d'insectes résistantes.

Dans un premier temps, le niveau d'expression a été comparé en fonction de deux conditions : la souche d'appartenance, et l'exposition à un insecticide. Les analyses génomiques préalables ont permis de mettre en évidence que les deux souches étudiées ont un fond génétique très similaire, ce qui représente un atout majeur de notre comparaison, la plupart des études réalisées ayant contrasté des souches extrêmement différenciés d'un point de vue géographique. Lors de leur comparaison, seuls 20 gènes ont été détectés comme sur-exprimés chez London Field, dont six protéines cuticulaires potentiellement impliquées dans la résistance. Les gènes dont l'expression est modifiée suite à une exposition aux pyréthriinoïdes sont bien plus nombreux. Ainsi, 375 gènes ont été détectés comme sur-exprimés chez les individus ayant survécu à l'exposition à la deltaméthrine par rapport aux non-exposés. Parmi ces gènes, deux catégories de gènes de résistance se sont détachées comme étant significativement enrichies, dont les transporteurs ABC. Quatre des six transporteurs ABC ainsi détectés étaient également co-localisés sur le même scaffold. Un seul gène, codant pour une *leucine-rich repeat protein* (LRR) *soc-2*, a enfin été identifié comme présentant une interaction entre les conditions souche et traitement, et également comme

sur-exprimé chez London Field et chez les survivants traités.

Dans un second temps, nous avons corrélié les variants candidats préalablement détectés au niveau du génome avec les gènes candidats identifiés par l'analyse transcriptomique entre souches.

Deux des protéines cuticulaires ainsi détectées comme étant sur-exprimées chez London Field font partie d'une duplication inversée particulièrement différenciée (F_{ST} dans le top 10%). Quatre protéines cuticulaires, dont une sur la duplication décrite précédemment, ont également un ou plusieurs SNPs fortement différenciés à proximité (avec une fréquence de l'allèle alternatif plus élevée chez London Field). Ces loci sont donc des candidats pouvant moduler la réponse transcriptionnelle du gène. De même, le gène codant pour la LRR *soc-2* présente des SNPs fortement différenciés dans sa région promotrice ainsi que dans les séquences intergéniques avoisinantes.

En revanche, nous n'avons pas identifié de signal transcriptomique particulier au niveau du superlocus précédemment détecté. Différentes hypothèses permettant d'expliquer ce résultat sont discutées dans ce manuscrit.

ORIGINAL ARTICLE

Journal Section

Transcriptomic response to pyrethroid treatment in resistant and susceptible bed bug strains

Chloé Haberkorn^{1,2,*} | Zaïnab Belgaidi¹ | Romain Lasseur² | Fabrice Vavre¹ | Julien Varaldi¹

¹Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, UMR 5558, Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France

²IZInnovation, 13 Rue des Émeraudes, 69006 Lyon, France

Correspondence

Chloé Haberkorn
Email: chloehbk@gmail.com

Funding information

This work was supported by a CIFRE grant N° 2019/0800, funded by Association Nationale de la Recherche et de la Technologie, from the collaboration between the Laboratoire de Biométrie et Biologie Évolutive and IZInnovation SAS ; and the Scientific Breakthrough Project Micro-be-have (Microbial impact on insect behavior) of Université de Lyon, within the programme 'Investissements d'Avenir' (ANR-11-IDEX-0007; ANR-16-IDEX-0005)

The common bed bug, *Cimex lectularius*, is one of the primary human parasite. The world-wide resurgence of this pest is mainly due to globalization, and the spread of insecticide resistance. Although we have a relatively clear picture of the genes that might be involved, little is know about their genomic basis. In the present study, we sequenced the transcriptomes of two Londonian strains, one pyrethroid susceptible and one resistant, for insecticide survivors and unexposed individuals in both. In sharp distinction to previous studies, differences in expression between resistant and susceptible strains seemed to be related to a small number of genes, mainly coding for cuticular proteins, through duplications or SNPs in non-coding regions. In contrast, insecticide exposure induces the expression of a multitude of genes, such as ABC transporters. We also raised the need to explore post-transcriptional response, as highly differentiated regions detected in previous genomic analysis were not associated with transcriptional signal in our data.

KEYWORDS

Cimex lectularius; insecticide resistance; RNAseq; transcripts

1 | INTRODUCTION

Cimex lectularius, also known as the common bed bug, is an obligate parasite that mostly feed on humans. Physical and psychological disorders induced in their human preys can range from allergic reactions (Alexander, 1984), to psychosis and paranoia (Goddard and Deshazo, 2009). The current resurgence of this species, starting late 90s (Potter,

2011), seems to be due to globalization, together with a widespread second-hand market (Doggett et al., 2004) and growing insecticide resistance (Davies et al., 2012). Pyrethroids are the main insecticides used to fight this pest and are therefore suspected to be the main pressure selecting resistance (Romero et al., 2007).

Pyrethroid resistance in *Cimex lectularius* is thought to be due to several mechanisms. Cuticular resistance can first impair insecticide penetration in the body. Studies based on RT-qPCR assays showed that cuticular protein genes, such as chitin synthase (CHS) or cuticle protein, had higher transcript levels in resistant bed bug strains (Mamidala et al., 2012; Koganemaru et al., 2013). Following penetration of the cuticular barrier, detoxification metabolism may be mobilized to degrade or excrete the insecticide. Indeed, an increased activity of several enzymes such as cytochromes P450s (Romero et al., 2009) and esterases (Lilly et al., 2016a) has been observed in pyrethroid resistant strains. In addition, administration of piperonyl butoxide (PBO), a primary inhibitor of some cytochrome P450 monooxygenases, was associated with a significant decrease in resistance, further suggesting their involvement in resistance (Romero et al., 2009). Increased expressions of glutathione-S-transferases (GSTs) were also spotted in resistant juvenile bed bugs (Mamidala et al., 2011). Finally, mutations can occur in the nervous system, and more specifically in the voltage-dependent sodium channels (VGSC), targeted by pyrethroid (Dang et al., 2014, 2015; Akhondi et al., 2015; Balvín and Booth, 2018). A single non-synonymous mutation can alter VGSC conformation and hinder pyrethroid binding, thus conferring knock-down resistance (*kdr* mutations). The *kdr* mutant L925I (leucine 925 to isoleucine) has been identified in most studies (88% of the 117 strains tested in USA in Zhu et al. 2010, 100% of individuals in France in Durand et al. 2012), although some bed bugs had additional V419L (valine 419 to leucine, 40.9% in Zhu et al. 2010) or, more rarely, V419L alone (2.7% in Zhu et al. 2010).

Because insecticide resistance can involve multiple mechanisms, each of them relying on several genes, it is critical to have, as far as possible, a comprehensive view of them. This holistic view may be obtained on the DNA ground through whole genome sequencing. Since 2016, thanks to Benoit et al. (2016) & Rosenfeld et al. (2016), we now have a reference genome for *C. lectularius*. Based on the knowledge acquired on other insect systems, this priceless genomic resource allowed to delineate the set of genes possibly involved in insecticide resistance in bed bugs.

To provide an insight into the underlying genomic sequences, we recently performed a DNA pool-sequencing study which contrasted resistant and susceptible strains of *C. lectularius* (Haberhorn et al., 2023). A superlocus showing high genetic differentiation between strains was identified, corresponding to the large genomic region detected by Fountain et al. (2016), enriched in outlier single nucleotide polymorphisms (SNPs) and structural variants (SVs) putatively involved in resistance. However, the functional validation of those candidates remains to be investigated. Transcriptomic approaches are particularly relevant for genes involved in detoxification, since their efficiency is often related to their level of expression (Li et al., 2007). A few studies, all published before the first genome was released, have addressed this question. In a first attempt to do so, several studies released between 2011 and 2013 produced whole transcriptomic data for strains of *C. lectularius* with differential pyrethroid resistance status (Adelman et al., 2011; Bai et al., 2011; Zhu et al., 2013). However, since they all used 454 technology, quantification of each transcript was not possible (because of the relatively low throughput of this now ancient technology). Instead, a small set of transcripts/genes was further investigated using qRT-PCR. Since then, to our knowledge, only one resistance study using RNA-seq was conducted, leading to the identification of a huge quantity of transcripts differentially expressed between resistant/susceptible strains (15540 out of 51492 expressed sequence tags (ESTs), Mamidala et al. 2012). Although these studies provided significant insights into mechanisms possibly involved in resistance in bed bugs, they all suffered from several caveats. First, all the comparative studies mentioned so far were conducted on strains having different genetic backgrounds. Indeed, most of them used the Harlan strain (sampled in Fort Dix, NJ in 1973) as the susceptible reference strain, and compared it to strains sampled from distant locations (800km away for Columbus, OH in Mamidala et al. 2012 and Bai et al. 2011, and 440km for Richmond, VA in Adelman et al. 2011). Consequently,

whether the genes or transcripts identified in these studies are differentially expressed because of selection for resistance, selection by other selective factors or even non adaptive genetic differentiation (genetic drift) is unclear. This may explain why so many transcripts were detected as differentially expressed in Mamidala et al. (2012). Additionally, these studies focused on constitutive differences between strains, since no exposure to insecticide was performed. This may limit our power to detect important genes involved in resistance, since resistance genes are often induced upon insecticide exposure (Guedes et al., 2017; Poupardin et al., 2008).

In the present study, we take advantage of the recently published reference genome (Benoit et al., 2016; Rosenfeld et al., 2016) and analyzed the whole transcriptome of two bed bugs strains differing in resistance to pyrethroids. The two strains, that were either exposed or not to pyrethroids, were genetically very similar, since the overall index of differentiation (F_{ST}) was only 0.018 (Haberhorn et al., 2023). Expression levels were compared in order to identify (i) genes showing constitutive differences between strains, (ii) genes whose expression is altered by pyrethroid exposure, and (iii) genes showing interaction between both strain and treatment conditions. We then linked these results with previous whole genome analysis on the very same strains, which identified outlier SNPs and structural variants putatively involved in pyrethroid resistance (Haberhorn et al., 2023).

2 | MATERIALS AND METHODS

2.1 | Insects

The two strains used in this study were provided by Cimex Store Ltd (Chepstow, United Kingdom). London Lab (collected in Great Britain) was susceptible to insecticides, as collected before the massive use of insecticide and raised in laboratories for over than 40 years. London Field (collected in 2008 in Great Britain) however was moderately resistant to pyrethroids.

Insects were kept isolated before imaginal moult in 24-wells Petri dishes containing accordion-folded blotting papers, serving as harborage. Bed bugs were maintained at 25°C, 40% relative humidity (RH), and a photoperiod of 12:12h. Since males perform the so-called traumatic insemination to copulate, which can be costly for their partners (Stutt and Siva-Jothy, 2001), insecticide exposures were conducted only on 7-day-old unfed virgin females.

2.2 | Topical Assay

Insecticidal assays were carried out with deltamethrin (98% purity, Cluzeau, Sainte-Foy-La-Grande, France), a pyrethroid, as it remains the most used insecticide family recently for bed bug control. Four replicates of 8 insects were performed both for control and 1 ng/ μ L insecticide exposure. Bed bugs were previously immobilized by placing them in a Petri dish on ice for 5 minutes. Topical applications were then made onto the ventral surface of the thorax, between the coxae, with a 50- μ L glass syringe attached to a repeating dispenser (Hamilton Co., Reno, NV). Treated insects were exposed to 1 μ L of insecticide powder diluted in acetone, whereas control insects received 1 μ L of acetone only. Insects were treated with a discriminating dose of 1 ng/ μ L of deltamethrin, as assumed to reflect the LD₅₀ of Cimex-Store resistant strains (Haberhorn et al., 2023).

Mortality was assessed after 24 h by flipping each insect on dorsal side with a featherweight forceps, to see if it was able to reverse on the ventral side (alive) or if its movements were not coordinated enough to do so (moribund, considered as dead). A fisher test was used to assess the significance of mortality difference between pairs of strains.

2.3 | RNA extraction and sequencing

For each replicate, three individuals were randomly sampled out of survivors for insecticide-exposed and control conditions, and immediately flash-frozen in liquid nitrogen and stored at -80°C. Extraction were therefore realized on 3 individual survivors per replicates (n=4) per condition (treated or untreated) for control and survivors bugs, resulting in 24 bed bugs per strain. Bed bugs were extracted individually, using RNAeasy Mini Kit (Qiagen, Hilden, Germany) with Turbo DNase treatment on column (ThermoFisher, Waltham MASS, USA). RNA concentration was measured using Nanodrop, and Qubit with RNA HS Kit (Agilent, Santa Clara CA, USA).

Both reverse-transcription and sequencing were performed by MacroGen Europe (Amsterdam, Netherlands). Libraries were prepared using TruSeq stranded mRNA kit (Illumina, San Diego CA, USA), and sequencing was performed on Illumina Novaseq6000 with 40 million reads 2*100PE. The raw data have been submitted to the Sequence Read Archive (SRA) database of NCBI under BioProject PRJNA832557.

2.4 | cDNA processing

The adapters had already been removed by MacroGen, and quality was checked with Fastqc. STAR v2.7.3a workflow was then applied, first to generate genome indexing files, using reference genome *Clec_2.1* in FASTA and associated GTF annotation file (GCF_000648675.2), and then to map reads to this genome. Parameters for mapping were as follow: "sjdbOverhang" of 100 (read size - 1), a computed optimal "genomeSAindexNbases" of 13, "genomeChrBinNbits" of 18 and "quantMode" on "GeneCounts". This last option allowed STAR to count the number of reads per gene while mapping, which required the GTF annotation provided above. Read counts per gene were then outputted into the file "ReadsPerGene.out.tab". Out of the 96 samples sequences (24 samples per strain), the mean of uniquely mapped reads was 84,09%.

2.5 | Mapping scaffolds on linkage groups

Fountain et al. (2016) constructed a genetic map for *C. lectularius*, using RAD-seq QTL protocol. Nevertheless, the genome assembly of *C. lectularius* has been updated in the meantime (v1.0 to 2.1). In order to improve the genetic map produced, we decided to use Fountain's RAD-seq data and R scripts to generate a genetic map based on the latest genome assembly. This allowed us to obtain putative autosomes, called linkage groups (LG).

For assembling LG, we re-aligned RAD sequencing results available on Dryad (<http://dx.doi.org/10.5061/dryad.d4r50>) with GSNAP version 2020-06-01 on the RefSeq genome version 2.1, rather than version 1.0 used by Fountain. By following the same pipeline, we were able to keep 338 out of 441 RAD markers, distributed over 14 LG. Markers detected on scaffolds that appeared on more than one single LG were removed. Hence, out of 1573 nuclear scaffolds, only 151 scaffolds were kept. While Fountain managed to map >65% of the genome on LG, we only obtained 46% (i.e. 234.7 Mb out of 510 Mb). One explanation could reside in the fragmentation of the new genome version. Indeed, the latest version of the genome was split into more scaffolds (*Clec_1.0*: 1402 scaffolds, *Clec_2.1*: 1574 scaffolds - see https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Cimex_lectularius/101/). RAD markers therefore potentially fall into smaller scaffolds, which can lead to this loss of information.

2.6 | Differential Expression analysis on London strains

We focused only on London Lab and London Field samples (48 out of the 96 sequenced). Inside the "ReadsPerGene.out.tab" file, different counts are provided depending on data strandedness. As having stranded data, we used the 4th column, from whose sum was greater than in the 3rd column. Column 3 is therefore giving the count of anti-sense reads. Genes annotated as transfer RNA (tRNA) were removed from the global gene pool, before following the R package DESeq2 v1.34.0 pipeline (Love et al., 2014). A matrix with reads count by gene for each sample was first built, together with a table containing data conditions (untreated or treated, lab or field strain). Factors were reordered to have "lab" and "untreated" as reference levels. A model with interaction between conditions treatment and strain was built. Before running the DESeq function, all lines (i.e. gene) with less than 10 reads in total (for all samples) were removed, leading to a total of 12187 genes out of 13208 (without tRNA).

Genes were considered as significantly over or under-expressed when showing an adjusted p-value below 5% ($p_{adj} < 0.05$), and a fold change in expression above 1.5, i.e. Log_2 fold change (LFC) > 0.58 (up) or < -0.58 (down). The 1.5 fold change threshold corresponded to at least half the individuals of a condition carrying a gene duplication.

2.7 | Functional analysis

In order to test whether some relevant biological functions were enriched in a set of candidate genes, we defined ten categories of genes potentially involved in resistance, similarly to Faucon et al. (2015). Six of them corresponded to metabolic resistance : "Binding/Sequestration" (6 genes) , "GST" (14), "CCE" (39), "UDPGT" (39), "ABC transporter/MRP" (55), and "P450" (56). Other categories were separated as: "Other detox" (64), "Cuticle" (113), "Insecticide target and nervous system" (31), and "Redox homeostasis" (14 genes). Details about those 431 genes are provided in Haberhorn et al. 2023.

3 | RESULTS

3.1 | Assessment of resistance or susceptible phenotype

London Lab and London Field strains were first exposed to a discriminating dose of deltamethrin (1 ng per insect), in order to assess their pyrethroid resistance or susceptible status. As expected, the mean mortality was significantly higher for London Lab compared to London Field (59.4% versus 31.3%, $n=24$ for each strain, $p\text{-value}=0.044$).

3.2 | Principal component analysis did not reveal clear pattern between samples

RNA seq was performed in order to compare the transcriptome of the two strains in the absence of insecticide and after insecticide exposure. 12 replicates were performed for each condition. The read counts table obtained after RNA-seq sequencing was first analyzed by Principal Component Analysis (PCA) (Figure 1). The sum of these two axes explained 45% of the variance (53% by adding PC3). No clear pattern was observed neither between strains, nor between treated and untreated samples. This suggests that the 48 transcriptomes were relatively similar among all samples.

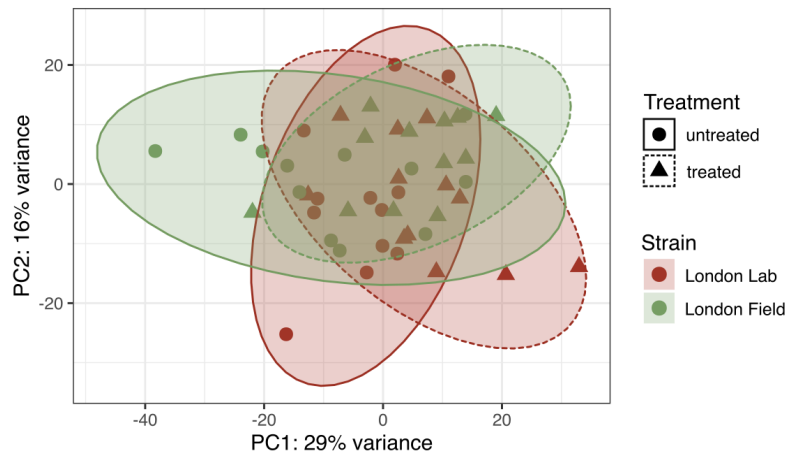


FIGURE 1 Projection of London Lab and London Field populations on the top two principal components using PCA. Read counts obtained from DESeq were used, and transformed with vst. Individuals are separated within a population between survivors of insecticide treatment and untreated. One point represents the top 500 most variable genes of a sample.

3.3 | Analysis of differential gene expression

Variance in read counts was decomposed using a model including a strain effect (London Lab and London Field), a treatment effect (treated and untreated) and their interaction. Genes were considered as differentially expressed when the fold change exceeded 1.5 (in absolute value) and when $p_{adj} < 0.05$.

3.3.1 | Constitutive differences between resistant and susceptible strains

The "strain" effect was first analyzed in order to identify genes showing constitutive differences in expression between the two strains. Out of 12187 genes, 20 genes were detected as up-regulated and 4 as down-regulated in London Field compared to London Lab (Figure 2).

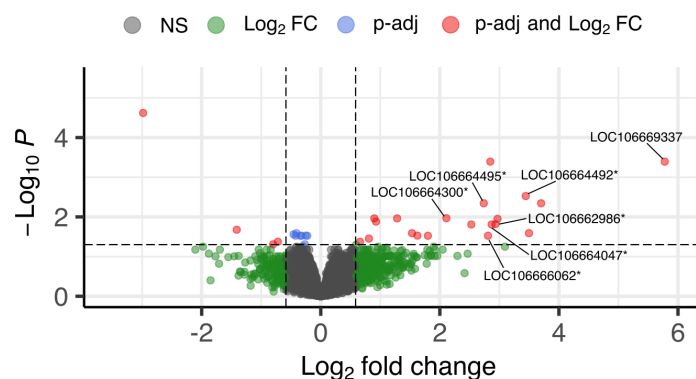


FIGURE 2 Volcano plot using enhanced volcano on strain condition ($n_{genes} = 12187$, $p_{adj} < 0.05$, $LFC > \text{or} < |0.58|$). Genes coding for cuticular proteins are labelled with a star.

No genes potentially involved in insecticide resistance were detected among the down-regulated genes. On the

contrary, among the genes overexpressed in London Field, six were putatively involved in insecticide resistance which represents an highly significant over-representation for this category (p -value = $5.13e-10$). Strikingly, all six genes were coding for several cuticular proteins (Figure 3, Supplementary table 1), which also represents an over-representation of this specific category (p -value < $2.2e-16$). Among over-expressed genes, the one showing simultaneously the highest LFC and lowest p_{adj} encoded a putative leucine-rich repeat-containing protein DDB_G0290503 (106669337). To date, this gene is not known to be involved in insecticide resistance.

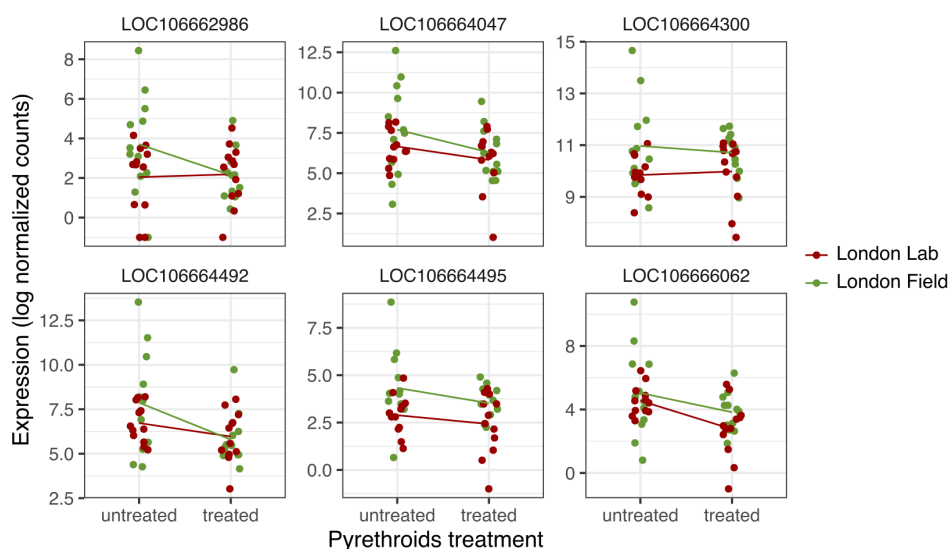


FIGURE 3 Cuticular proteins detected as significantly over-expressed in London Field. Expression is given in \log_2 of normalized count.

3.3.2 | Insecticide-altered genes

The "treatment" effect was then analyzed, i.e. the difference between untreated and treated survivors. Genes were similarly filtered on LFC and adjusted p -values, leading to respectively 375 genes significantly up-regulated and 388 down-regulated after treatment (Figure 4).

Among the genes over-expressed in treated survivors, 20 were putatively involved in insecticide resistance (Figure 5, Supplementary table 2), which constitutes a significant enrichment (p -value = 0.02). Two out of the six resistance categories among over-expressed genes were enriched, namely ABC transporters (p_{adj} = $4.06E-03$), and other detox (p_{adj} = $8.19E-03$).

Furthermore, most of the induced genes identified in a resistance category were involved in the detoxification metabolism: six ABC transporters, six various other detox (including transcription factor cap'n'collar and MAF, and three sulfotransferases), two GST and two P450. Three acetylcholinesterase-like and one cuticular protein were also detected. All acetylcholinesterase-like genes detected (labelled with a star on fig. 4) were in the top 15 of genes having the highest LFC. The up-regulated gene showing the highest LFC encoded a peroxidase-like protein (106671105), whereas the up-regulated gene showing the lowest p_{adj} encoded an elongation of very long chain fatty acids protein 7-like (106664652).

Of the down-regulated genes, 15 were detected in an insecticide resistance category (Supplementary table 3), which is not more than what is expected under the null hypothesis of no association between resistance and down-

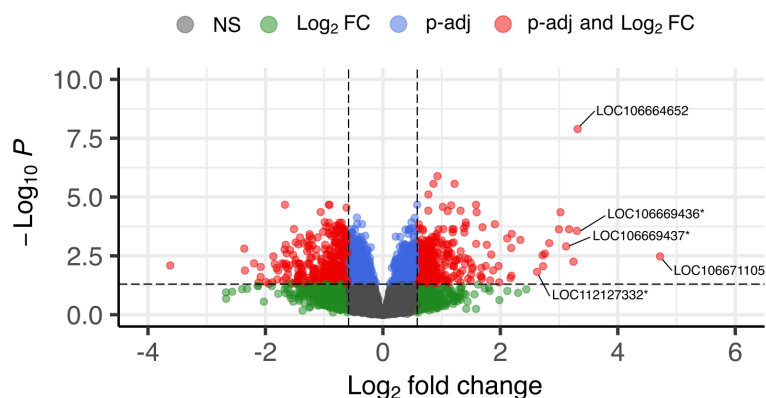


FIGURE 4 Volcano plot using enhanced volcano on treatment condition ($n_{genes} = 12187$, $p_{adj} < 0.05$, $LFC > \text{or} < |0.58|$). Genes coding for acetylcholinesterase-like proteins are labelled with a star.

regulation ($p\text{-value} = 0.30$). Similarly, no significant enrichment in any resistance gene category was detected (all $p_{adj} > 0.05$). Among the 15 genes identified, seven were coding for cuticular proteins, two for ABC transporters (both on LG 2), two for insecticide target and nervous system (one sodium channel protein Nach and one acetylcholinesterase-like), one from other detox category (sulfotransferase), two P450, and one Red/Ox gene (superoxide dismutase [Cu-Zn]-like). The down-regulated gene showing both the lowest LFC and lowest p_{adj} was an uncharacterized protein (106673864).

3.3.3 | Interaction between treatment and strain

Only one gene showed a significant interaction between treatment and strain: a leucine-rich repeat protein soc-2 (106661213, Figure 6). This gene had opposite expression patterns in the two strains upon insecticide exposure. Leucine rich repeat (LRR) motif were detected in many immune receptors of animals, as in *Homo* and *Drosophila* (Dushay and Eldon, 1998).

3.4 | Crossing expression results with genomic data

A previous population genomic study was conducted on the same two strains, although on different individuals (Haberhorn et al., 2023). This work led to the identification of a set of SNPs and structural variants (inversions, duplications) showing extreme genetic differentiation between strains with contrasted resistance status. As such, these loci may delineate genomic regions involved in the resistance phenotype. One of the purposes of the present study was to establish possible connections between those genomic variants and genes detected in the present transcriptomic study as differentially expressed (DE). Genes identified by both analysis are expected to be good candidates for being involved in the phenotypic differences observed between strains.

3.4.1 | Distribution of DE genes

Our previous genomic analysis revealed a clustering of candidate SNPs in some genomic regions, and particularly in 3 adjacent scaffolds constituting a 6 Mb region. This "superlocus" was furthermore significantly enriched in putative resistance genes, and in particular for *GSTs* and *P450s*, suggesting a clustering of those detoxification genes involved in insecticide resistance. We first tested whether the significantly DE genes were distributed evenly along the genome,

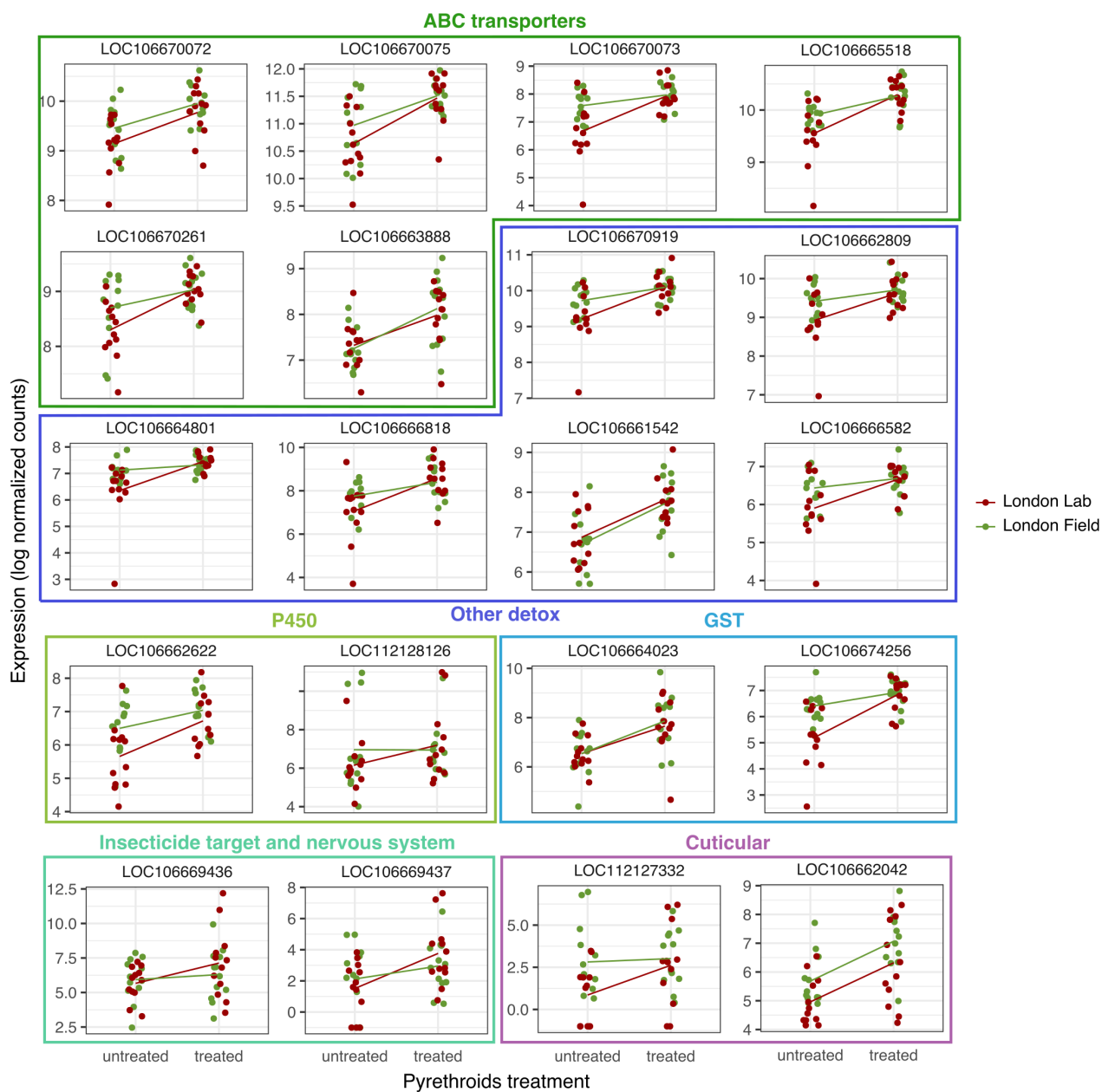


FIGURE 5 Resistance genes by category detected as significantly over-expressed in treated survivors. Expression is given in log₂ of normalized count.

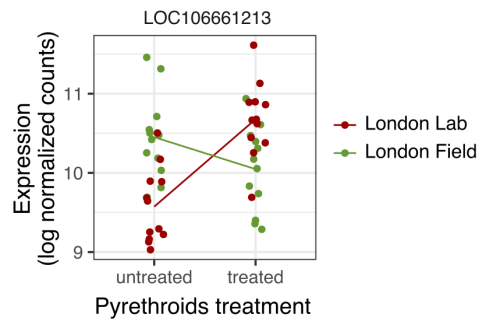


FIGURE 6 Gene significantly detected for the interaction between strain and treatment conditions. Expression is given in log₂ of normalized count.

at the chromosome (LG) or scaffold level. No pattern of enrichment was observed for up or down-regulated genes, at any scale (all $p_{adj} > 0.05$). Hence, contrary to the SNPs that were clustered in some genomic locations, differentially-expressed genes for treatment or strain conditions were scattered along the genome. Nevertheless, six significantly DE genes were identified in the 6Mb superlocus, all of them being induced upon insecticide exposure. One of them was putatively involved in resistance, namely a GST on scaffold NW_019392763.1 (106664023, Supplementary table 2).

3.4.2 | Genomic differentiation of DE genes

F_{ST} mean values were computed by gene as a proxy of genetic differentiation, and compared between differentially-expressed genes and the global pool. We hypothesized that F_{ST} values should be higher for differentially-expressed genes, since both the genomic and the transcriptomic approaches should identify loci possibly involved in insecticide resistance. Although no link was observed between F_{ST} value and differential expression for the strain condition (p-value = 0.926), there was a marginally significant effect for the treatment condition (p-value = 0.059): F_{ST} values were slightly higher for DE genes upon insecticide exposure compared to non-DE genes (F_{ST} DE = 0.01356 > F_{ST} non-DE = 0.00826). This suggests that DE genes upon insecticide exposure are located in genomic regions that are more differentiated than the rest of the genome. However, the same test computed at the level of each individual LG did not reveal any pattern.

3.4.3 | SNP variants associated with differences in expression between strains

In addition to this genome-wide approach, we tried to identify individual genes that pop up from both the genomic and transcriptomic analysis. Constitutive transcriptional differences between strains are expected to be the consequence of either SNP, or structural variation. Thus, we first searched for such variants.

Outlier SNPs obtained from our genomic analysis were defined based on the intersection of three criterion : (i) the differentiation between the two strains originating from London should be high (top 5% F_{ST} s), (ii) alleles should be in a derived state, if one considers that the alleles of the reference genome (the insecticide susceptible Harlan strain) are ancestral (at least for loci involved in insecticide resistance), (iii) (derived) alleles should be in higher frequency in London Field than in London Lab. Nevertheless, we draw the reader's attention to the fact that these criteria are less stringent than those used by Haberkorn et al. (2023) to define outlier SNPs, since we focus here on the two London

strains only, whereas two additional strains were included in this previous study. With these criteria, 100,941 outlier SNPs were detected in London strains, versus 576 in Haberhorn et al. (2023).

Out of the 24 genes showing constitutive differences between LL and LF, 12 carried at least one outlier SNP, including four out of the six cuticular protein detected. Two cuticular proteins (106662986 and 106664047) carried only outlier SNPs in intergenic regions, whereas the others also had SNPs in promoter (106664492) or intronic region (106666062) (Table 1).

We then studied the only gene showing interaction between strain and treatment, i.e. the leucine-rich repeat protein soc-2. This gene was associated with five outlier SNPs. Two of them were localized in intergenic regions, one in a promoter and two in introns (Table 1).

Scaffold	Position	Gene ID	Dist.	Annotation	Location	R>A	F_{altLL}	F_{altLF}	pF_{ST}
NW_019392649.1	1018810	106664492	17	larval cut. A3A	P Inter	A>G	0.22	0.64	0.044
	1019770		977		P Inter	C>T	0.08	0.82	4.36E-04
NW_019392685.1	759447	106662986	1832	cuticle prot. 12.5	Inter	T>C	0.39	0.89	0.017
NW_019392920.1	61585	106664047	15825	larval cut. A1A	Inter	A>G	0.45	0.88	0.032
	62751		14659		Inter	G>A	0.73	1	0.027
	67008		10402		Inter	A>G	0.64	1	0.037
	68629		8781		Inter	A>T	0.46	0.87	0.037
NW_019392993.1	2430659	106666062	7112	cuticle protein	Inter	A>C	0.10	0.69	5.91E-03
	2430749		7022		Inter	C>T	0.15	0.63	0.017
	2430766		7005		Inter	T>G	0.17	0.59	0.039
	2431463		6308		Inter	C>T	0.24	0.70	0.032
	2431466		6305		Inter	C>T	0.19	0.71	0.015
	2431520		6251		Inter	A>G	0.24	0.75	0.016
	2431528		6243		Inter	G>A	0.27	0.76	0.021
	2434684		3087		Inter	A>T	0.18	0.67	0.028
	2434689		3082		Inter	T>C	0.20	0.70	0.025
	2434733		3038		Inter	T>G	0.27	0.80	0.017
	2435135		2636		Inter	T>C	0.27	0.73	0.037
	2438872		-		I	T>A	0.15	0.80	1.85E-03
	2438882		-		I	G>A	0.13	0.70	6.55E-03
	2445654		-		I	G>T	0.23	0.68	0.031
2452655	-	I	T>C	0.06	0.64	3.79E-03			
NW_019392653.1	1288539	106661213	1213	LRR soc-2	Inter	A>C	0.29	0.89	3.92E-03
	1288598		1154		Inter	A>G	0.38	0.81	0.042
	1291912		-		I P	C>A	0.43	0.91	0.023
	1299742		-		I	G>C	0.58	1	9.70E-03
	1302774		-		I	T>C	0.39	0.83	0.030

TABLE 1 Outlier SNPs falling in resistance genes **over**-expressed in London Field when compared to London Lab strain. SNPs were selected with the frequency of alternative allele (F_{alt}) higher in London Field (LF) than London Lab (LL) and differentiation index F_{ST} value in top 5% quantile (i.e. empirical $pF_{ST} < 0.05$). The position of each SNP was given, together with distance to nearest gene and its gene ID, gene annotation, reference to alternative nucleotide (R>A) and the location of the SNP (promoter=P, intergenic=Inter, intron=I). Whenever position were litigious due to the gap between transcripts annotation, both position were given and separated by a pipe.

3.4.4 | Structural variants associated with differences in expression between strains

Using a combination of read-pair orientation, insert size information and read depth, the genomic analysis also revealed 650 inverted duplications, 509 tandem duplications and 1,338 inversions in London Field compared to the reference genome, with higher frequencies in London Field compared to London Lab strain. These structural variants were overlapping with 4,118 genes. Here, we tested whether the constitutive differentially expressed genes identified in the present transcriptomic analysis were more likely to be located inside SVs. The rationale was that structural variants may underlie transcriptomic changes, either through gene duplications or through rearrangements in promoter sequences. Overall, this was not the case (p -value = 0.901). However, among the 20 over-expressed genes in London Field compared to London Lab, five were detected inside SVs (Supplementary table 4), including two cuticle protein A3A-like within a single inverted duplication located on LG5 (106664492 and 106664495). The three other genes were not part of any resistance category. Two genes out of the four under-expressed in LF compared to LL were part of SVs, but none of them were thought to be involved in insecticide resistance. Finally, no structural variants were detected as overlapping with the only gene detected with an interaction between treatment and strain condition, LRR soc-2.

3.4.5 | Outlier non-synonymous SNPs in DE genes for treatment condition

Genes that are similarly induced upon insecticide treatment in both susceptible and resistant strains may still underlie the resistance difference if their coding sequence differ.

We thus tested whether DE gene for treatment (but not strain) carried non-synonymous outlier SNPs, and identified a total of 13 genes (Supplementary table 5). Out of the 375 genes up-regulated, seven were carrying at least one non-synonymous special SNP, but none of them were part of an insecticide resistance category. However, we note that two of them (FAR wat and serpin B3 proteins) had three to four amino acid changes in their sequence. Interestingly, all three non-synonymous mutations found in the serpin B3 (gene ID 106664949) were fixed in London Field strain.

Similarly, among the 388 genes down-regulated, four were carrying at least one non-synonymous special SNP, although no one belonged to a putative resistance category.

4 | DISCUSSION

The global resurgence of *C. lectularius* seemed to be mainly due to the evolution of insecticide resistance (Davies et al., 2012). A few transcriptomic analysis have been conducted on bed bugs, but only by comparing constitutive expression in resistant and susceptible strains with different genetic backgrounds (Zhu et al., 2012; Mamidala et al., 2012; Koganemaru et al., 2013). Because of this difference, it is unclear to what extent the genes or transcripts identified are truly relevant for insecticide resistance, for other stress factors or even has no adaptive significance if the differentiation arose as a consequence of genetic drift. Here, we provide a comprehensive transcriptomic analysis by comparing two sympatric bed bug strains with very similar genetic backgrounds (average genome-wide $F_{ST} = 0.018$, Haberkorn et al. 2023) but different resistance phenotypes. In addition to baseline constitutive gene expression, we also used deltamethrin treatment to identify genes whose expression is altered upon exposure. We interpret our results in the light of a previous genome-wide analysis performed on the same strains, which identified genomic islands of exceptional differentiation that are expected to be enriched in adaptive loci (Haberkorn et al., 2023).

4.1 | Constitutive difference in cuticular gene expression is associated with resistance

We first compared the two strains in the absence of any treatment in order to detect constitutive difference between strains that may underlie the observed difference in resistance. Only 24 genes (out of 12187) were detected as differentially-expressed, among which 20 were over-expressed in the resistant strain. Importantly, up-regulated genes were enriched in a single resistance gene category, namely the cuticular category. Six cuticular genes were indeed detected, including three A3A larval cuticle proteins, one protein A1A and two cuticle proteins. Cuticular resistance has been observed in many pyrethroid-resistant insects, such as *Helicoverpa armigera* (Ahmad et al., 2006), *Anopheles gambiae* (Yahouédo et al., 2017) or *Culex pipiens* (Pan et al., 2009). In bed bugs also, several results suggest that cuticle plays a central role in insecticide resistance. For instance, a morphological study showed a positive correlation between cuticle thickness and resistance status of the strains (Lilly et al., 2016b). Additionally, three transcripts encoding cuticular proteins (C2, C10 and C13) were detected as over-expressed in different resistant populations of bed-bugs (Zhu et al., 2013). Other studies showed that chitin synthase or larval/pupal cuticular protein had higher transcript level in pyrethroid resistant bed bug strains (Mamidala et al., 2012; Koganemaru et al., 2013). In particular, the larval cuticle protein A3A has also been detected among up-regulated transcripts in resistant strains (see Supplementary table 6 in Mamidala et al. 2012). Interestingly, we also identified two A3A proteins as up-regulated in London Field strain that were overlapping with a putative inverted duplication (Haberhorn et al., 2023). However, this inverted duplication was found in both London Lab and London Field strains, although in higher frequency in LF. Whether the relative small difference in frequency for this inverted duplication (12.3% versus 5.7%) underlies the observed difference in expression is unclear but surely deserves additional investigations.

4.2 | Genetic variants associated with differences in expression between strains

Overall, the set of genes showing constitutive difference in expression was not particularly enriched in structural variants, as could have been expected. However, we detected outlier SNPs in non-coding regions, including regions annotated as promoters, that might be involved in the constitutive expression difference observed. These SNPs were part of the most differentiated loci detected in the whole genome analysis we conducted previously (Haberhorn et al., 2023).

Four of such mutations were detected in introns of the over-expressed cuticle protein (106666062). This same gene was also associated with 11 outlier SNPs located in a small nearby flanking region (2636 bp to 7112 bp upstream). Both flanking and intronic mutations may lead to transcriptomic changes, since both may serve as enhancers that ultimately recruit the RNA polymerase II on the promoter (Cannavò et al., 2016).

In one of the up-regulated larval cuticular A3A gene (106664492), two mutations were found in a flanking region annotated as promoter or intergenic sequence, depending on the transcript considered. In *Nilaparvata lugens*, the up-regulation of the P450 genes CYP6AY1 and CYP6ER1 were associated with resistance to imidacloprid together with several mutations in their respective promoters. By site-directed mutagenesis, the authors were able to ascribe most of variance in up-regulation to a few set of single point mutations in the promoter sequences (Liang et al., 2018). The "resistant alleles" were associated with a higher binding activity for nuclear proteins, suggesting a higher recruitment of transcription factors (Liang et al., 2018).

The unique gene showing an interaction between the strain and treatment effect encoded a leucine-rich repeat soc-2 protein. This gene was associated with an outlier SNP in a promoter region, as well as with intergenic (around 1 kb upstream) and intronic outliers.

Although *soc-2* is not directly known to be involved in insecticide resistance, it seems to play a role in nicotinic

acetylcholine receptor (nAChR) sensitivity, the target site of neonicotinoids (Gottschalk et al., 2005). High levels of neonicotinoid resistance have been observed in bed bug strains and attributed partially to enzymatic activities, without exploring the involvement of nAChR (Romero and Anderson, 2016). Although our resistant London Field strain was only tested for deltamethrin resistance, it could also have been exposed to neonicotinoids in the field. More studies should be done to detect potential SNPs in nAChRs from neonicotinoid-resistant populations, and to disentangle the putative involvement of *soc-2*.

The results we obtained from this part are in sharp contrast with previous comparative whole transcriptome studies performed in bed bugs. Indeed, in Mamidala et al. (2012), several thousands of differentially-expressed transcripts were identified in the comparison between the two strains they used (Fort Dix, NJ and Columbus, OH). In Adelman et al. (2011) and Bai et al. (2011) which both relied on 454 sequencing followed by targeted RT-qPCR analysis, differences in transcripts abundance were detected for several categories of putative resistance genes: esterases, GSTs and several P450s. In contrast with these studies, we identified only a small set of genes constitutively DE, and a single category of putative resistance genes was detected (cuticular genes). Although part of the difference may be explained by the fact that they did not have a reference genome at that time (and had thus to rely on *de novo* assembly of the transcripts which can inflate the number of "genes"), another explanation could be the lack of genetic homogeneity in the two strains they used. Indeed, in all three cases, the strains were collected in distant locations. This means that they may differ in allelic composition just because of other selective pressures experienced in the different environment they encountered (apart from insecticide) or because of genetic drift due to independent evolution. On the contrary, from our previous genetic analysis, it appeared that London Lab and London Field strains were the closest among the four strains we analyzed with an overall low genetic differentiation as measured by the F_{ST} . This indicates that these strains are highly similar in their allelic composition. Genetic difference (that should underlie the constitutive expression differences observed) are thus more likely to be the consequence of selection for the major difference they have, i.e. the difference in pyrethroid resistance. This result thus challenges the view that transcriptomic changes involving various detoxification mechanisms (esterases, GSTs, P450) has evolved in pyrethroid-resistant bed bugs populations. However, we acknowledge that different mechanisms may have evolved in different contexts, which may possibly explain part of the difference observed.

4.3 | Deltamethrin as a trigger for ABC transporters

Next, we aimed at identifying genes whose expression is modified by insecticide exposure. In total, 375 genes were jointly induced upon pyrethroid exposure in both strains. Both ABC transporter and other detox categories of resistance genes were detected as significantly enriched in induced genes. Indeed, among the 20 resistance genes up-regulated upon insecticide exposure, six were ABC transporters (Supplementary table 2).

Four ABC transporters were identified in sub-family ABCG4, all clustered on LG 5. ABCG4 transporters have been shown to be over-expressed in pyrethroid treated larvae (Epis et al., 2014) and adults (Mastrantonio et al., 2017) of *An. stephensi*. Gene silencing of ABCG4 resulted in an increased pyrethroid efficacy against *An. stephensi* larvae (Negri et al., 2019), further demonstrating the involvement of this gene in pyrethroid resistance.

4.4 | Advantageous variants in genes induced upon insecticide exposition can be further selected in resistant populations

The genes similarly induced after insecticide exposure in both strains are not expected to explain any difference in resistance phenotype between strains, unless they encode different alleles. Among the 375 genes that are similarly

induced in both resistant and susceptible strains, we identified a few outliers showing this pattern (Supplementary table 5). However, none of them were part of a resistance category. Still, some of them deserve attention.

First, a gene encoding a serpin B3 protein was associated with three non-synonymous outlier SNPs, with the putative derived allele being fixed in the resistant strain. Additionally, this gene was associated in total with 28 non-synonymous mutations (including non-outlier SNPs) in our previous genomic analysis (Haberhorn et al., 2023). This gene was part of the top 1% of genes having the highest ratio between non-synonymous and synonymous mutations ($\pi_n/\pi_s=4.67$), suggesting that it has experienced positive selection in the past. Additionally, the gene was overall highly differentiated on its whole sequence, since the differentiation index computed at the scale of the whole gene was part of the top 5% ($F_{ST}=0.15$). Overall, this strongly suggests that this gene has experienced positive selection in the recent past, leading to an increase in frequency of a putative derived allele. Although to our knowledge this gene has never been associated with pyrethroid resistance, the fact that it is also induced upon deltamethrin exposure renders its possible involvement in insecticide resistance plausible. The function of this particular gene is unknown, but serpins are involved in insects immune response, and thus thought to have a potential effect in insecticide resistance, as in *Culex pipiens* mosquitoes (Li et al., 2016).

A second gene might be of interest. This gene is encoding a fatty acyl-CoA reductase wat-like protein, and was associated with four non-synonymous outlier SNPs, in coding or intronic region depending on the transcripts. However, with a total of 12 non-synonymous and 11 synonymous mutations ($\pi_n/\pi_s = 0.46$), and an overall relatively low differentiation index ($F_{ST}=0.0042$), this gene involved in lipid synthesis seemed less compelling.

4.5 | No superlocus pattern in transcriptomic data

Because of the overall high genetic proximity between the two strains, genomic regions involved in insecticide resistance are expected to be enriched in outlier SNPs. Consistently with that hypothesis, the large 6 Mb superlocus identified as showing extreme genetic differentiation in Haberhorn et al. (2023) corresponds to the unique QTL that Fountain et al. (2016) associated with pyrethroid resistance. However, differentially-expressed genes, either between strains or between untreated and treated conditions, were dispersed across putative chromosomes (LG) and scaffolds. Furthermore, almost no differential expression was detected in the superlocus. The unique exception among putative resistance genes was a GST up-regulated upon insecticide treatment in both strains.

Ours result show that overall, the two strains had very similar response to insecticide exposure, suggesting that the genes affected by pyrethroid exposure did not differ in their regulatory sequences between strains. Based on the previous genomic analysis conducted on the same strains, this result is somehow surprising. Indeed, this analysis revealed numerous outlier SNPs, showing an extreme genetic differentiation between pyrethroid resistant and susceptible strains (Haberhorn et al., 2023). Nevertheless, in the genomic analysis, the great majority of outlier SNPs were synonymous, and localized in intergenic regions or introns. Indeed, in the original stringent analysis between four stains, only one non-synonymous outlier SNP was detected out of 576 (i.e. 0.002%). Similarly, the subset of outlier SNPs based on the comparison between London Lab and London Field highlighted 303 non-synonymous mutations out of 100,941 (0.003%). We thus expected part of these non-coding candidate SNPs to be involved in the regulation of the expression of resistance genes. However, this does not seem to be the case. This discrepancy may be explained by either (i) the fact that the SNPs identified are not associated with insecticide resistance (ii) the resistance difference between the strains is neither mediated by protein changes nor by transcriptomic changes but rather by other mechanisms. In support for hypothesis (i), we expected under selective sweep that alleles physically associated with the allele under selection (in linkage disequilibrium) would also increase in frequency in a phenomenon known as hitch-hiking. This may explain why part of the outlier loci we detected are not directly involved in insecticide

resistance. In support of hypothesis (ii), we acknowledge that resistance may be triggered by transcriptomic changes that we did not quantify in this work, i.e. non-coding RNAs changes.

A growing interest for non-coding RNAs has brought to light many examples of their involvement in insecticide resistance. A recent study demonstrated that a long non-coding RNA (lncRNA) modulated the expression of a GST detoxification enzyme by competing with micro RNA (miRNA) binding, thus mediating cyflumetofen resistance (Feng et al., 2020). In *Aphis gossypii*, RT-qPCR and RNAi studies assessed the role of lncRNAs in acetyl-CoA carboxylase regulation, the target site of spirotetramat insecticide Peng et al. (2021). MiRNAs are an other type of ncRNAs which can affect insecticide resistance by itself, through repressing gene expression, as shown in *Spodoptera frugiperda* Yang et al. (2022). We believe future studies should investigate the presence of ncRNAs in all remaining candidates enlightened by the genomic study, especially within the superlocus, and of their putative impact on post-transcriptional regulation and finally on insecticide resistance.

5 | CONCLUSION

In conclusion, most differences in expression between field-resistant and lab-susceptible pyrethroid strains were due to cuticular genes, and we provide a set of SNPs that may underlie these constitutive differences in expression. After deltamethrin treatment, a majority of detoxification genes and especially ABC transporter G4 seemed to be at play for both strains, but without evident protein changes in putative resistance genes. Nevertheless, we identified a set of 5 genes that may be involved in resistance since they are induced after insecticide exposure, and have numerous non-synonymous mutations, thus representing new candidates that should be further studied. Finally, little overlap was found between differentially expressed genes and regions with high genomic differentiation. We presented several hypothesis (hitch-hiking, non-coding RNAs) that may explain the pattern and should be tested in the future.

acknowledgements

This work was performed using the computing facilities of the CC LBBE/PRABI.

conflict of interest

The authors declare no conflicts of interest.

data availability statement

The data that support the findings of this study are openly available on SRA under BioProject PRJNA832557.

references

- Adelman, Z. N., Kilcullen, K. A., Koganemaru, R., Anderson, M. A., Anderson, T. D. and Miller, D. M. (2011) Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a single population. *PLoS ONE*, **6**, 1–9.
- Ahmad, M., Denholm, I. and Bromilow, R. H. (2006) Delayed cuticular penetration and enhanced metabolism of deltamethrin in pyrethroid-resistant strains of *Helicoverpa armigera* from China and Pakistan. *Pest management science*, **62**, 805–810.
- Akhoundi, M., Kengne, P., Cannet, A., Brengues, C., Berenger, J. M., Izri, A., Marty, P., Simard, F., Fontenille, D. and Delaunay, P.

- (2015) Spatial genetic structure and restricted gene flow in bed bugs (*Cimex lectularius*) populations in France. *Infection, Genetics and Evolution*, **34**, 236–243. URL: <http://dx.doi.org/10.1016/j.meegid.2015.06.028>.
- Alexander, J. O. (1984) Infestation by Hemiptera. In *Arthropods and Human Skin*, 57–74. London: Springer London. URL: https://doi.org/10.1007/978-1-4471-1356-0_6.
- Bai, X., Mamidala, P., Rajarapu, S. P., Jones, S. C. and Mittapalli, O. (2011) Transcriptomics of the bed bug (*Cimex lectularius*). *PLoS ONE*, **6**, 1–10.
- Balvín, O. and Booth, W. (2018) Distribution and frequency of pyrethroid resistance-associated mutations in host lineages of the bed bug (Hemiptera: Cimicidae) across Europe. *Journal of Medical Entomology*, **55**, 923–928.
- Benoit, J. B., Adelman, Z. N., Reinhardt, K., Dolan, A., Poelchau, M., Jennings, E. C., Szuter, E. M., Hagan, R. W., Gujar, H., Shukla, J. N., Zhu, F., Mohan, M., Nelson, D. R., Rosendale, A. J., Derst, C., Resnik, V., Wernig, S., Menegazzi, P., Wegener, C., Peschel, N., Hendershot, J. M., Blenau, W., Predel, R., Johnston, P. R., Ioannidis, P., Waterhouse, R. M., Nauen, R., Schorn, C., Ott, M. C., Maiwald, F., Johnston, J. S., Gondhalekar, A. D., Scharf, M. E., Peterson, B. F., Raje, K. R., Hottel, B. A., Armisén, D., Crumière, A. J. J., Refki, P. N., Santos, M. E., Sghaier, E., Viala, S., Khila, A., Ahn, S. J., Childers, C., Lee, C. Y., Lin, H., Hughes, D. S., Duncan, E. J., Murali, S. C., Qu, J., Dugan, S., Lee, S. L., Chao, H., Dinh, H., Han, Y., Doddapaneni, H., Worley, K. C., Muzny, D. M., Wheeler, D., Panfilio, K. A., Vargas Jentzsch, I. M., Vargo, E. L., Booth, W., Friedrich, M., Weirauch, M. T., Anderson, M. A., Jones, J. W., Mittapalli, O., Zhao, C., Zhou, J. J., Evans, J. D., Attardo, G. M., Robertson, H. M., Zdobnov, E. M., Ribeiro, J. M., Gibbs, R. A., Werren, J. H., Palli, S. R., Schal, C. and Richards, S. (2016) Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature Communications*, **7**, 1–10.
- Cannavò, E., Khoueiry, P., Garfield, D., Geeleher, P., Zichner, T., Gustafson, E., Ciglar, L., Korbel, J. and Furlong, E. (2016) Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. *Current Biology*, **26**, 38–51. URL: <https://www.sciencedirect.com/science/article/pii/S0960982215014281>.
- Dang, K., Toi, C. S., Lilly, D. G., Bu, W. and Doggett, S. L. (2015) Detection of knockdown resistance mutations in the common bed bug, *Cimex lectularius* (Hemiptera: Cimicidae), in Australia. *Pest Management Science*, **71**, 914–922.
- Dang, K. A. I., Toi, C. S., Lilly, D. G., Lee, C.-y., Naylor, R., Tawatsin, A., Thavara, U., Bu, W. and Doggett, S. L. (2014) Detection of knockdown resistance (kdr) in *Cimex lectularius* and *Cimex hemipterus* (Hemiptera: Cimicidae). *The Eighth International Conference on Urban Pests*, **71**, 914–922.
- Davies, T. G., Field, L. M. and Williamson, M. S. (2012) The re-emergence of the bed bug as a nuisance pest: Implications of resistance to the pyrethroid insecticides. *Medical and Veterinary Entomology*, **26**, 241–254.
- Doggett, S. L., Geary, M. J. and Russell, R. C. (2004) The resurgence of bed bugs in Australia: with notes on their ecology and control. *Environmental health*, **4**, 30–38.
- Durand, R., Cannet, A., Berdjane, Z., Bruel, C., Haouchine, D., Delaunay, P. and Izri, A. (2012) Infestation by pyrethroids resistant bed bugs in the suburb of Paris, France. *Parasite (Paris, France)*, **19**, 381–387.
- Dushay, M. S. and Eldon, E. D. (1998) *Drosophila* immune responses as models for human immunity. *The American Journal of Human Genetics*, **62**, 10–14.
- Epis, S., Porretta, D., Mastrantonio, V., Urbanelli, S., Sasseria, D., De Marco, L., Mereghetti, V., Montagna, M., Ricci, I., Favia, G. and Bandi, C. (2014) Temporal dynamics of the ABC transporter response to insecticide treatment: insights from the malaria vector *Anopheles stephensi*. *Scientific Reports*, **4**, 7435. URL: <https://doi.org/10.1038/srep07435>.
- Faucon, F., Dusfour, I., Gaude, T., Navratil, V., Boyer, F., Chandre, F., Sirisopa, P., Thanispong, K., Juntarajumng, W., Poupardin, R., Chareonviriyaphap, T., Girod, R., Corbel, V., Reynaud, S. and David, J.-p. (2015) Unravelling genomic changes associated with insecticide resistance in the dengue mosquito *Aedes aegypti* by deep targeted sequencing. *Genome Research*, 1–13.
- Feng, K., Liu, J., Wei, P., Ou, S., Wen, X., Shen, G., Xu, Z., Xu, Q. and He, L. (2020) lincRNA_Tc13743.2-miR-133-5p-TcGSTm02 regulation pathway mediates cyflumetofen resistance in *Tetranychus cinnabarinus*. *Insect Biochemistry and Molecular Biology*, **123**, 103413. URL: <https://www.sciencedirect.com/science/article/pii/S0965174820301028>.

- Fountain, T., Ravinet, M., Naylor, R., Reinhardt, K. and Butlin, R. K. (2016) A linkage map and QTL analysis for pyrethroid resistance in the bed bug *Cimex lectularius*. *G3: Genes, Genomes, Genetics*, **6**, 4059–4066.
- Goddard, J. and Deshazo, R. (2009) Bed Bugs (*Cimex lectularius*) and Clinical Consequences of Their Bites. *JAMA: the journal of the American Medical Association*, **301**, 1358–1366.
- Gottschalk, A., Almedom, R. B., Schedletzky, T., Anderson, S. D., Yates III, J. R. and Schafer, W. R. (2005) Identification and characterization of novel nicotinic receptor-associated proteins in *Caenorhabditis elegans*. *The EMBO Journal*, **24**, 2566–2578. URL: <https://doi.org/10.1038/sj.emboj.7600741>.
- Guedes, R. N. C., Walse, S. S. and Throne, J. E. (2017) Sublethal exposure, insecticide resistance, and community stress. *Current Opinion in Insect Science*, **21**, 47–53.
- Haberkorn, C., David, J.-P., Henri, H., Delpuech, J.-M., Lasseur, R., Vavre, F. and Varaldi, J. (2023) Identification of new candidates for insecticide resistance in bed bugs through genome-wide analysis. *in prep*, **xx**, **xx**.
- Koganemaru, R., Miller, D. M. and Adelman, Z. N. (2013) Robust cuticular penetration resistance in the common bed bug (*Cimex lectularius* L.) correlates with increased steady-state transcript levels of CPR-type cuticle protein genes. *Pesticide Biochemistry and Physiology*, **106**, 190–197. URL: <http://dx.doi.org/10.1016/j.pestbp.2013.01.001>.
- Li, C.-x., Guo, X.-x., Zhang, Y.-m., Dong, Y.-d., Xing, D., Yan, T., Wang, G., Zhang, H.-d. and Zhao, T.-y. (2016) Identification of genes involved in pyrethroid-, propoxur-, and dichlorvos- insecticides resistance in the mosquitoes, *Culex pipiens* complex (Diptera: Culicidae). *Acta Tropica*, **157**, 84–95. URL: <https://www.sciencedirect.com/science/article/pii/S0001706X16300195>.
- Li, X., Schuler, M. A. and Berenbaum, M. R. (2007) Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. *Annual Review of Entomology*, **52**, 231–253. URL: <http://www.annualreviews.org/doi/10.1146/annurev.ento.51.110104.151104>.
- Liang, Z.-K., Pang, R., Dong, Y., Sun, Z.-X., Ling, Y. and Zhang, W.-Q. (2018) Identification of SNPs involved in regulating a novel alternative transcript of P450 CYP6ER1 in the brown planthopper. *Insect Science*, **25**, 726–738. URL: <https://doi.org/10.1111/1744-7917.12472>.
- Lilly, D. G., Dang, K., Webb, C. E. and Doggett, S. L. (2016a) Evidence for metabolic pyrethroid resistance in the common bed bug (Hemiptera: Cimicidae). *Journal of Economic Entomology*, **109**, 1364–1368.
- Lilly, D. G., Latham, S. L., Webb, C. E. and Doggett, S. L. (2016b) Cuticle thickening in a pyrethroid-resistant strain of the common bed bug, *Cimex lectularius* L. (Hemiptera: Cimicidae). *PLoS ONE*, **11**, 6–16.
- Love, M. I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550. URL: <https://doi.org/10.1186/s13059-014-0550-8>.
- Mamidala, P., Jones, S. C. and Mittapalli, O. (2011) Metabolic resistance in bed bugs. *Insects*, **2**, 36–48.
- Mamidala, P., Wijeratne, A. J., Wijeratne, S., Kornacker, K., Sudhamalla, B., Rivera-Vega, L. J., Hoelmer, A., Meulia, T., Jones, S. C. and Mittapalli, O. (2012) RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug. *BMC Genomics*, **13**.
- Mastrantonio, V., Ferrari, M., Epis, S., Negri, A., Scuccimarra, G., Montagna, M., Favia, G., Porretta, D., Urbanelli, S. and Bandi, C. (2017) Gene expression modulation of ABC transporter genes in response to permethrin in adults of the mosquito malaria vector *Anopheles stephensi*. *Acta tropica*, **171**, 37–43.
- Negri, A., Ferrari, M., Nodari, R., Coppa, E., Mastrantonio, V., Zanzani, S., Porretta, D., Bandi, C., Urbanelli, S. and Epis, S. (2019) Gene silencing through RNAi and antisense Vivo-Morpholino increases the efficacy of pyrethroids on larvae of *Anopheles stephensi*. *Malaria Journal*, **18**, 294. URL: <https://doi.org/10.1186/s12936-019-2925-5>.

- Pan, C., Zhou, Y. and Mo, J. (2009) The clone of laccase gene and its potential function in cuticular penetration resistance of *Culex pipiens pallens* to fenvalerate. *Pesticide Biochemistry and Physiology*, **93**, 105–111.
- Peng, T., Pan, Y., Tian, F., Xu, H., Yang, F., Chen, X., Gao, X., Li, J., Wang, H. and Shang, Q. (2021) Identification and the potential roles of long non-coding RNAs in regulating acetyl-CoA carboxylase ACC transcription in spirotetramat-resistant *Aphis gossypii*. *Pesticide Biochemistry and Physiology*, **179**, 104972. URL: <https://www.sciencedirect.com/science/article/pii/S0048357521002030>.
- Potter, M. F. (2011) The history of bed bug management- With lessons from the past. *American Entomologist*, **57**, 14–25.
- Poupardin, R., Reynaud, S., Strode, C., Ranson, H., Vontas, J. and David, J.-P. (2008) Cross-induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito *Aedes aegypti*: Impact on larval tolerance to chemical insecticides. *Insect Biochemistry and Molecular Biology*, **38**, 540–551. URL: <https://www.sciencedirect.com/science/article/pii/S0965174808000234>.
- Romero, A. and Anderson, T. D. (2016) High levels of resistance in the common bed bug, *Cimex lectularius* (hemiptera: Cimicidae), to neonicotinoid insecticides. *Journal of Medical Entomology*, **53**, 727–731.
- Romero, A., Potter, M. F. and Haynes, K. F. (2009) Evaluation of Piperonyl Butoxide as a Deltamethrin Synergist for Pyrethroid-Resistant Bed Bugs. *Journal of Economic Entomology*, **102**, 2310–2315.
- Romero, A., Potter, M. F., Potter, D. A. and Haynes, K. F. (2007) Insecticide Resistance in the Bed Bug: A Factor in the Pest's Sudden Resurgence? *Journal of Medical Entomology*, **44**, 175–178.
- Rosenfeld, J. A., Reeves, D., Brugler, M. R., Narechania, A., Simon, S., Durrett, R., Foox, J., Shianna, K., Schatz, M. C., Gandara, J., Afshinnekoo, E., Lam, E. T., Hastie, A. R., Chan, S., Cao, H., Saghbini, M., Kentsis, A., Planet, P. J., Kholodovych, V., Tessler, M., Baker, R., DeSalle, R., Sorkin, L. N., Kolokotronis, S. O., Siddall, M. E., Amato, G. and Mason, C. E. (2016) Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nature Communications*, **7**.
- Stutt, A. D. and Siva-Jothy, M. T. (2001) Traumatic insemination and sexual conflict in the bed bug *Cimex lectularius*. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5683–5687.
- Yahouédo, G. A., Chandre, F., Rossignol, M., Ginibre, C., Balabanidou, V., Mendez, N. G. A., Pigeon, O., Vontas, J. and Cornelie, S. (2017) Contributions of cuticle permeability and enzyme detoxification to pyrethroid resistance in the major malaria vector *Anopheles gambiae*. *Scientific reports*, **7**, 11091.
- Yang, Y., Zhang, Y., Wang, A., Duan, A., Xue, C., Wang, K., Zhao, M. and Zhang, J. (2022) Four MicroRNAs, miR-13b-3p, miR-278-5p, miR-10483-5p, and miR-10485-5p, Mediate Insecticide Tolerance in *Spodoptera frugiperda*. *Frontiers in Genetics*, **12**, 1–11.
- Zhu, F., Gujar, H., Gordon, J. R., Haynes, K. F., Potter, M. F. and Palli, S. R. (2013) Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Scientific Reports*, **3**, 1–8.
- Zhu, F., Sams, S., Moural, T., Haynes, K. F., Potter, M. F. and Palli, S. R. (2012) RNA interference of NADPH-cytochrome P450 reductase results in reduced insecticide resistance in the bed bug, *Cimex lectularius*. *PLoS ONE*, **7**.
- Zhu, F., Wigginton, J., Romero, A., Moore, A., Ferguson, K., Palli, R., Potter, M. F., Haynes, K. F. and Palli, S. R. (2010) Widespread distribution of knockdown resistance mutations in the bed bug, *Cimex lectularius* (Hemiptera: Cimicidae), populations in the United States. *Archives of Insect Biochemistry and Physiology*, **73**, 245–257.

6 | SUPPLEMENTARY

LG	Scaffold	Start	End	Gene ID	LFC	<i>p</i> _{adj}	Annotation
5	NW_019392649.1	1017882	1018792	106664492	3.44	2.97E-03	larval cuticle protein A3A
		1046176	1046953	106664495	2.74	4.53E-03	larval cuticle protein A3A
13	NW_019392685.1	756768	757614	106662986	2.94	1.52E-02	cuticle protein 12.5
-	NW_019392920.1	77411	82689	106664047	2.87	1.55E-02	larval cuticle protein A1A
-	NW_019392993.1	2437772	2454768	106666062	2.81	2.98E-02	cuticle protein
-	NW_019393130.1	673262	674032	106664300	2.11	1.08E-02	larval cuticle protein A3A

SUPPLEMENTARY TABLE 1 Resistance genes **over**-expressed in London Field when compared to London Lab strain for both untreated and treated survivors. LFC corresponds to the Log2 Fold Change between conditions (i.e. LFC of 3.44 means 10.85 fold change).

LG	Scaffold	Start	End	Gene ID	LFC	<i>p</i> _{adj}	Annotation
4	NW_019392661.1	361896	365391	106662622	0.90	2.60E-02	cytochrome P450 6k1
5	NW_019392644.1	647119	719373	106670072	0.62	1.03E-02	ABC transporter G 4
		749946	794199	106670075	0.78	1.88E-03	ABC transporter G 4
		1346254	1453427	106670073	0.99	3.20E-03	ABC transporter G 4
	NW_019392682.1	973424	1009568	106665518	0.66	6.15E-04	ABC transporter G 4
9	NW_019392638.1	815190	988569	106670919	0.79	1.54E-03	segmentation protein cap'n'collar
	NW_019392675.1	595496	597255	106662042	1.77	1.07E-03	cuticle protein 12.5
11	NW_019392763.1	493269	495040	106664023	1.30	4.07E-03	GST
-	NW_019392631.1	346739	535797	106662809	0.61	1.56E-02	deacetylase/sulfotransferase
-	NW_019392668.1	644977	687162	106670261	0.69	3.65E-03	ABC transporter G 1
-	NW_019392709.1	60749	80839	106663888	0.67	2.70E-02	ABC transporter G 23
-	NW_019392742.1	1238897	1243972	106664801	0.88	4.89E-04	protein-tyrosine sulfotransferase
		1771166	1773352	106666818	1.25	4.51E-03	transcription factor MafB
-	NW_019392764.1	512150	523899	106661542	0.89	8.26E-03	cytochrome b5-related protein
-	NW_019392803.1	7	886	112127332	2.63	1.50E-02	AChE 1
		7080	7993	106669437	3.12	1.25E-03	AChE
-	NW_019392929.1	126420	138992	106674256	1.41	1.20E-04	GST
-	NW_019393228.1	8	1146	112128126	1.88	1.31E-02	probable cytochrome P450 9f2
-	NW_019394179.1	823	3373	106669436	3.30	2.74E-04	acetylcholinesterase
-	NW_019394202.1	1886643	1895153	106666582	0.60	2.86E-02	carbohydrate sulfotransferase 11

SUPPLEMENTARY TABLE 2 Resistance genes **over**-expressed in treated survivors when compared to untreated individuals for both London Lab and London Field strains. LFC corresponds to the Log2 Fold Change between conditions (i.e. LFC of 0.90 means 1.87 fold change).

LG	Scaffold	Start	End	Gene ID	LFC	p_{adj}	Annotation
1	NW_019392641.1	990484	1009113	106669934	-1.25	7.83E-03	SOD [Cu-Zn]
2	NW_019392653.1	1404658	1420128	106660996	-0.81	1.09E-02	ABC transporter G 23
		1427987	1437993	106661276	-0.80	4.20E-02	ABC transporter G 23
	NW_019392979.1	33885	34615	106665194	-0.96	4.62E-02	endocuticle structural glycoprot.
4	NW_019392730.1	1347727	1365982	106662446	-1.41	5.20E-04	sodium channel protein Nach
		54557	58081	106670437	-0.77	1.53E-02	cuticle protein 18.7
		220047	220929	106670525	-1.24	3.87E-03	pupal cuticle protein
		329822	331325	106670323	-0.75	3.13E-02	cuticle protein 18.7
6	NW_019392719.1	1392866	1401379	106661029	-0.96	2.62E-03	pupal cuticle protein 20
13	NW_019392637.1	3045820	3055263	106671932	-0.80	3.18E-02	AChE
		529446	602890	106665955	-0.74	2.84E-02	sulfotransferase
-	NW_019392668.1	1276458	1279945	106670277	-0.87	2.53E-02	probable cytochrome P450 6a14
-	NW_019392801.1	1599	16364	106669268	-1.39	1.92E-02	cytochrome P450 4C1
-	NW_019392822.1	780550	784904	106665184	-0.97	6.62E-03	endocuticle structural glycoprot.
-	NW_019393130.1	625738	626515	106664854	-0.85	2.03E-02	larval cuticle protein A2B

SUPPLEMENTARY TABLE 3 Resistance genes **under**-expressed in treated survivors when compared to untreated individuals for both London Lab and London Field strains. LFC corresponds to the Log2 Fold Change between conditions (i.e. LFC of -1.25 means -2.37 fold change).

LG	Scaffold	Gene ID	Annot.	TypeEvent	Size	F_{LL}	F_{LF}	F_{ST}
3	NW_019392860.1	LOC106674398	uncharacterized	Inv	15 kb	0.10	0.21	0.022
5	NW_019392649.1	106664492	cuticle A3A	InvDup	248 kb	0.057	0.123	0.013
		106664495	cuticle A3A			0.057	0.123	0.013
10	NW_019392795.1	106667662	GRP wall structural	InvDup	1 M	0.03	0.08	0.016
				InvDup	241 kb	0.16	0.17	1.18E-4
				Inv	550 kb	0.16	0.19	1.99E-03
-	NW_019392939.1	112127884	uncharacterized	Inv	1 M	0.07	0.09	1.02E-03

SUPPLEMENTARY TABLE 4 Genes up-regulated in London Field when compared to London Lab individuals detected inside a structural variant (inversion "Inv", inverted duplication "InvDup" or tandem duplication "TanDup"). Size and frequencies of those variants in both strain were showed together with associated differentiation index (F_{ST}). Gene annotations have been shortened, but can be found using the gene ID on NCBI.

DE	LG	Scaffold	Position	Gene ID	Annotation	Location	R>A	F_{altLL}	F_{altLF}	pF_{ST}
up	12	NW_019392813.1	4400166	112127381	uncharacterized	C I P	C>T	0.20	0.68	0.023
up	-	NW_019392656.1	2931030	106664583	FAR wat	C I	T>G	0.15	0.72	0.006
			2931057			C I	C>A	0.23	0.83	0.005
			2931088			C I	G>T	0.25	0.90	0.003
			2944507			C	T>A	0.10	0.52	0.036
up	-	NW_019392664.1	1542428	106662188	venom allergen	C I	T>G	0.30	0.85	0.011
			1542429			C I	T>G	0.30	0.83	0.014
up	-	NW_019392712.1	552763	106669984	protein goliath	C	T>C	0.06	0.50	0.030
up	-	NW_019392775.1	745973	106663301	AF4/FMR2 4	C	A>G	0.52	1	0.013
up	-	NW_019392805.1	43356	106673166	BCAT	C	G>C	0.43	0.83	0.046
up	-	NW_019392993.1	70657	106664949	serpin B3	C	A>T	0.60	1	0.015
			70664			C	C>G	0.55	1	0.015
			70665			C	T>C	0.55	1	0.015
do	3	NW_019392650.1	261464	106670335	uncharacterized	C	A>T	0.17	0.60	0.040
do	7	NW_019392715.1	430046	106665593	titin homolog	C	A>T	0.48	0.93	0.029
do	8	NW_019393011.1	184388	106665395	solute carrier f26 m10	C	C>A	0.22	0.67	0.033
do	-	NW_019392712.1	537273	106669983	forkhead box P1	C	T>C	0.12	0.54	0.036

SUPPLEMENTARY TABLE 5 Non-synonymous outlier SNPs in differentially expressed (DE) genes (up or down-regulated) for treatment condition. SNPs were selected with the frequency of alternative allele (F_{alt}) higher in London Field (LF) than London Lab (LL) and differentiation index F_{ST} value in top 5% quantile (i.e. empirical $pF_{ST} < 0.05$). The position of each SNP was given, together with gene ID, gene annotation, reference to alternative nucleotide (R>A) and other putative location of the SNP whenever position was litigious due to the gap between transcripts annotation (coding=C, intron=I, promoter=P).

4

Métagénomique infectieuse

4.1	Contexte - encore une sombre histoire d'élevage	142
4.2	Exploration de l'ADN hors punaise de lit	144
4.2.1	Assemblage et analyse des contigs	144
4.2.2	Identification taxonomique	146
4.3	Analyse des punaises de lit au phénotype oxydé	148
4.3.1	Extraction et séquençage des acides nucléiques viraux putatifs	149
	Insectes	149
	Broyage et filtration	149
	Digestion des acides nucléiques de <i>C. lectularius</i>	150
	Extraction des acides nucléiques restants	150
	Digestion enzymatique des échantillons obtenus	151
4.3.2	Assemblage et alignement des contigs	153
4.3.3	Identification taxonomique	154
	Contigs ADN	154
	Contigs ARN	157
4.4	Validation de candidats	159
4.4.1	Préparation des échantillons	159
4.4.2	Réalisation de la PCR	160
4.4.3	Migration sur gel d'agarose	161
4.5	Discussion & Perspectives	163
4.5.1	Potentiels agents pathogènes détectés chez la punaise de lit	163
4.5.2	Exploration des bases de données SRA	164
4.5.3	Gestion de la mortalité en laboratoire	165
4.6	Matériel supplémentaire	168

Chloé Haberkorn^{1,2}, David Lepetit¹, Zaïna Belgaidi¹, Élodie Rosinski¹,
Natacha Kremer¹, Romain Lasseur², Fabrice Vavre¹, Julien Varaldi¹

¹ Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, UMR 5558,
Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France

² IZInovation, 13 Rue des Émeraudes, 69006 Lyon, France

4.1 Contexte - encore une sombre histoire d'élevage

Chez d'autres insectes hématophages, certains virus peuvent avoir des conséquences désastreuses sur les élevages. C'est le cas du virus de l'hypertrophie des glandes salivaires chez la mouche tsé-tsé *Glossina pallidipes* (GpSGHV, *Hytrosaviridae*).

Les infections virales sont connues pour pouvoir être symptomatiques (avec des symptômes observables de "maladies") ou asymptomatiques (sans symptômes visibles) (Abd-Alla *et al.*, 2021). Ces symptômes peuvent aller d'une perte de valeur sélective, à la mort des insectes infectés. *A contrario*, les infections latentes cryptiques sont indétectables jusqu'à leur activation par des conditions abiotiques défavorables, comme un changement dans la température ou l'hygrométrie, ou une augmentation de la densité de population (Tanada and Omi, 1974).

C'est ce qui a été observé dans de nombreux élevages de mouches tsé-tsé, où les colonies stagnaient, présentant une forte mortalité et une faible fécondité (Jaenson, 1986; FAO/IAEA, 1987). Des dissections ont ensuite révélé une hypertrophie des glandes salivaires chez une majorité des insectes, associée à une dégénérescence testiculaire et une aspermie chez les mâles (Kokwaro, 2006). Pourtant, la prévalence du virus différait fortement entre les élevages, l'augmentation de la prévalence étant corrélée négativement au déclin d'une souche. Il s'est alors avéré que la principale différence entre les colonies plus ou moins touchées résidait dans les techniques de nourrissage utilisées. En effet, ce virus aurait été détecté comme particulièrement abondant au sein de certains élevages pratiquant les nourrissages successifs des insectes sur un système *in vitro*. Il a été démontré par la suite que ce système, qui semble très similaire à l'Hemotek utilisé dans nos élevages, favorisait l'accumulation et la transmission de particules virales de GpSGHV (Abd-Alla *et al.*, 2021).

Au sein de notre élevage de punaises de lit, nous avons aussi eu affaire à une mortalité suspecte. Le phénotype de cette mortalité était caractérisé ainsi : suite au repas de sang, les punaises de lit ne semblent pas le digérer, le tube digestif se rompt et le sang emplit alors la cavité principale de l'insecte, faisant virer la couleur du corps de l'insecte au rouge sombre puis au noir (Figure 4.1).

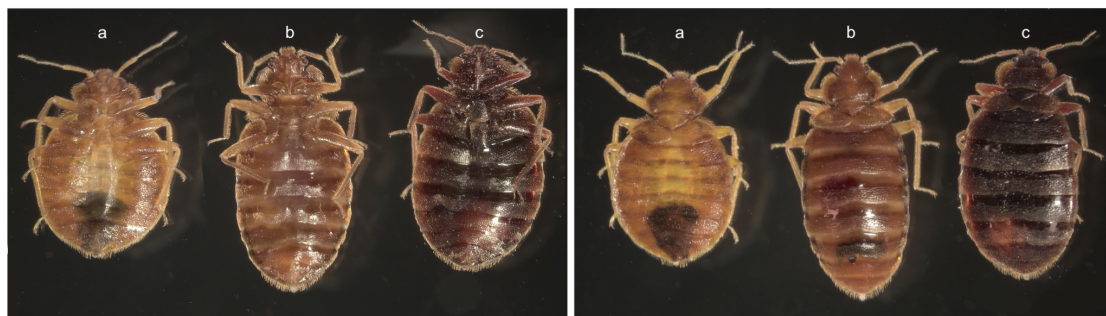


Figure 4.1 – Apparence des punaises de lit touchées par le phénomène de mortalité anormal au sein des élevages. Photographies de femelles mortes à jeun (a), après gorgement (b), et présentant le phénotype oxydé post-gorgement (c) en face ventrale (à gauche) et dorsale (à droite).

Suite à nos sollicitations, plusieurs laboratoires Européens (en Allemagne, Angleterre, et République Tchèque) ont reconnu avoir expérimenté des épisodes de mortalité similaires au sein de leurs élevages de punaises de lit, avec des conséquences parfois catastrophiques :

*“We have experienced mortality as you have observed. (...) It seems to occur, if the humidity is high in our cultures, shortly after feeding. I’m not sure whether it is a feature of increased metabolism immediately after feeding coupled with the need to excrete water from their bodies quickly, which leads to high humidity. We almost lost the population because we had fed the bedbugs and then left the culture for a week, whilst on holiday and we returned to **90% of the bugs like this.** (...) We manage the humidity by reducing the density of bedbugs.”*
(Communication personnelle, Nov. 2021)

Ce phénotype est observé dans différents types d'élevages avec des conditions physiques (température et hygrométrie) et physiologiques variables (gorgement sur volontaires humains, rongeurs vivants, ou système artificiel, avec du sang provenant de différents animaux et en ajoutant divers types d'anticoagulants). À notre connaissance, la littérature ne le mentionne pas comme ayant été observé, que ce soit en laboratoire ou *in natura* (aucune publication, seulement des bruits de couloir). Le seul dénominateur commun semble donc être le fait de garder les punaises dans des espaces clos relativement restreints (pots en plastiques ou en

verre). Néanmoins, il semblerait que la prévalence de ce phénotype reste à niveau très faible, sans épisode majeur de mortalité, au sein des élevages nourris sur volontaires humains, renforçant ainsi les similitudes avec les observations sur *G. pallidipes*. Nous avons ainsi formulé l'hypothèse qu'un agent infectieux pouvait être impliqué dans ce phénomène.

L'exploration de cette hypothèse est guidée par deux principaux objectifs : (i) faciliter l'élevage de ce modèle d'étude, rendu important par son fort pouvoir de nuisance sociétal, et (ii) identifier de potentiels futurs agents de bio-contrôle.

Nous avons donc tout d'abord réalisé une analyse du métagénome de punaises de lit au phénotype *a priori* sain à partir de données pool-seq, avant d'étudier le cortège de micro-organismes de punaises porteuses du phénotype oxydé.

4.2 Exploration de l'ADN hors punaise de lit

Dans un premier temps, nous avons donc utilisé les données pool-seq explorées lors des analyses génomiques (Chapitre 2). Ces punaises ont été achetées en 2018, bien avant donc les gros événements de mortalité observés par la société CimexStore. Elles sont envoyées par la société juste après leur avoir fait prendre un repas de sang, et l'extraction avait été réalisée sur des punaises vivantes d'aspect sain - on peut donc supposer que l'agent infectieux n'était pas présent, ou alors en quantité moindre, et donc que ces données peuvent s'approcher d'un "contrôle négatif". Ces données correspondant à des extraits ADN, on s'attend à ce qu'elles nous permettent d'identifier des micro-organismes ayant un génome ADN.

4.2.1 Assemblage et analyse des contigs

Parmi les reads des données ADN pool-seq déjà à notre disposition, tous n'appartiennent pas à la punaise de lit, et ne sont donc pas cartographiables sur son génome de référence. Après avoir enlevé les duplicats de PCR avec FastUniq (Xu *et al.*, 2012), il est donc possible d'extraire tous les reads ne s'alignant pas sur le génome de référence de *Cimex lectularius* grâce à Hisat2 (Zhang *et al.*, 2021), puis de les apparier à l'aide de Trimmomatic (Bolger *et al.*, 2014). Les pourcentages de ces reads non-mappés par rapport à l'ensemble des reads apparierés sont de 3,40% pour London Lab, 1,80% pour German Lab, 1,87% pour

London Field et 5,22% pour Sweden Field. Ils peuvent correspondre à des contaminants divers, mais également à des micro-organismes infectant les punaises (bactéries, virus ou autres).

Dans un second temps, ces reads ont été assemblés en séquences appelées contigs à l'aide de Megahit (Li *et al.*, 2015) (Figure 4.2). Les contigs ainsi formés ont été comparés à une base de données généralistes protéiques nr (non-redondante) avec MMseqs2 (Steinegger and Söding, 2017) afin d'identifier les différents micro-organismes.

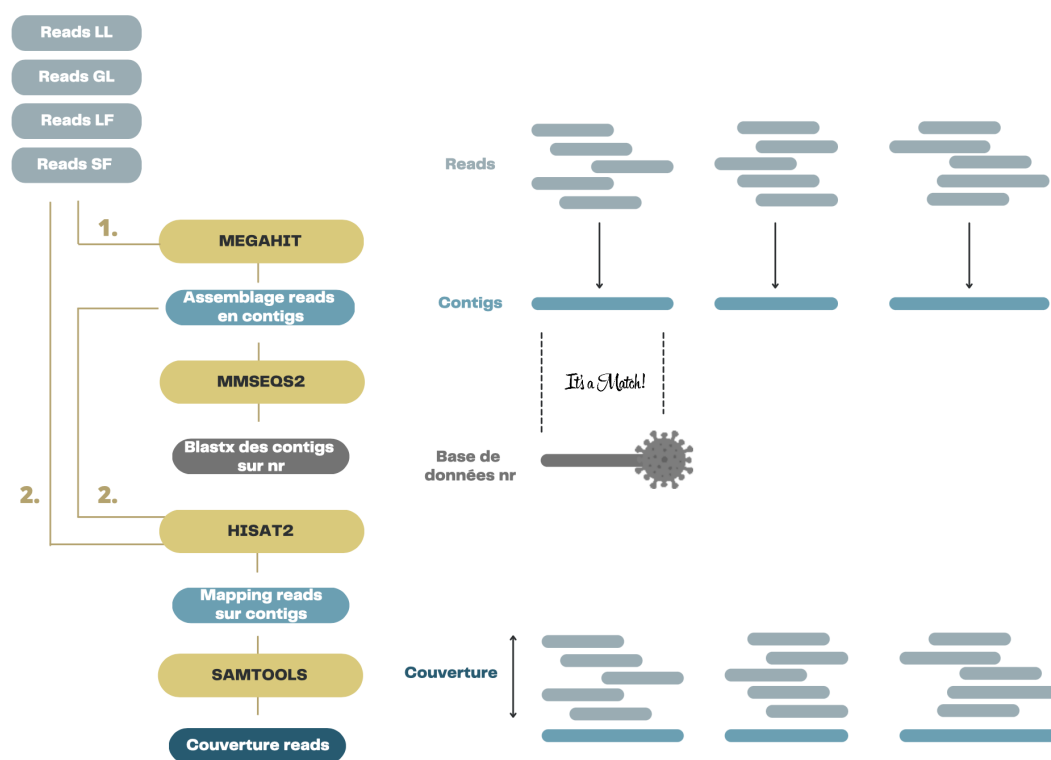


Figure 4.2 – Pipeline de traitement des reads pour la détection d'organismes hors punaise de lit. Dans les cadres jaunes, on retrouve les noms des programmes utilisés, avec dans les cadres en dessous les actions qu'ils réalisent. Principales étapes : (1.) Utilisation des reads pour former des contigs qui seront alignés sur la base de donnée nr, (2.) Utilisation des reads et des contigs formés pour calculer la couverture desdits contigs.

Les reads ont ensuite été réalignés sur les contigs grâce à Hisat2, ce qui a permis de calculer la profondeur de la couverture à chaque position nucléotidique avec Samtools (Li *et al.*, 2009). Une annotation taxonomique des résultats du blast a été réalisée, afin de détecter les potentiels virus, mais également les autres agents infectieux, bactériens ou eucaryotes (parasites, champignons, protozoaires...). Enfin, les résultats ont été filtrés avec des seuils draconiens, comme un score d'identité supérieur à 90%, ainsi qu'un bit-score (mesure de la

similarité des séquences indépendamment de la taille de la séquence en entrée ou de la base de données) supérieur à 50, considéré comme le seuil de significativité (Pearson, 2013).

4.2.2 Identification taxonomique

Ces analyses ont permis d'identifier 38 876 contigs, dont 38 148 annotés jusqu'au rang de l'espèce. Les contigs étaient principalement assignés à des eucaryotes (32 764), puis à des bactéries (5 383), et enfin à un seul virus. Pour chaque espèce, la taille cumulée (t_c) des contigs détectés est calculée, afin de visualiser les plus représentés (Figure 4.3). Néanmoins, seules les espèces présentant une t_c supérieure à 500 kb sont ici décrites dans le texte.

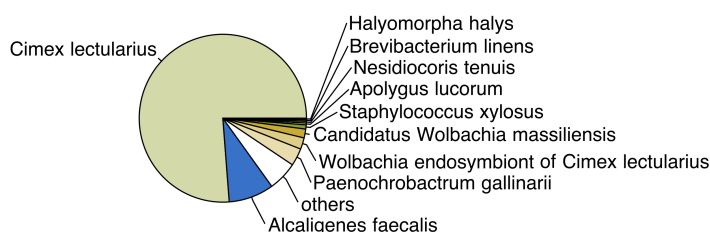


Figure 4.3 – Tailles cumulées des contigs détectés pour chaque espèce au sein des données pool-seq, hors alignement sur le génome de *C. lectularius*. Les 10 espèces les plus représentées sont annotées, et les autres sont regroupées dans "others".

Chez les eucaryotes, on retrouve ainsi sans surprise une très large majorité de contigs de *C. lectularius* ($n=31\ 151$; $t_c=39$ Mb). Ces données sont cohérentes avec les pourcentages de mapping des données pool-seq observés sur le génome de référence de la punaise de lit (entre 94 et 98% en fonction des souches, avec les 6% restants de la taille totale du génome, soit 510 Mb, correspondant à 30 Mb).

Du côté des bactéries, les contigs à la taille cumulée la plus importante appartiennent à *Alcaligenes faecalis* ($n=661$, $t_c=4,4$ Mb), suivie de *Paenochrobactrum gallinarum* ($n=1\ 263$, $t_c=1,7$ Mb). *Paenochrobactrum gallinarum* est une alpha-protéobactérie (Ordre : Hyphomicrobiales, Famille : Brucellaceae), isolée dans l'air d'un élevage de canard, dont on ne connaît pas grand-chose de plus que son métabolisme oxydatif, et la grande variété des lieux dans lesquels les bactéries de ce genre ont été détectées (Kämpfer *et al.*, 2010).

Étonnamment, les dix plus grand contigs détectés (262 kb à 103 kb) sont attribués à la bactérie *Alcaligenes faecalis*. *A. faecalis* est un bacille nommé *faecalis* d'après sa découverte originelle dans des fèces, mais présent également de

manière plus globale dans le sol et l'eau.

Une étude a exploré l'hypothèse d'une symbiose entre la bactérie *A. faecalis* et un nématode, retrouvés ensemble à l'intérieur d'un cadavre de Lépidoptère (Quiroz-Castañeda *et al.*, 2015). Plusieurs espèces de nématodes (*Steinernema feltiae*, *S. carpocapsae*, *Heterorhabditis bacteriophora*) ont été mises en contact avec *A. faecalis* et une bactérie servant de contrôle négatif, *E. coli*. *A. faecalis* a ensuite été détectée dans le tube digestif des nématodes, alors que *E. coli* était retrouvé à l'extérieur de la cuticule. La pathogénicité et la toxicité d'*A. faecalis* a ensuite été démontrée chez des larves du Lépidoptère *Galleria mellonella*. Ces bactéries possèdent en effet de nombreuses enzymes (chitinases, estérases) impliquées dans la toxicité, et pouvant participer à l'infection et la dégradation du corps d'un insecte. Les interactions décrites dans cette étude sont résumées dans un schéma (Figure 4.4).

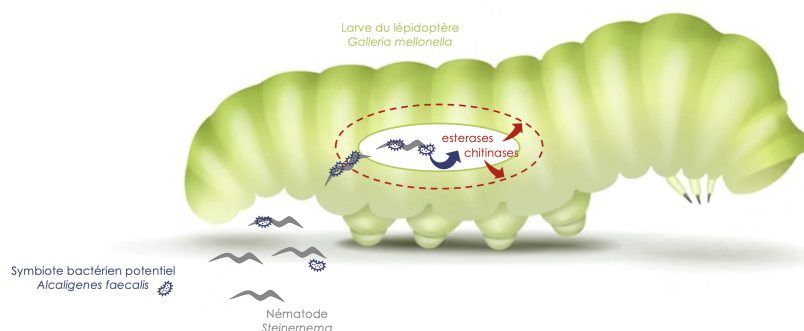


Figure 4.4 – Interactions décrites entre le nématode *Steinernema*, la bactérie *Alcaligenes* et la larve de Lépidoptère *Galleria*, d'après Quiroz-Castañeda *et al.* (2015).

Ce type de symbiose mutualiste a déjà été décrite entre la bactérie *Xenorhabdus* et le nématode *Steinernema* (Awori, 2022). Le nématode infecte l'insecte par ses orifices naturels, comme les spiracles qui permettent la respiration. Une fois dans la cavité interne de l'insecte, le nématode relâche son symbiote intestinal par les fèces. La bactérie va alors sécréter des molécules boostant la fécondité du nématode, et le protéger en inhibant les phénoloxydases du système immunitaire de l'insecte. Une fois que la progéniture du nématode a atteint le troisième stade juvénile, elle s'associe à nouveau avec la bactérie *Xenorhabdus*, avant de quitter le cadavre de l'insecte à la recherche d'un nouvel hôte.

Parmi les bactéries détectées, on retrouve ensuite le symbiote obligatoire *Wolbachia* décrit chez la punaise de lit (n=236, $t_c=1,1$ Mb) (Hosokawa *et al.*, 2010), ainsi que celui détecté chez la punaise de lit tropicale *C. hemipterus* : *Wolbachia massiliensis* (n= 611, $t_c=868$ kb) (Laidoudi *et al.*, 2020). La confusion entre ces deux espèces très proche est possible puisqu'il ne devrait y avoir que

Wolbachia dans ces données, malgré le plus fort pourcentage d'identité moyen mesuré pour *Wolbachia massiliensis* (95,2% contre 94,96% pour la *Wolbachia* inféodé à *C. lectularius*).

D'autres bactéries ont été identifiées par de précédentes études chez *C. lectularius* (voir le chapitre 1, section 1.1.2), telles qu'une protéobactérie BEV-like de la famille des *Enterobacteriaceae* (Meriweather *et al.*, 2013), ainsi qu'une bactérie du genre *Rickettsia* (Potts *et al.*, 2020). Aucun contig de protéobactérie BEV-like n'a été identifié dans les données pool-seq ici explorées, alors que quelques-uns semblent témoigner de la présence du genre *Rickettsia*, se rapprochant plus particulièrement de la *Rickettsia* symbiotique de la mouche des sables *Culicoides newsteadi* (n=17, t_c =119 kb). Néanmoins, il faut souligner que le génome de la bactérie *Rickettsia* inféodée à la punaise de lit n'était pas encore publié lorsque la base de données nr a été téléchargée, puisqu'il ne l'a été que courant 2022 (Voir Davison *et al.* 2022, GenBank ID CP084572.1, taille=1,6 Mb), ce qui peut expliquer le faible nombre de contigs détectés ayant un pourcentage d'identité supérieur à 90%. Quant au génome de la protéobactérie BEV-like de *C. lectularius*, il n'a à notre connaissance toujours pas été publié.

Le seul contig viral détecté est celui d'un *Siphoviridae* (taille=1074 pb, %ID=90,2), un virus à ADN double brin linéaire d'une cinquantaine de kilobases de long. Ce phage, de l'ordre des Caudovirales, a pour hôte naturel les bactéries et les archées plutôt que les insectes (Brüssow and Desiere, 2001).

Ensemble, ces analyses préliminaires ont permis d'identifier des micro-organismes à ADN présents chez des punaises au phénotype *a priori* sain. Dans un second temps, l'ADN et l'ARN de punaises porteuses du phénotype oxydé a donc été analysé.

4.3 Analyse des punaises de lit au phénotype oxydé

Le phénotype oxydé a touché toutes les souches que nous avons en élevage au sein du laboratoire (Souches Cimexstore : London Lab, London Field, German Lab, Sweden Field ; Souches obtenues dans le cadre d'un partenariat avec O.Otti :

F4¹, S1²). Aussi, il est très compliqué d'avoir des contrôles négatifs. Lors de cette analyse, nous avons donc choisi de réaliser uniquement un contrôle positif, afin de détecter des candidats potentiels.

4.3.1 Extraction et séquençage des acides nucléiques viraux putatifs

À ce stade, notre hypothèse principale était la présence d'un virus, similaire à celui de l'hypertrophie des glandes salivaires, provoquant une forte mortalité dans les élevages de mouche tsé-tsé. Ainsi, nous avons utilisé un protocole de purification virale reposant sur des filtrations et traitements enzymatiques.

Insectes

Afin d'identifier de potentiels virus pathogènes pour la punaise de lit, des punaises de lit présentant le phénotype de mortalité oxydé ont été congelées à -80°C dès les premiers signes d'une telle mortalité imminente (individus toujours gorgés quelques jours après le repas de sang qui commencent à foncer du rouge au noir). Ces punaises étaient de différents stades (stade 1 à adulte), sexes (M/F), et souches (London Lab/London Field).

Broyage et filtration

Douze lots de 45 mg d'insectes ont ensuite été placés avec une bille dans 500 µL de tampon GBVPE-1X (Grinding Buffer for Viral Particles Extraction) à pH 7,5 (Concentration 1X : 20 mM Tris-HCl pH 7.5, 10 mM KCl, 4 mM MgCl₂, 6 mM NaCl, 1 mM Dithioerythrol, 0.1% Tween-20). Le but de ce tampon est de permettre le maintien de l'intégrité des capsides virales, tout en détruisant les bicouches lipidiques composant les membranes cellulaires, nucléaires et mitochondriales. Les échantillons ont été homogénéisés 30 sec à 30 Hertz avec le broyeur à bille (Tissue-Lyser, Qiagen) afin d'en extraire l'ADN et l'ARN. Les échantillons ont été centrifugés 2 min à 800 g à 4°C. Les surnageants obtenus ont été collectés, puis chaque culot a été rincé avec 300 µL de GBVPE-1X. Deux centrifugations supplémentaires ont permis de collecter à chaque fois le surnageant restant. L'échantillon ainsi obtenu a été poussé par une seringue

1. La souche F4 a été échantillonnée à Londres en 2008, en même temps que la souche London Field de Cimextore.

2. La souche S1 a été échantillonnée dans les années 1970 à Londres également, probablement en même temps que la souche London Lab de Cimextore.

hydrophile à travers deux filtres successifs de taille 5 microns, puis 1,2 μm (Minisart, Sartorius). Le filtre a ensuite été rincé avec 3 mL de tampon GBVPE-1X, pour récupérer ce qu'il restait d'échantillon.

Digestion des acides nucléiques de *C. lectularius*

Le volume final de cet échantillon a été estimé ($V_f = 24 \text{ mL}$) afin d'ajouter : (i) de la RNase-A (Ribonucléase pancréatique bovine) pour une concentration finale 200 $\mu\text{g/mL}$ (soit 170 μL de RNase-A 20 mg/mL) et (ii) de la DNase-I (Désoxyribonucléase I pancréatique bovine, D5025, Sigma-Aldrich) pour une concentration finale de 0.1 unité/ μL (soit 34 μL de 50 U/ μL de DNase-I). L'objectif de ces traitements à la DNase et à la RNase est de dégrader respectivement l'ADN génomique et l'ARN total de l'insecte. L'échantillon a ensuite été incubé 2h à 20°C (température de la pièce), puis centrifugé 30 min à 13 000 g et 4°C pour précipiter les potentielles particules virales. Le surnageant a été retiré sans perturber le culot, puis ce culot a été nettoyé des DNase-I et RNase-A en ajoutant 20 mL de tampon GBVPE-1X. Une nouvelle centrifugation a été réalisée pendant 10 min à 13 000 g et 4°C, et le surnageant retiré.

Extraction des acides nucléiques restants

Les ADN et ARN viraux putatifs présents dans le culot ont été extraits séparément à l'aide du kit All-In-One DNA/RNA/Protein Miniprep (BioBasic).

Pour simplifier, nous parlerons d'ADN et d'ARN viraux par la suite (ADN_v et ARN_v), mais gardons à l'esprit que si ce protocole doit protéger puis digérer les supposées capsides virales, certaines cellules peuvent tout de même potentiellement y résister, comme les *Wolbachia* enveloppées de plusieurs membranes (Chang and Musgrave, 1973; Fischer *et al.*, 2014). De plus, le plus petit filtre utilisé n'ayant des pores que de 1,2 microns, il peut laisser passer certaines bactéries de petite taille, telle que *Wolbachia*, le symbiote obligatoire de la punaise de lit.

Le culot pesant 138,2 mg, 2100 μL de tampon de lyse DRP (BioBasic) y ont été ajoutés, suivant les recommandations du kit, puis séparé en quatre échantillons pour chaque type d'acides nucléiques. Chaque échantillon a été suspendu dans 50 μL de tampon CE pour l'ADN, et 30 μL d'eau RNase-free pour l'ARN.

Une fois les ARN_v et ADN_v extraits, un gel d'agarose à 1% avec TBE 0.5X (Tris, Borate, EDTA) a été réalisé afin de les visualiser (Figure 4.5). Le Ladder 1kb

Plus DNA (Invitrogen) a servi d'échelle. Pour l'ARNv, on observe plusieurs bandes représentant de gros fragments (entre 1 et 8 kb), ce qui pourrait correspondre au profil attendu pour un virus à ARN segmenté, tel qu'un réovirus (Shatkin *et al.*, 1968). Concernant l'ADNv, on observe une bande de haut poids moléculaire (taille >15 kb) qui pourrait traduire la présence d'un grand virus à ADN double brin, et/ou d'une contamination génomique.

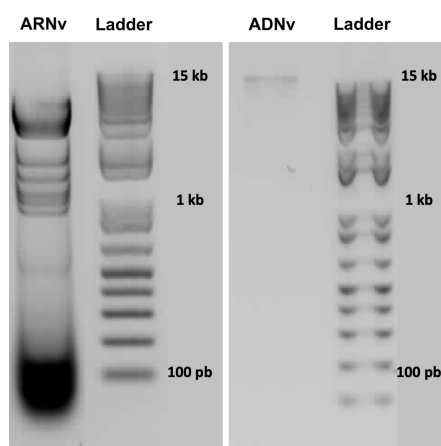


Figure 4.5 – Migration des échantillons d'ARN viral (ARNv) et d'ADN viral (ADNv) sur gel d'agarose.

Digestion enzymatique des échantillons obtenus

Afin d'écarter l'hypothèse d'une contamination avec de l'ADN génomique de la punaise dans l'échantillon ADNv, nous avons ensuite réalisé une digestion enzymatique avec EcoRI. Cette enzyme ayant de nombreux sites de reconnaissance sur l'ADN génomique (constitué de plusieurs mégabases), elle devrait le couper en de nombreux fragments, ce qui résulterait en l'observation d'un « smear » (tâche représentative d'ADN dégradé). En revanche, si notre hypothèse d'un génome viral est avérée, plusieurs bandes distinctes devraient être observées en raison de la faible taille du génome viral (typiquement < 15 kb dans le cas d'un réovirus).

Dans le but d'avoir un contrôle positif, nous avons extrait l'ADN et l'ARN génomique de quatre punaises de lit femelles (sans les étapes de purification virale) en utilisant le même kit Biobasic (concentrations Nanodrop : 13,6-28,2 ng/μL).

Les échantillons d'ADNg et d'ADNv ont ensuite été traités avec plusieurs enzymes : MspI (site de coupure C/CGG, enzyme 4 pb soit digestion toutes les 256 pb pour un génome à 50% de GC), MseI (site T/TAA, enzyme 4 pb) et EcoRI (site G/AATTC, enzyme 6 pb, soit une digestion toutes les 4 096 pb). La

digestion enzymatique de l'ADN génomique permet d'observer des smears assez nets, tandis que la digestion de l'ADN viral permet d'observer quelques bandes, en particulier à environ 10 kb et moins avec l'enzyme MspI (Figure 4.6). Une contamination à l'ADN génomique semble donc assez peu probable. Enfin, tous les échantillons d'ARN viraux ont été traités à la RNase pour vérifier qu'ils n'étaient ensuite plus visibles sur un gel d'agarose. Nous avons ainsi confirmé que ces échantillons étaient bien constitués exclusivement d'ARN, et qu'il n'y avait pas de contamination ADN.

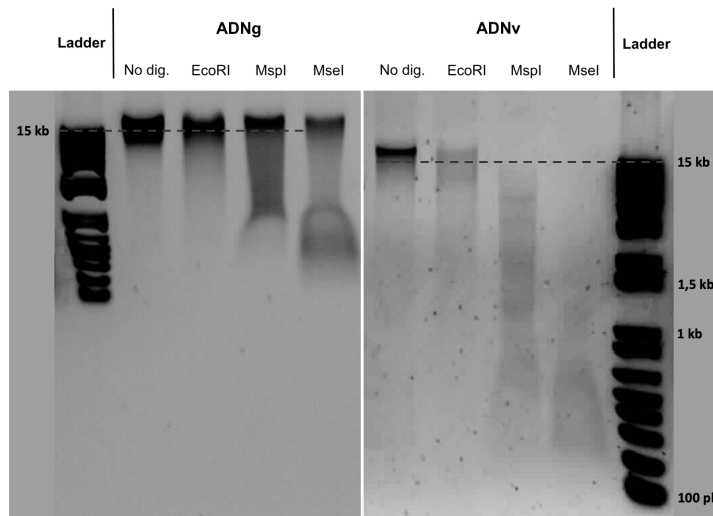


Figure 4.6 – Migration des échantillons d'ADN génomique (ADNg) et d'ADN viral (ADNv) sur gels d'agarose, avec digestion par plusieurs enzymes successives (EcoRI, MspI, MseI) ou sans digestion enzymatique (No dig.).

Les échantillons d'ARNv et d'ADNv ont ensuite été séquencés (Macrogen, Corée du Sud). Pour l'ARN, un sizing sur colonne a tout d'abord été réalisé, afin de retirer tous les petits fragments de moins de 200 pb (Figure 4.5). Le séquençage a été réalisé à l'aide du kit TruSeq Stranded Total RNA (Illumina, San Diego CA, USA) pour produire des reads paired-end de 2 x 150 pb. Concernant l'ADN, un kit TruSeq Nano LMW DNA a été utilisé (Illumina, San Diego CA, USA) pour produire également 2*150PE. Les échantillons ont été dosés après les différentes étapes (Table 4.1), et leur intensité visualisée sur des électrophérogrammes avec le système TapeStation (Agilent, Santa Clara CA, USA) (Supplementary figure 4.1).

	Concentration (ng/μL)	Nombre de reads
ARNv - avant sizing	402,8	-
ARNv - après sizing	228,6	19 803 544
ADNv	0,753	17 823 270

Table 4.1 – Concentration et nombre de reads séquencés pour les échantillons d'ARNv et ADNv extraits des punaises de lit présentant un phénotype oxydé. Le dosage a été réalisé par Macrogen avec un système TapeStation (Agilent, Santa Clara CA, USA).

4.3.2 Assemblage et alignement des contigs

Les reads ADN et ARN obtenus ont été nettoyés à l'aide de Trimmomatic (suivant les mêmes filtres que pour les données pool-seq) et respectivement assemblés en contigs à l'aide de Megahit, puis alignés sur nr (version de novembre 2019) avec MMseqs2 taxonomy (Voir partie 4.2, Figure 4.2). MMseqs2 taxonomy utilise un modèle dit du "dernier ancêtre commun" (Last Common Ancestor) basé sur le double blast (2bLCA, d'après Hingamp *et al.* 2013) pour assigner chaque contig à un nœud taxonomique (Figure 4.7).

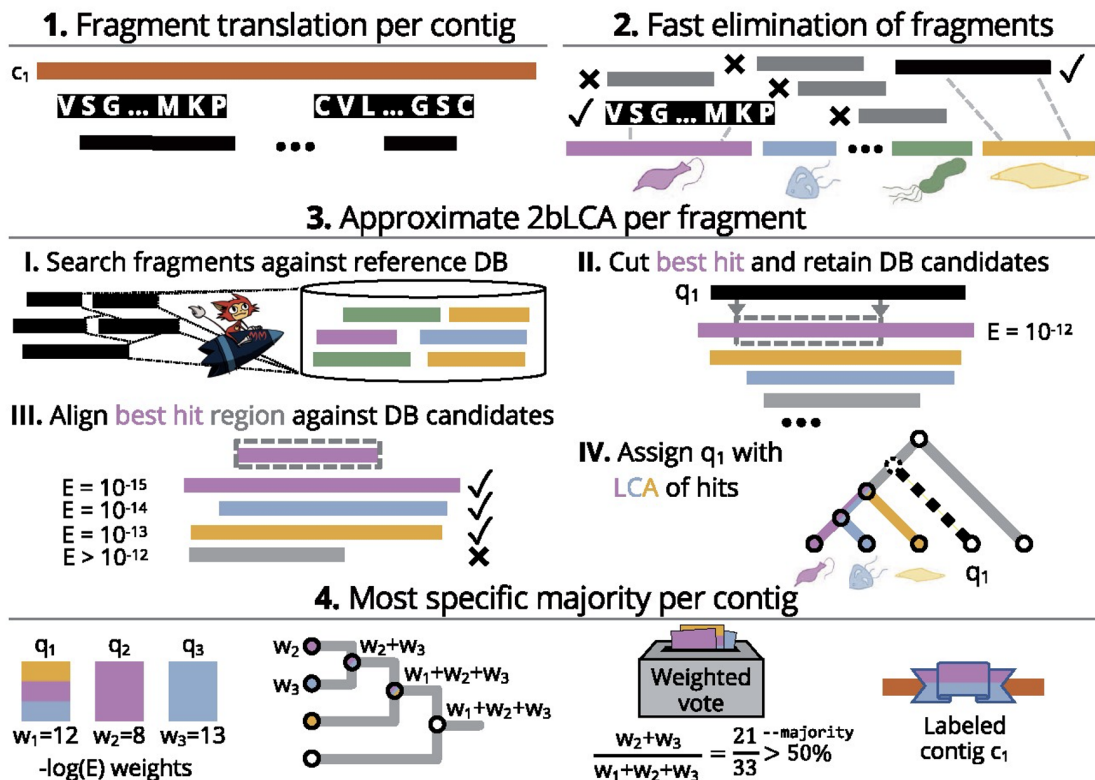


Figure 4.7 – Algorithme MMseq2 d'attribution taxonomique. Ces assignations sont faites en quatre étapes : (1) Traduire tous les fragments protéiques possibles pour tous les contigs ; (2) Rejeter les fragments peu susceptibles de trouver une correspondance taxonomique ; (3) Attribuer les nœuds taxonomiques. (I) Chaque fragment inconnu q (en noir) est recherché dans la base de données de référence. (II) Pour un fragment q_1 , cela donne ce qui donne une liste l de toutes ses cibles homologues. (III) La région alignée (encadrée en pointillés) entre q_1 et le meilleur résultat t (en rose), ayant la plus faible e-valeur (ici, 10^{-12}), est retenue, puis alignée sur toutes les séquences cibles dans l . (IV) On attribue à "q" l'ancêtre commun le plus récent des labels taxonomiques de toutes les séquences l ayant une e-valeur inférieure à celle de t ($<10^{-12}$) ; (4) Le nœud taxonomique du contig est déterminé en fonction des taxons associés majoritairement aux fragments sur ce contig (en fonction des e-valeurs associées).

Les résultats de MMseqs2 taxonomy peuvent être convertis afin d'être visualisé avec le package R Pavian v1.2.0. (Breitwieser and Salzberg, 2020). Dans un second

temps, les contigs ont été également annotés suivant la méthodologie classique MMseqs2 search. Seul le meilleur hit de chaque contig, ayant la e-value la plus faible, a ensuite été conservé en utilisant MMseqs2 filterdb.

4.3.3 Identification taxonomique

Les différents taxons identifiés sont donc détaillés, tout d'abord pour les contigs assemblés à partir des données ADN, puis ARN.

Contigs ADN

À partir des données ADN, 4 689 contigs ont été assemblés, dont 9,36% n'ont pas pu être assignés à un règne par mmseqs taxonomy. Le règne le plus représenté est le bactérien, avec 3 730 contigs, suivi des Eucaryotes avec 246 contigs et des Virus avec seulement 5 contigs (Figure 4.8).

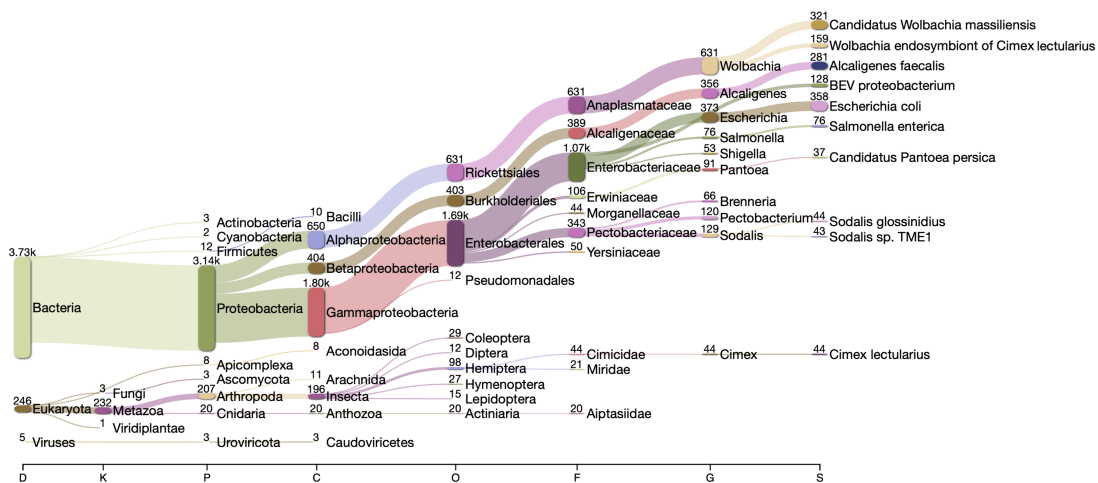


Figure 4.8 – Diagramme de Sankey des contigs ADN détectés.

Parmi les bactéries, la majorité des contigs détectés font partie du phylum des Protéobactéries, des bactéries à Gram négatif très diverses. Découpées en neuf classes, les Protéobactéries englobent des bactéries fermentaires, saprophytes, dénitrifiants, méthyliotrophes, fixatrices d'azote et pathogènes (Garrity *et al.*, 2006). Dans les données d'ADNv explorées, on retrouve les classes α , β , et γ -proteobacteria.

Les α -proteobacteria détectées sont quasi-exclusivement du genre *Wolbachia* (631 contigs/650), une bactérie connue pour former une symbiose mutualiste obligatoire chez *C. lectularius* Hosokawa *et al.* (2010). Concernant les

β -proteobacteria, on retrouve une majorité d'*Alcaligenes* (281 contigs/404), et plus précisément *Alcaligenes faecalis*. Enfin, les γ -proteobacteria détectées sont majoritairement de l'ordre des Enterobacterales, et plus particulièrement des Enterobacteriaceae (1 701 contigs), avec les espèces *Escherichia coli* (358 contigs) et la protéobactérie BEV (128 contigs).

E. coli est une bactérie fréquemment retrouvée dans les intestins des mammifères, en forme de bâtonnet. Certaines souches produisent des toxines les rendant pathogènes, et pouvant causer chez l'homme des maladies comme la gastro-entérite.

Récemment, une étude a montré qu'une symbiose mutualiste pouvait se mettre en place entre la punaise *Plautia stali* et la bactérie *E. coli* (Koga *et al.*, 2022). Au cours de cette étude, des nymphes de *P. stali* ont été stérilisées, les débarrassant ainsi de leur symbiote obligatoire *Pantoea*, puis infectées avec *E. coli*. Chez les *P. stali* survivantes, *E. coli* a été spécifiquement détectée dans les organes symbiotiques, et transmise verticalement à la descendance. Étonnamment, plusieurs lignées d'*E. coli* ont évolué indépendamment pour restaurer l'émergence élevée des adultes de l'hôte.

La seconde Enterobacteriaceae principalement détectée, la protéobactérie BEV, est un symbiote secondaire de la punaise de lit, encore mal décrit. Lors de sa détection chez *C. lectularius*, la bactérie BEV-like a été nommée ainsi, car elle était très proche (> 97.7% de similarité) du symbiote Bactérien d'*Euscelidius variegatus* (BEV), la cicadelle (Hemiptera, Cicadellidae). Cette bactérie se transmet de manière transovarienne entre générations de punaises de lit (Hypša and Aksoy, 1997). La symbiose de *C. lectularius* avec la bactérie BEV-like semble facultative, cette bactérie n'étant pas systématiquement retrouvée chez son hôte : seuls 53% des individus l'arboraient dans l'étude d'Hosokawa *et al.* 2010 (voir également Meriweather *et al.* 2013).

Parmi les γ -proteobacteria, de nombreuses bactéries de la famille des Pectobacteriaceae ont également été identifiées, avec les genres *Pectobacterium* (120 contigs) et *Sodalis* (129 contigs).

Le genre *Pectobacterium* est principalement connu pour être un pathogène de plantes, telles que *Solanum tuberosum*, la pomme de terre (Charkowski, 2015). Les bactéries du genre *Sodalis* en revanche sont devenues régulièrement et de manière indépendante des symbiotes de nombreuses espèces d'insectes, comme la

punaise *Cantao ocellatus* (Kaiwa *et al.*, 2010), le mélophage du mouton *Melophagus ovinus* (Chrudimský *et al.*, 2012), ou encore le cercope des prés *Philaenus spumarius* (Koga and Moran, 2014).

L'espèce principalement détectée parmi les contigs du genre *Sodalis* est *Sodalis glossinidius*, le symbiote secondaire d'un autre hématophage : la mouche tsé-tsé *Glossina pallipides*. La mouche tsé-tsé porte trois symbiotes : *Wigglesworthia glossinida*, *Sodalis glossinidius*, et *Wolbachia pipientis* (Balmand *et al.*, 2013). *S. glossinidius* est un symbiote secondaire, qui serait impliqué dans le développement et la fécondité de la mouche tsé-tsé (Sassera *et al.*, 2013), dont la taille du génome est de 4.3 Mb (Toh *et al.*, 2006).

Le but premier de cette analyse étant d'identifier de potentiels virus, les annotations des cinq différents contigs viraux sont détaillées ci-après : *Bos-associated insect adintovirus*, *Myoviridae*, *Pectobacterium phage ZF40*, *Plutella xylostella nucleopolyhedrovirus multiple*, et *phage phiSG1 Sodalis*. Tous ces contigs viraux appartiennent à des virus à ADN double brin (dsDNA), et sont de taille relativement petites (<1 kb, sauf le phage du *Pectobacterium*, de 1,9 kb). Trois de ces virus sont en réalité des phages, et infectent donc les bactéries (*Myoviridae*, *Pectobacterium phage ZF40*, *phage phiSG1 Sodalis*). Nous n'avons pas trouvé de littérature traduisant une potentielle pathogénicité du *Bos-associated insect adintovirus*. En revanche, le *nucleopolyhedrovirus multiple* détecté est un baculovirus pathogène de la teigne des crucifères *Plutella xylostella* (Harrison and Lynn, 2007). Les Baculovirus sont en effet connus pour infecter et tuer de nombreuses espèces de lépidoptères (Grzywacz, 2017), et utilisés à ce titre pour le bio-contrôle des ravageurs de certaines cultures, comme celle du pommier avec le granulovirus permettant de lutter contre *Cydia pomonella* (Eberle *et al.*, 2009). Il s'agit donc de notre meilleur candidat viral, bien que sa pathogénicité n'ait jamais été observée à notre connaissance chez les Hétéroptères.

Néanmoins, de nombreux contigs n'ont pas pu être assignés à une espèce, et en particulier les cinq plus grands contigs détectés (361 kb à 1,77 Mb). Pour y remédier, nous avons donc également analysé les meilleurs résultats de l'alignement avec mmseqs easy-search, en les filtrant sur les mêmes paramètres que pour les données pool-seq (>90% identité, bit score >50). De plus, le diagramme de Sankey précédemment présenté ne prends pas en compte la taille des contigs, et donc leur importance relative par rapport aux données. Les tailles relatives des génomes des espèces détectées ont donc été calculées en faisant la somme des paires de bases des contigs, en fonction de ces nouvelles assignations.

Les 10 plus grands sont annotés, et les autres regroupés dans une même catégorie ("others") (Figure 4.9). Les séquences retrouvées comme étant les plus abondantes appartiennent ainsi à *Alcaligenes faecalis*, avec 3,6 Mb (dont un contig de 1,77 Mb et 96,1% d'identité), sachant que la taille estimée de son génome est de 4,2 Mb³). Les deuxièmes séquences les plus abondantes sont celles de *Wolbachia*, le symbiote obligatoire de *C. lectularius*, avec une t_c de 872 kb, pour une taille de génome estimée à 1,2 Mb⁴). On retrouve ensuite la BEV protéobactérie, avec 770 kb, pour une taille de génome estimée à 3,8 Mb chez la bactérie originelle d'*E. variegatus* (Degnan *et al.*, 2011) Les deux candidats dont les estimations de taille semblent les plus proches sont donc *Alcaligenes faecalis* et *Wolbachia*.

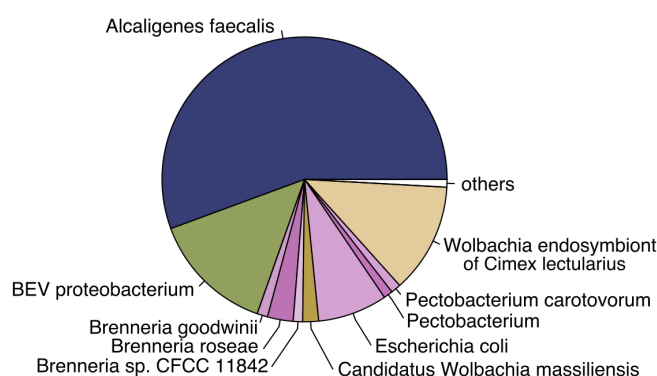


Figure 4.9 – Tailles cumulées des contigs détectés pour chaque espèce au sein des données ADN suite à la purification virale. Les 10 espèces les plus représentées sont annotées, et les autres sont regroupées dans "others".

Plusieurs grands contigs de *Brenneria*, une bactérie du genre *Pectobacterium* qui n'a pas été décrite précédemment, sont également identifiés. Cette bactérie étant connue comme phytopathogène (Naushad *et al.*, 2014), son identification pourrait s'expliquer par une proximité génomique avec la BEV protéobactérie, dont à peine 1/3 du génome est correctement détecté. Quant à *E. coli*, l'ensemble des contigs mesurés est de 462 kb, soit bien loin de la taille estimée de son génome de 4,6 Mb⁵.

Contigs ARN

Concernant les données ARN, 705 contigs ont été détectés, dont 7,66% n'ont pas de rang (non-classés). Parmi ces contigs, 137 appartiennent à des Bactéries, 495 à des Eucaryotes, et 3 à des Virus (Figure 4.10).

Une majorité de contigs ($n = 366$) semblent témoigner de la présence de *C.*

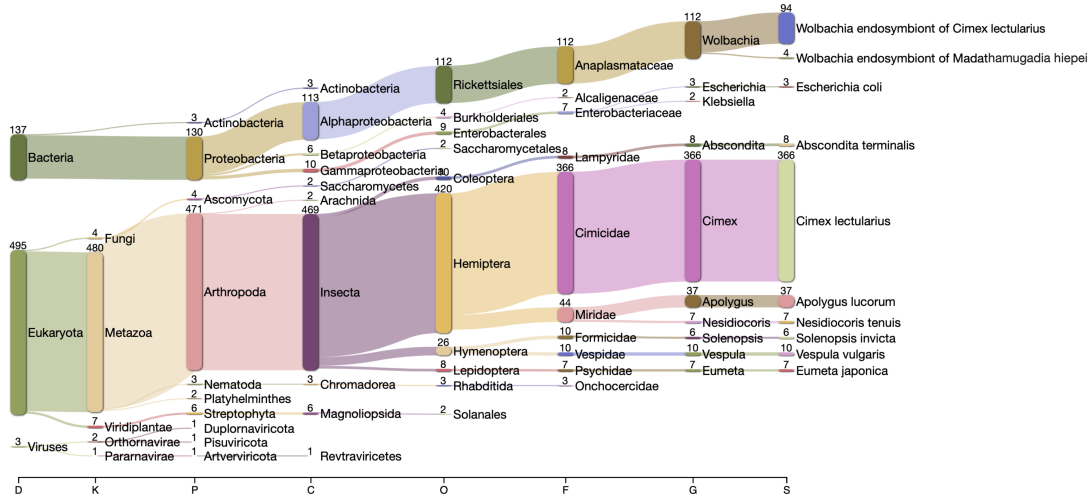
3. Assemblage du génome d'*Alcaligenes faecalis* : ASM96730v2

4. Génome de *Wolbachia* : ASM82931v1

5. Assemblage du génome d'*E.coli* : ASM584v2

lectularius. Chez les bactéries, on retrouve principalement des α -proteobacteria, et plus spécifiquement du genre *Wolbachia* (112 contigs).

Enfin, seuls 3 contigs viraux ont été identifiés dans ces données de "purification virale", comme appartenant aux genres réovirus, partitivirus et errantivirus. Ce dernier contig semble trop petit pour soutenir la présence d'un errantivirus (k141_77 = 338 pb), et correspond plus probablement à un artefact. En revanche, le contig de partitivirus détecté fait 1,4 kb de long, soit environ la moitié de la taille de ces génomes >3 kb à ARN double brin et découpés en deux fragments, et a un pourcentage d'identité relativement élevé (59,9%) (Vainio *et al.*, 2018).



4.11). Si les contigs les plus représentés sont ceux de *C. lectularius* et *Wolbachia*, on retrouve également dans ce top 10 le contig du réovirus décrit ci-dessous.

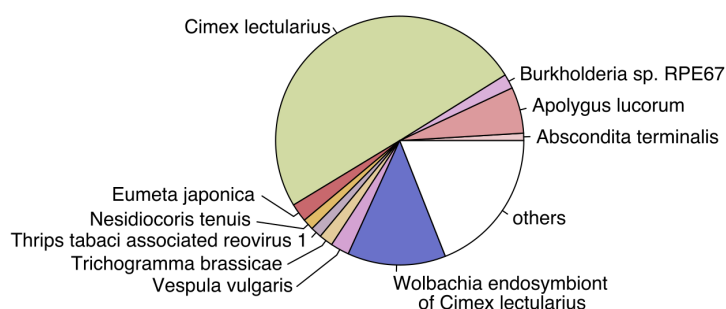


Figure 4.11 – Tailles cumulées des contigs détectés pour chaque espèce au sein des données ARN suite à la purification virale. Les 10 espèces les plus représentées sont annotées, et les autres sont regroupées dans "others".

4.4 Validation de candidats

Une étape de validation par PCR a ensuite été entreprise, afin de confirmer la présence de certains agents pathogènes candidats dans les élevages de punaises de lit. Les organismes se détachant des données bio-informatiques sont donc la bactérie *Alcaligenes faecalis*, le virus à ADN *Plutella xylostella nucleopolyhedrovirus multiple*, et le réovirus à ARN. L'objectif était également de vérifier si certains de ces candidats étaient détectés de manière différentielle (présence/absence) entre les extraits de punaises de lit au phénotype sain ou oxydé.

4.4.1 Préparation des échantillons

L'ARN et l'ADN ont été extraits à partir des mêmes quatre individus sains (DNA-sn 1 à 4 et RNA-sn 1 à 4) et oxydés moribonds (DNA-ox 1 à 4 et RNA-ox 1 à 4). Après une lyse des individus entiers 30 secondes à 30 Hz dans 350 µL de tampon Tissue Lyser, le kit BioBasic All-in-One présenté précédemment a été utilisé pour tous les extraits. L'ADN final a été suspendu dans 50 µL de tampon CE, l'ARN final dans 30 µL d'eau RNase-free, et tous les échantillons ont été dosés au Nanodrop (Table 4.2). Une transcription inverse a été réalisée sur les échantillons d'ARN avec le SuperScript™ IV VILO™ Master Mix, contenant des amorces aléatoires (random primers), afin de ne pas rater les ARNs dépourvus de queue poly A, cette dernière étant spécifique des eucaryotes (Iturriza-Gomara *et al.*, 1999).

Nous avons ensuite besoin de témoins positifs de PCR. Pour cela, nous

souhaitions utiliser les volumes restants des échantillons d'ADN et ARN viral purifiés envoyés à MacroGen pour le séquençage.

Échantillon	Concentration ADN	Concentration ARN
sn-1	13,6	234,6
sn-2	25,4	409,9
sn-3	28,2	285,4
sn-4	14,1	304,9
ox-1	4,2	312,0
ox-2	2,8	196,5
ox-3	2,4	-
ox-4	3,3	-

Table 4.2 – Concentration (en ng/μL) des échantillons d'ADN génomique et d'ARN total extraits des punaises de lit présentant un phénotype sain (sn) ou oxydé (ox).

Néanmoins, si une fraction de l'échantillon d'ARN a pu être conservée, le premier échantillon d'ADN envoyé a quant à lui été perdu par le transporteur. L'entièreté du volume d'ADN restant a donc déjà été utilisé en remplacement. En revanche, le surnageant pré-extraction des acides nucléiques viraux putatif avait été conservé à -80°C. Ce surnageant a donc été de nouveau centrifugé 30 min à 13 000 g et 4°C, avant d'extraire l'ARN et d'ADN du culot et du nouveau surnageant produit, toujours selon le kit All-In-One de BioBasic (sup-1 à 3 pour l'ADN, sup-1 et sup-2 pour l'ARN). Néanmoins, les quantités d'acides nucléiques contenues dans ces échantillons étaient toutes extrêmement faibles (<1,5 ng/μL pour l'ADN, <1 ng/μL pour l'ARN après dosage au Nanodrop), et aucune amplification par PCR n'a fonctionné.

4.4.2 Réalisation de la PCR

Les amorces de *C. lectularius* (protéine ribosomale RPL18) et *Wolbachia* (Fisher *et al.*, 2019) ainsi que de la bactérie BEV-like (Hosokawa *et al.*, 2010) sont tirées de la littérature. Des amorces ont ensuite été conçues avec Primer Blast pour nos pathogènes candidats⁶, en prenant garde qu'elles soient bien spécifiques, en vérifiant qu'elles ne s'alignent ni sur le génome de *C. lectularius*, ni sur celui de *Wolbachia* (Tableau supplémentaire 4.1).

Tous les échantillons d'ADN et d'ADNc (ARN post-transcription inverse) ont ensuite été dilués au 1/20^e, et un mix-PCR a été préparé. Ce mix était composé pour chaque échantillon de 18,9 μL d'eau ultrapure, 2,5 μL de tampon 10X, 0,5 μL

6. Primer Blast est un outil disponible sur le site NCBI pour trouver des amorces spécifiques : <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>

d'amorces A1, A2, et de dNtp afin de former les brins d'ADN complémentaires, et 0,1 µL de DreamTaq DNA polymérase. Les températures de demi-dénaturation (T_m) des amorces ont été ajustées afin qu'elles soient les plus appropriées pour la Taq polymérase utilisée⁷. Le témoin positif utilisé pour les données ADN est un échantillon ayant servi au séquençage en pool du chapitre 2, un individu de la souche London Lab. Le témoin positif des données ARN est une fraction de l'échantillon d'"ARN viral" ayant été séquençé par Macrogen (après transcription inverse).

4.4.3 Migration sur gel d'agarose

Suite à la PCR, une migration sur gel d'agarose a été réalisée afin de visualiser les amplicons des échantillons. Les contrôles positifs correspondant à *C. lectularius* et au symbiote bactérien obligatoire *Wolbachia* ont été détectés avec succès dans tous les échantillons (Figure 4.12).

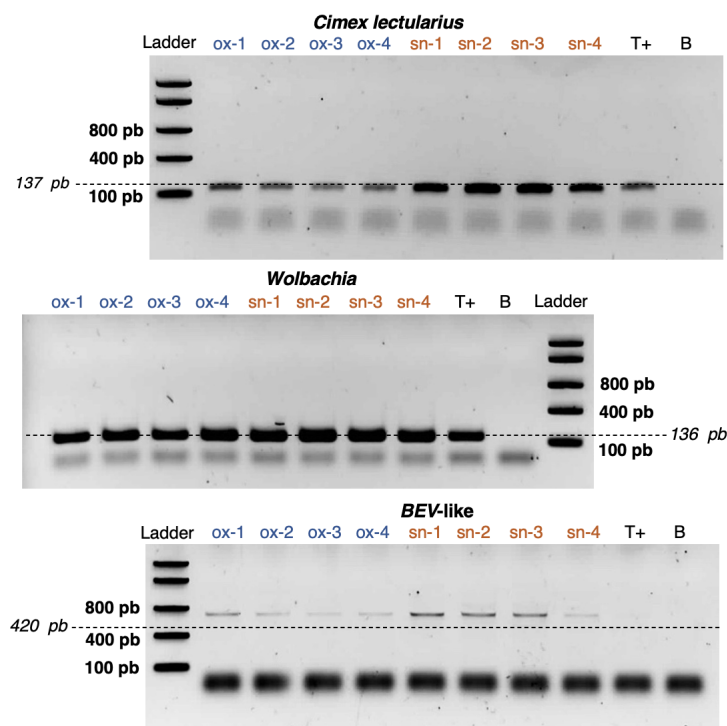


Figure 4.12 – Migration sur gel des produits de PCR pour différentes amorces : *Cimex lectularius*, *Wolbachia*, et BEV-like proteobacterium. Les échantillons d'ADN utilisés proviennent de punaises au phénotype oxydé (ox, en bleu) ou sain (sn, en orange). Le témoin positif (T+) est une punaise de la souche London Lab, utilisée lors du séquençage pool-seq. Un contrôle négatif (B) ne contenant pas d'ADN a également été réalisé. La taille d'amplicon attendue est indiquée avec une ligne discontinue.

7. Vérification de la T_m : <https://www.thermofisher.com/fr/fr/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html>

Le symbiote BEV-like quant à lui a été détecté dans les échantillons oxydés et sains, mais pas dans le témoin positif, ce qui est logique au vu de son absence lors de l'analyse des données pool-seq (voir section 4.2.2). Tous les ADN ayant été dilués au 1/20^e, et les échantillons sains étant entre trois et dix fois plus concentrés que les échantillons oxydés (Table 4.2), les bandes observées pour ces derniers étaient relativement plus fines. Aucune bande n'a été observée pour les échantillons contrôles négatifs (blanc de PCR). Néanmoins, aucune bande ne semble être à la hauteur prévue pour ce symbiote, ainsi que pour les pathogènes candidats (Figure 4.13).

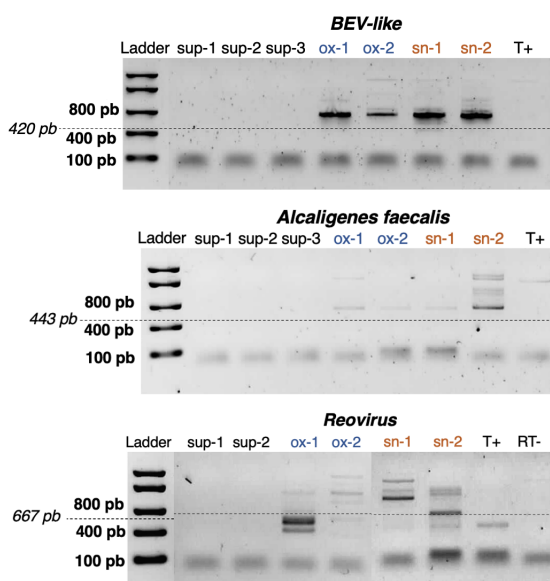


Figure 4.13 – Migration sur gel des produits de PCR pour différentes amorces : BEV-like proteobacterium, *Alcaligenes faecalis* et réovirus. Les échantillons d'ADN (BEV et *Alcaligenes*) et d'ARN (réovirus) utilisés proviennent du surnageant de l'échantillon "purification virale" ensuite envoyé au séquençage (sup), de punaises au phénotype oxydé (ox, en bleu) ou sain (sn, en orange). Le témoin positif (T+) est une punaise de la souche London Lab pour l'ADN, et une fraction de l'échantillon de "purification virale" pour l'ARN. Un contrôle négatif de transcription inverse a également été réalisé (RT-). La taille d'amplicon attendue est indiquée avec une ligne discontinue.

Néanmoins, quatre bandes semblent se détacher pour les échantillons oxydés et sains pour l'*Alcaligenes*, à environ 800 pb. Le même décalage est observé pour la détection de la bactérie BEV-like (Figure 4.12 et 4.13), et pourrait s'expliquer par un problème venant de l'échelle (Ladder) utilisée. Toutefois, la présence d'autres bandes pourrait également traduire un manque de spécificité des amorces, qui amplifient également d'autres fragments.

Comme expliqué précédemment, les échantillons ré-extraits du surnageant (appelés sup-1 à 3) n'ont pas pu être amplifiés avec succès, du fait de leur faible concentration. Nous n'avons donc pas réellement de témoin positif pour la

présence de l'*Alcaligenes*, étant donné que même s'il a été observé dans les données pool-seq, il s'agissait d'analyses bio-informatiques conduites sur l'ensemble des individus des quatre souches, ce qui ne reflète donc pas nécessairement sa présence dans chaque individu.

Concernant le réovirus, une seule bande (un peu plus basse que l'amplicon prévu) est détectée au niveau du témoin positif, et semble également présente chez les individus ox-1, sn-1 et sn2, pouvant traduire la détection de ce virus (Figure 4.13). Enfin, aucune bande n'a été observée avec les amorces du *nucleopolyhedrovirus*, ni pour les témoins négatifs (blanc) de PCR (données non montrées). Afin de valider la détection de ces différents candidats, les produits PCR devront être séquencés, et leur homologie avec des séquences connues de ces candidats de nouveau analysée par Blast.

4.5 Discussion & Perspectives

Ce chapitre de thèse visait à détecter de potentiels candidats engendrant la forte mortalité observée au sein de nos élevages de punaises de lit, présentant un phénotype oxydé caractéristique. Après avoir discuté les potentiels agents détectés, nous proposerons plusieurs perspectives à ces travaux, utilisant aussi bien des outils bio-informatiques qu'expérimentaux et relatifs à la gestion de l'élevage.

4.5.1 Potentiels agents pathogènes détectés chez la punaise de lit

Une synthèse des principales espèces bactériennes et virales détectées dans les différentes séquences ADN explorées a tout d'abord été réalisée (Tableau 4.3). Le symbiote obligatoire *Wolbachia* a été identifié dans les deux jeux de données. La bactérie BEV-like, identifiée uniquement dans les données de purification virale, a été précédemment décrite comme étant un symbiote secondaire. Quant aux bactéries *Rickettsia* et *Paenochrobactrum*, elles n'ont été identifiées que dans les données pool-seq, soit le contraire de l'attendu pour une bactérie impliquée dans le phénotype oxydé.

Quoique jamais décrite auparavant chez la punaise de lit, la présence d'un *Alcaligenes* dans les données semble fortement probable, avec la totalité du génome couvert (taille estimée à 4,2 Mb). Enfin, *E. coli* a été observée lors des

deux alignements, avec une abondance néanmoins nettement supérieure chez les punaises oxydées. Toutefois, il est difficile de comparer ces abondances et d'en tirer des conclusions quelconques. En effet, nous n'avons aucune idée de la dynamique potentielle de répllication de l'agent pathogène au sein de l'insecte - nous ne pouvons qu'émettre l'hypothèse que ce potentiel pathogène serait présent de manière plus importante chez les insectes présentant des signes visibles de mortalité imminente. Plutôt que des PCR, il faudrait donc réaliser des PCR quantitatives, permettant de mesurer l'abondance relative des candidats. Il en va de même pour notre meilleur candidat ARN, le réovirus détecté.

Espèces d'intérêt		Pool-seq			Purification virale		
		Contigs	t_c	%ID	Contigs	t_c	%ID
Bactéries	<i>Alcaligenes faecalis</i> *	661	4,4 Mb	94,44	219	3,6 Mb	95,18
	<i>Paenochrobactrum gallinarum</i>	1,263	1,7 Mb	94,19	-	-	-
	<i>Wolbachia endosymb. C. lectularius</i> *	236	1,1 Mb	94,96	168	872 kb	95,24
	<i>Wolbachia massiliensis</i>	611	868 kb	95,20	232	112 kb	95,58
	<i>Rickettsia endosymb. Cul. newsteadi</i>	17	119 kb	93,51	-	-	-
	<i>Escherichia coli</i>	13	18 kb	96,86	912	462 kb	95,50
	BEV-like *	-	-	-	96,00	770 kb	93,69
Virus	Siphoviridae	1	1074 pb	90,20	-	-	-
	<i>P. xylostella</i> nucleopolyhedrovirus m.	-	-	-	1	878 pb	95,90

Table 4.3 – Résumé des principales espèces d'intérêt à génome ADN détectées dans les données métagénomiques. Les données issues du séquençage en pool sont séparées des données ADN de "purification virale" sur punaises présentant un phénotype oxydé. Le nombre de contigs, la taille cumulée (t_c) de ces contigs ainsi que le pourcentage d'identité moyen (%ID) sont donnés pour chaque espèce dans chaque jeu de données (mmseqs search, bit score >50, pourcentage d'identité >90%). Les espèces dont la présence a été validée par PCR sont identifiées par une étoile.

La couverture a également été calculée pour chacun des contigs analysés dans ce chapitre, sans que cette donnée n'ait pour l'instant été exploitée. De futures analyses pourraient utiliser ces données de couvertures pour estimer l'abondance relative des différents candidats, ou tenter de détecter des virus en fonction de leur couverture relativement aux contigs d'insectes. En effet, le réovirus détecté a potentiellement plusieurs autres fragments présents dans l'ARN séquencé à partir des extraits de purification virale, peut être trop divergents des réovirus dont les génomes sont déjà répertoriés, et n'ayant donc pas été identifiés lors de l'alignement sur nr.

4.5.2 Exploration des bases de données SRA

Internet regorge de données prêtes à être analysées, encouragé par un nombre croissant de travaux publiés en "accès libre". En janvier 2021, un total de plus de 900 jeux de données étaient référencés pour *Cimex lectularius* sur la base de

données SRA ⁸. Parmi ces jeux de données, 716 sont par exemple référencés comme "métagénomés", et 21 sont issus de séquençages ARN. De plus, dans le cadre de cette thèse, nous avons produit quatre jeux de données de génomes complets en pool-seq (groupe de 30 individus) pour les quatre souches Cimexstore explorées (London Lab, German Lab, London Field et Sweden Field), mais également 96 séquençages ARN pour ces mêmes souches (soit 24 échantillons par souche). Un séquençage long-reads a également été réalisé sur 5 individus de la souche London Lab.

Toutes ces données pourront être exploitées afin de rechercher la présence de virus et/ou de bactéries dans des punaises de lit de différents élevages, et parfaire l'assemblage des génomes de micro-organismes déjà décrits, et en particulier le réovirus et l'*Alcaligenes* détectés.

4.5.3 Gestion de la mortalité en laboratoire

Plusieurs méthodes peuvent permettre de réduire l'impact négatif d'un potentiel pathogène sur l'élevage. Tout d'abord, s'occuper d'un élevage en laboratoire implique un protocole strict de désinfection régulière du matériel et la mise en place de conditions de travail les plus stériles possibles.

Dans le cas d'insectes hématophage, ces consignes s'appliquent tout particulièrement au système d'alimentation, afin d'éviter de potentielles contaminations lors du repas de sang. La méthode du "Système d'Alimentation Propre" (CFS pour Clean Feeding System) a notamment prouvé son efficacité dans la gestion du virus de l'hypertrophie des glandes salivaires, qui décimait des colonies entières de la mouche tsé-tsé *Glossina pallidipes* élevée en laboratoire (Abd-Alla *et al.*, 2010). En effet, la transmission de ce virus est horizontale : les particules virales sont relâchées par les insectes infectés dans le réservoir de sang dans lequel leurs congénères non-infectés se nourrissent, et vont donc les contaminer (Figure 4.14). Un seul individu infecté peut excréter jusqu'à 10^7 particules virales infectieuses dans un réservoir de sang en 10 à 15 minutes de repas (Abd-Alla *et al.*, 2010). Le CFS consiste à fournir différents réservoirs de sang à différentes cages de mouches tsé-tsé à chaque repas, pour éviter les contaminations croisées, et a permis de réduire significativement la charge virale de $10^{7.4}$ à $10^{2.4}$ copies en moyenne dans la colonie (Abd-Alla *et al.*, 2013).

8. <https://www.ncbi.nlm.nih.gov/sra/?term=cimex+lectularius>

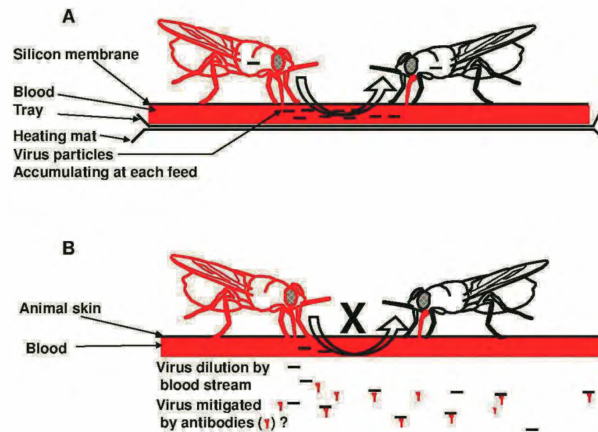


Figure 4.14 – Transmission horizontale du virus de l’hypertrophie des glandes salivaires chez *G. pallidipes*. (A) Repas de sang *in vitro* sur une membrane utilisé dans l’élevage de masse de la mouche tsé-tsé; (B) Simulation du repas de sang *in vivo* sur animaux vivants. En rouge, les mouches infectées avec le virus; en noir, les mouches non-infectées (Abd-Alla *et al.*, 2011).

Une fois l’agent pathogène identifié, l’ajout d’antibiotiques (qui n’auraient pas d’effet sur le symbiote obligatoire *Wolbachia*) ou d’antiviral ciblé au repas de sang peut permettre de maîtriser la mortalité observée en laboratoire chez la punaise de lit. Certains antiviraux, comme l’aciclovir et le valaciclovir, peuvent être utilisés à des concentrations qui ne nuisent pas à la synthèse de l’ADN de l’hôte (300 µg/mL de sang), tout en étant convertis en métabolites actifs par l’enzyme thymidylate synthase codée par le virus (Abd-Alla *et al.*, 2012). Ces antiviraux, en réduisant la répllication de l’ADN viral, sont particulièrement efficaces pour empêcher la prolifération du virus chez les insectes avant qu’ils ne présentent les premiers signes d’infection. En effet, une fois les premiers signes de l’infection visibles, le virus s’est déjà répliqué de manière suffisamment importante, et l’utilisation de ces antiviraux n’est plus indiquée (Abd-Alla *et al.*, 2021).

Chez *G. pallipies*, la gestion des infections virale est donc réalisée en (1) désinfectant strictement l’environnement de travail, l’équipement et les cages des insectes, (2) suivant l’apparition des symptômes (dissection régulières pour observer la présence de glandes salivaires hypertrophiées, diagnostic PCR), (3) mettant en quarantaine les colonies infectées/conservé une colonie "de secours" séparée, présentant une faible prévalence (Abd-Alla *et al.*, 2021).

Au laboratoire, les élevages de punaises de lit sont désinfectés de manière stricte (autoclave des réservoirs, stérilisation des paillasse et du matériel) et les différents pots des punaises de lit sont nourris sur des réservoirs séparés (acquisition de réservoirs aux volumes plus restreints - 0.5 mL au lieu de 3 mL), mais cela ne suffit pas à éviter de très forts pics de mortalité ponctuellement. Un

laboratoire avec qui nous avons été en contact et qui nourri ses punaises sur bras humain a reconnu observer une mortalité similaire, présente de manière régulière, mais toujours en très faible pourcentage (<2-5% des individus). Cette opposition diamétrale de notre expérience et la sienne nous encourage à poursuivre les efforts pour détecter un potentiel agent se transmettant via le repas de sang. Une fois l'agent identifié de manière formelle, un séquençage du sang post-repas pourrait également être réalisé, afin de confirmer ce mode de transmission.

4.6 Matériel supplémentaire

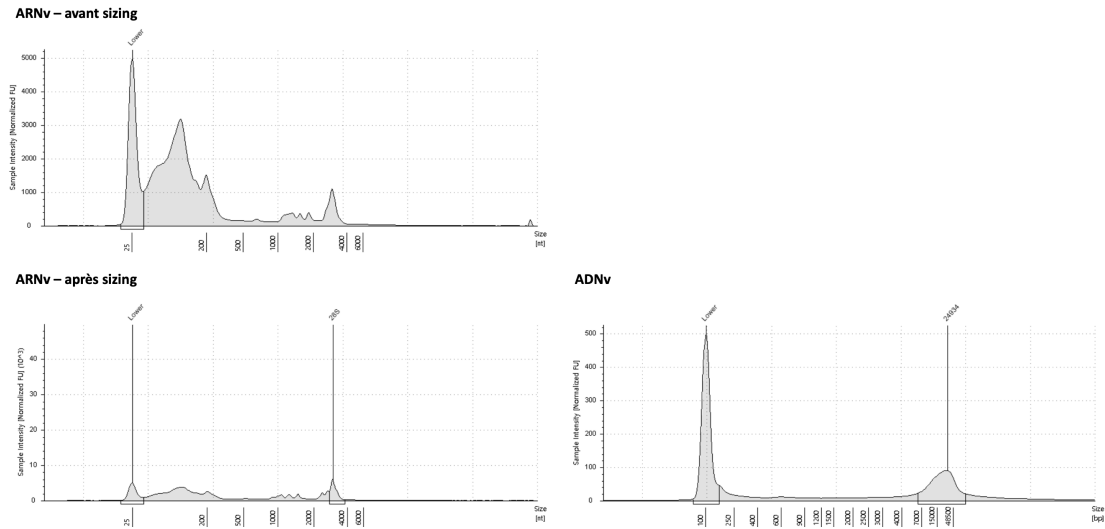


Figure supplémentaire 4.1 – Intensité de l'échantillon en fonction de sa taille. À gauche - ARNv, avant et après sizing en dessous, à droite - ADNv.

Cible	Amorces	Amplicon	Tm
<i>Cimex lectularius</i>	F - GTA TGA CGG AGG CAG CTA GG R - AAC ATT CGA GCA AAT TCG GTA	137	56
<i>Wolbachia</i>	F - AGT CAT CAT GGC CTT TAT GGA R - TCA TGT ACT CGA GTT GCA GAG T	136	60
BEV-like	F - GCA CAA GGG AGG TTG CTC CCC R - CAG CAA GGT TAT TAA CCT TAC TG	420	57
<i>Alcaligenes faecalis</i>	F - AGT GCC GTC GAT GCC TTT TA R - TTG CGC AAT ACA TCG AAG CG	443	57
<i>P. x. nucleopolyhedrovirus m.</i>	F - GGA CGG CAA ATC CGT GAC TA R - AGC CTG CTG TCG TGA ATA CC	553	60
réovirus	F - ATT GTG ATG GCC CTA ACG GG R - ACT GTG GTA CGG CTC TCT CT	667	59

Tableau supplémentaire 4.1 – Amorces utilisées pour les PCR de validation des virus et bactéries détectées dans les données métagénomiques. Pour chaque agent ciblé, le contig dont les amorces ont été extraites, ainsi que les séquences Forward (F) et Reverse (R) de ces amorces, la taille de l'amplicon produit (en pb) et la température de demi-dénaturation (Tm, en °C) sont données.

5

Discussion générale

5.1	Résumé des principaux résultats de la thèse	170
5.2	Distribution des candidats associés au phénotype de résistance le long du génome	174
5.2.1	Un superlocus impliqué dans la résistance aux pyréthri- noïdes	174
5.2.2	Superlocus en auto-stop ou évolution d'un supergène? . .	176
5.2.3	Organisation du génome	180
5.3	Mécanismes de résistance aux pyréthri- noïdes chez <i>C. lectularius</i>	184
5.3.1	Rares SNPs non-synonymes	185
	Importance de la mutation <i>kdr</i>	185
	Variations des séquences codantes de gènes induits par l'exposition aux pyréthri- noïdes	186
5.3.2	Régulation transcriptionnelle	188
	Différences constitutives d'expression entre les souches résistantes et sensibles	188
	Quid du cluster de P450 détecté sur le superlocus? . . .	190
5.3.3	Régulation post-transcriptionnelle	191
	Explorer l'implication des ARN non-codants dans la résistance aux insecticides	192
	Des ARNnc dans le supergène?	193
5.4	Comprendre pour faciliter la lutte de demain	195
5.4.1	Développement de tests diagnostics	195
5.4.2	La recherche dans de beaux draps!	198
5.5	Conclusion générale	200

De la création d'un élevage à l'analyse bio-informatique des données, en passant par l'extraction d'ADN et d'ARN et l'exposition aux insecticides, la complétude de ce projet a nourri ma créativité, mes compétences, et mon développement en tant que scientifique. Avant de commencer cette discussion, je souhaiterais souligner la satisfaction personnelle que m'a apporté ce projet, en étant actrice de chacune de ces étapes, aussi hétéroclites soient-elles.

5.1 Résumé des principaux résultats de la thèse

Mon travail a consisté à étudier la résistance aux insecticides chez la punaise de lit *Cimex lectularius* à l'échelle de l'entièreté du génome et sans *a priori*, en comparant les séquences mais également l'expression des gènes entre plusieurs populations aux statuts de résistance contrastés. En plus de participer à la compréhension globale de ces mécanismes de résistance, l'objectif de cette thèse était également d'identifier le potentiel agent pathogène engendrant une étrange mortalité au sein de l'élevage, et pouvant représenter une piste d'amélioration de la lutte contre cet insecte (Figure 5.1).

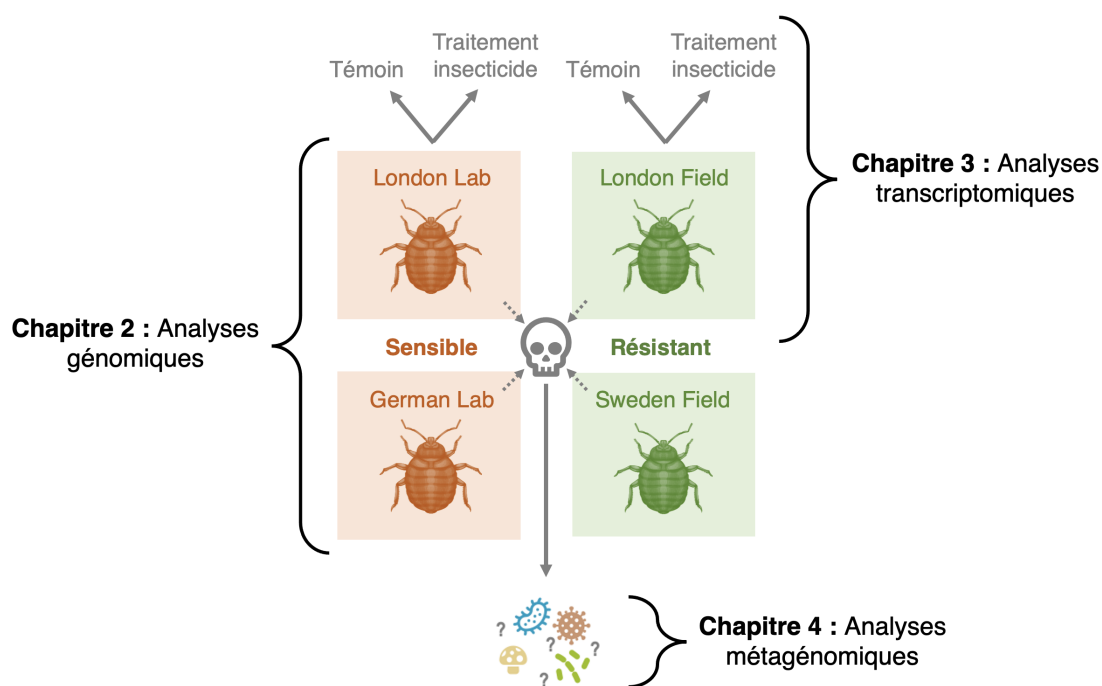


Figure 5.1 – Schéma de l'articulation des différents chapitres présentant les travaux réalisés dans le cadre de cette thèse.

Le chapitre 2 présente tout d'abord une analyse pool-seq des génomes de quatre populations de punaises de lit, aux statuts de résistance contrastés (Figure 5.2). Cette analyse a permis de détecter de nombreux SNPs outliers, fortement

différenciés entre les populations résistantes et sensibles. Un seul de ces variants a été détecté comme non-synonyme, la plupart étant identifiés dans les régions introniques ou intergéniques. Mais plutôt que d'être répartis de manière aléatoire le long du génome, un des chromosomes putatifs (LG 11) de la punaise de lit était très fortement enrichi en SNPs outliers, en particulier pour trois de ses scaffolds adjacents, formant un "superlocus" de 6 Mb. Ce superlocus portait un nombre significativement plus élevé de gènes impliqués dans la résistance aux insecticides (P450, GST) par rapport au reste du génome, et l'un des scaffolds le composant était également enrichi en variants structuraux, et plus particulièrement en inversions. L'hypothèse d'un supergène où les allèles impliqués dans la résistance aux pyréthrinoïdes seraient regroupés et maintenus associés une réduction de la recombinaison *via* une ou plusieurs inversions a ainsi été formulée.

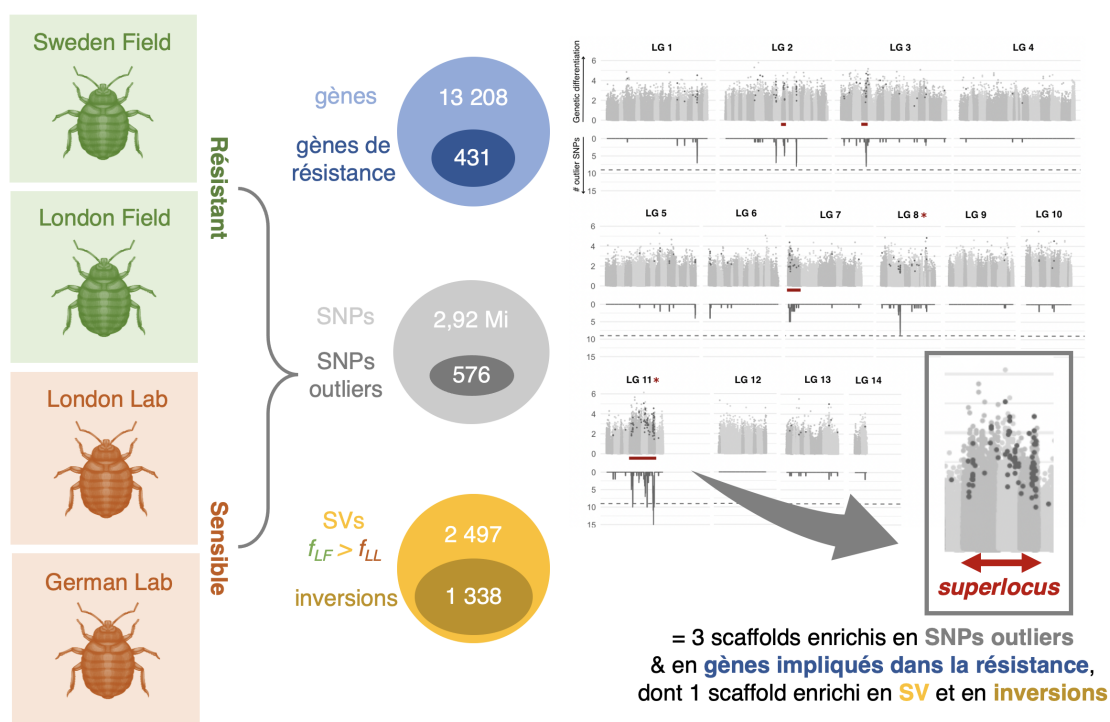


Figure 5.2 – Schéma bilan du chapitre 2 - analyses génomiques. Le nombre de gènes de résistance recherchés est donné par rapport au nombre de gènes total, ainsi que pour les SNPs outliers et les inversions (par rapport au nombre de variants structuraux, SV, ayant une fréquence supérieure chez London Field).

Le chapitre 3 consiste ensuite en une analyse transcriptomique de deux des quatre souches de punaise de lit précédemment étudiées d'un point de vue génomique, au fond génétique proche mais aux statuts de résistance opposés (Figure 5.3). Les gènes détectés comme différentiellement exprimés en fonction de la souche, mais également du traitement ou non aux pyréthrinoïdes, ont été analysés.

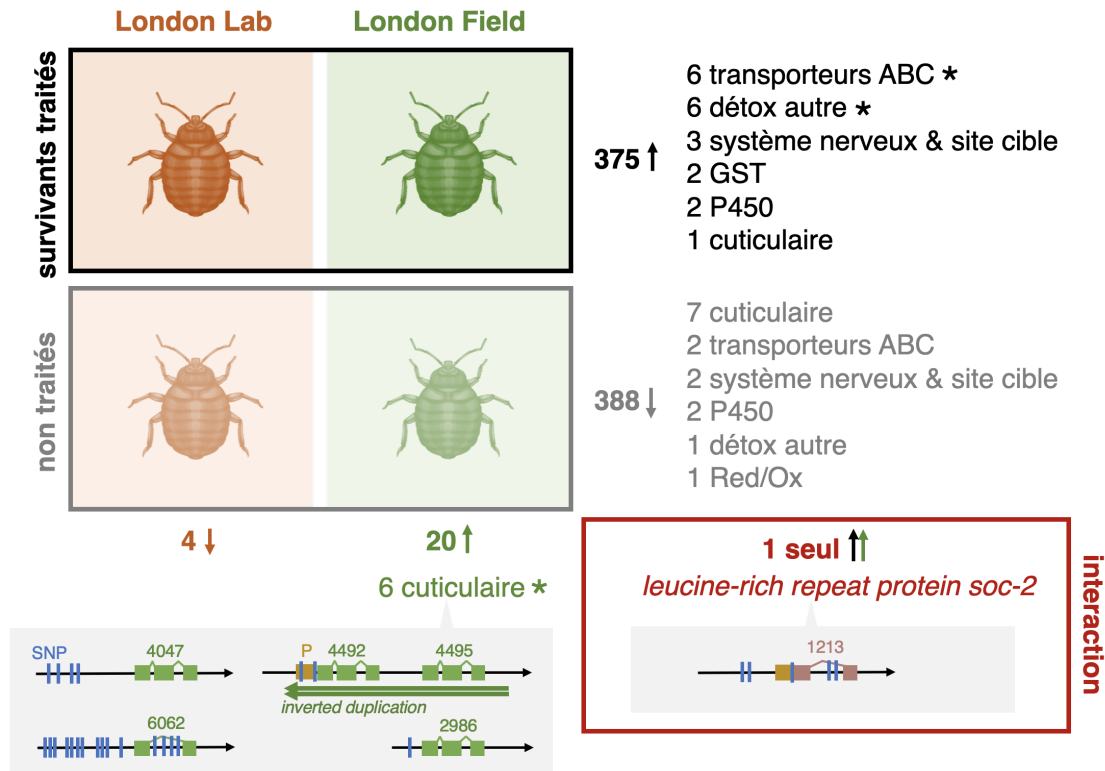


Figure 5.3 – Schéma bilan du chapitre 3 - analyses transcriptomiques. Le nombre de gènes significativement sur-exprimés et sous-exprimés pour les deux conditions (souche : London Lab/London Field et traitement insecticide : survivants traités/non traités) est donné, et détaillé en fonction des différentes catégories de gènes de résistance représentées. Ces catégories sont mises en évidence à l'aide d'une étoile si elles sont significativement enrichies. Concernant les gènes montrant un différentiel d'expression pour la condition "souche", les potentiels variants structuraux (flèche horizontale verte) et/ou SNPs différentiels (trait bleu) détectés à proximité lors de l'analyse génomique de ces souches sont annotés. Les quatre derniers chiffres de l'identifiant des gènes sont donnés (se référer à l'article pour retrouver l'identifiant complet). Les régions promotrices (en jaune) ne sont annotées que si des SNPs sont détectés à l'intérieur.

Parmi les gènes sur-exprimés au sein de la souche résistante par rapport à la souche sensible, on retrouve plusieurs protéines cuticulaires. Certaines font partie d'une duplication inversée ou présentent un ou plusieurs SNPs fortement différenciés à proximité, ces variants pouvant être responsables de la différence transcriptionnelle observée. Les gènes sur-exprimés chez les individus ayant survécu à l'exposition à la deltaméthrine étaient très nombreux, mais deux catégories de gènes de résistance se sont détachées comme étant significativement enrichies, dont les transporteurs ABC. Enfin, un seul gène a montré une interaction significative entre les conditions traitement et souche, suggérant que les souches diffèrent très peu dans leur réponse transcriptomique à une exposition à un pyréthrianoïde. À proximité de ce gène codant pour une *leucine-rich repeat protein*, de nombreux SNPs fortement différenciés ont également été détectés. En

revanche, nous n'avons pas identifié de signal particulier au niveau du superlocus précédemment détecté dans l'analyse de la différenciation génétique. En effet, l'ensemble des gènes différemment exprimés étaient répartis de manière relativement homogène le long du génome.

Enfin, le cortège de micro-organismes accompagnant la punaise de lit a été exploré dans le chapitre 4, afin de détecter de potentiels pathogènes. Si la présence de bactéries déjà décrites chez la punaise de lit a été mise en évidence (*Wolbachia*, *BEV-proteobacterium*, *Rickettsia*), celle d'une bactérie du genre *Alcaligenes* a été détectée de manière étonnement abondante dans toutes les données ADN. Un réovirus a également été identifié en explorant les données ARN. La détection de ces candidats ne semblant pas être limitée aux individus mourant avec un phénotype oxydé au sein des élevages, différentes perspectives ont été abordées, telle que la réalisation de PCR quantitatives afin d'estimer l'abondance relative des candidats, ou l'exploration de leur présence à plus large échelle *via* les bases de données accessibles librement en ligne.

L'évolution de la résistance aux insecticides semble être le facteur clé de la ré-émergence de la punaise de lit commune *C. lectularius* (Davies *et al.*, 2012). Nos travaux profitent de la mise à disposition d'un génome de référence pour proposer une analyse exhaustive et élargie à l'ensemble du génome des déterminants de cette résistance. Cette thèse a ainsi permis de compléter de précédentes études plus ciblées en apportant une vision précise des séquences, et des transcrits potentiellement impliqués dans la résistance aux pyréthrinoïdes.

Lors de la discussion générale, nous reviendrons sur la répartition des variants associés à la résistance aux pyréthrinoïdes le long du génome, avec en particulier l'évolution potentielle d'un supergène, et la nécessité d'améliorer l'assemblage disponible. Nous analyserons ensuite de manière plus précise les différents mécanismes de résistance aux pyréthrinoïdes, et discuterons les candidats détectés pour chacun de ces mécanismes, avant de revenir sur les principales limites identifiées à cette thèse. Enfin, nous explorerons diverses perspectives pour la lutte contre cet insecte à plus long terme, passant par le développement de tests diagnostics, et l'exploration de nouvelles méthodes de bio-contrôle.

5.2 Distribution des candidats associés au phénotype de résistance le long du génome

La distribution des candidats associés à la résistance aux pyréthriinoïdes, loin d'être aléatoire, montrait plutôt un signal fort de manière très localisée. En effet, le principal résultat obtenu lors de l'analyse génomique des données est l'identification d'un superlocus de 6 Mb, présentant une forte différenciation génétique en fonction du statut de résistance aux pyréthriinoïdes des souches.

5.2.1 Un superlocus impliqué dans la résistance aux pyréthriinoïdes

Le superlocus de 6 Mb identifié dans nos données génomiques contenait une large majorité de gènes communs à la région de 10,8 Mb identifiée par une précédente étude sur les mêmes souches. Dans cette étude, [Fountain et al. \(2016\)](#) ont construit une carte de liaison génétique et réalisé une étude d'association (QTL mapping) pour la résistance aux pyréthriinoïdes chez *C. lectularius*. Un séquençage de marqueurs RAD pour des recombinants de génération F2 a permis d'identifier une région de 10,8 Mb située sur le LG 12 et associée à la résistance aux pyréthriinoïdes, ce qui est traduit par un fort LOD score (correspondant à la vraisemblance de l'existence d'une liaison génétique) (Figure 5.4.A).

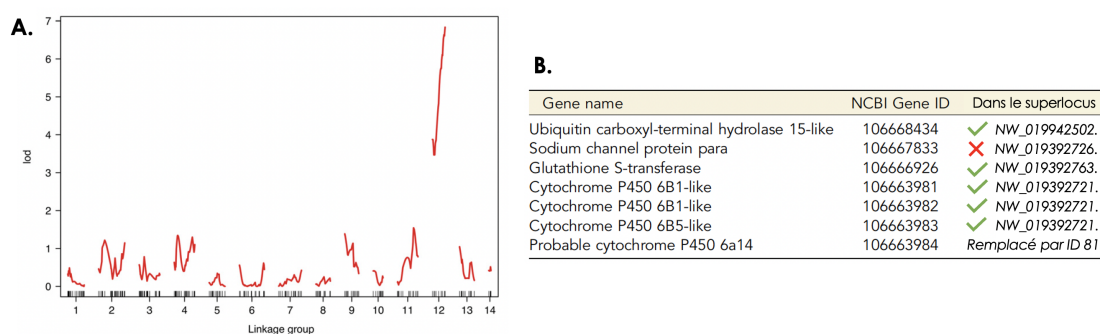


Figure 5.4 – Résultats extraits des travaux de [Fountain et al. \(2016\)](#). A. Les scores LOD des marqueurs RAD le long du génome mettent en évidence un pic significatif sur le LG 12. B. Nom et ID sur NCBI des gènes de résistance putatifs identifiés dans l'intervalle de confiance Bayésien à 95% calculé sur le QTL de résistance du LG 12. Leur présence au sein du superlocus ainsi que le scaffold potentiel sur lequel ils sont localisés dans la version 2.1 du génome de référence a été ajouté.

Un intervalle de confiance Bayésien à 95% a été calculé sur ce locus de caractères quantitatifs (QTL) associé au caractère de résistance, et sept gènes potentiellement impliqués dans la résistance y ont été identifiés (Figure 5.4.B).

Parmi ces sept gènes, cinq font partie du superlocus identifié par notre analyse génomique, un n'est plus considéré comme un gène à part dans la nouvelle version de l'assemblage, et un ne fait pas partie de ce superlocus : le canal sodium voltage-dépendant (*VGSC*, pour *voltage-gated sodium channel*). Pourtant, il s'agit de l'un des rares gènes dont l'implication dans le phénotype de résistance aux pyréthrinoides a été étudiée à de multiples reprises, relativement à la fréquence de la mutation *kdr* qu'il porte (Dang *et al.*, 2014, 2015; Akhouni *et al.*, 2015; Balvín and Booth, 2018). Néanmoins, son absence du superlocus est potentiellement artefactuelle : la différence entre la version du génome utilisée par Fountain *et al.* (2016) (Clec_1.0) et notre étude génomique (Clec_2.1) réside en partie dans le nombre de scaffolds, un peu plus élevé dans la version plus récente. Le scaffold portant le gène *VGSC* en particulier mesurait 21 Mb dans l'ancien assemblage (NW_014465016.1 sur Clec_1.0), et contenait 6 marqueurs RAD (voir le tableau supplémentaire S2 de Fountain *et al.* 2016), avant d'être divisé en plus petits morceaux dans l'assemblage Clec_2.1, laissant *VGSC* dans un scaffold de 1 Mb ne portant plus aucun marqueur RAD (NW_019392726.1 sur Clec_2.1). Cette différence entre les assemblages a donc mécaniquement empêché son intégration dans la nouvelle carte génétique.

Nous notons cependant que dans le scaffold de 21 Mb identifié par Fountain, *VGSC* était situé entre deux gènes de résistance qui sont tous deux inclus dans le superlocus (une carboxyl-terminal hydrolase sur le scaffold NW_019942502.1, et une glutathion S-transférase sur le scaffold NW_019392763.1). Le scaffold contenant le gène *VGSC* (NW_019392726.1) pourrait donc être localisé entre les deux scaffolds précédemment mentionnés, au milieu du superlocus. Si c'est le cas, cela renforcerait d'avantage le signal de différenciation génétique dans cette zone génomique. Ainsi, des efforts futurs devront être réalisés pour construire un assemblage de meilleure qualité, afin de clarifier la structure de cette région (voir partie 5.2.3).

La forte différenciation génétique mesurée dans cette région génomique, ainsi que le fractionnement des scaffolds entre les différentes versions du génome de référence laissent entrevoir qu'il s'agit d'une région potentiellement difficile à assembler.

5.2.2 Superlocus en auto-stop ou évolution d'un supergène ?

Au cours de notre analyse génomique, nous avons ainsi identifié une région fortement différenciée de 6 Mb. Plusieurs hypothèses peuvent expliquer la différenciation de ce superlocus. La première hypothèse impliquerait la présence de plusieurs loci répartis tout au long du superlocus contribuant au phénotype de résistance. La pression de sélection entraînerait donc mécaniquement une différenciation sur l'ensemble du locus de 6 Mb. En réalité, on sait que les fréquences alléliques des loci physiquement liés n'évoluent pas indépendamment les uns des autres. On peut donc imaginer un second modèle, caractérisé par la présence d'un seul ou un faible nombre de loci avantageux, dont la montée en fréquence serait récente, ayant entraîné une majorité de mutations neutres par effet d'auto-stop. Ce processus d'auto-stop consiste en une augmentation de la fréquence d'un allèle *a* du fait de sa proximité avec un allèle *b* sous sélection positive (Smith and Haigh, 1974).

Sous cette hypothèse, la mutation *kdr* et/ou le cluster de P450 pourraient être les seules mutations réellement impliquées dans la résistance aux insecticides au sein de ce superlocus, et le reste des variants détectés seraient simplement montés en fréquence par un effet d'auto-stop. Les données actuelles ne nous permettant pas de conclure quant aux impacts respectifs des gènes localisés sur le superlocus, il sera nécessaire d'entreprendre des étapes de validation fonctionnelle. Néanmoins, la littérature soutient plutôt l'hypothèse d'une réponse polygénique des punaises de lit aux insecticides, en identifiant de multiples mécanismes co-existants dans les souches résistantes (Dang *et al.*, 2017). Le superlocus détecté porte de nombreux gènes candidats à la résistance aux insecticides, dont le *VGSC*, un cluster de *CYP6B*, mais également également des GST (voir section 5.2.1). L'implication des P450 dans le phénotype de résistance aux pyréthrinoïdes a entre autres été mise en évidence en rétablissant partiellement le phénotype de sensibilité suite à l'utilisation de leur inhibiteur enzymatique, le PBO (Romero *et al.*, 2009b; Gonzalez-Morales and Romero, 2019). De plus, il est très rare de ne trouver que *kdr* dans une population résistante d'insectes, notamment à cause du potentiel coût engendré par cette mutation (Wang *et al.*, 2021; Freeman *et al.*, 2021), y compris chez les punaises de lit (Gordon *et al.*, 2015). Ces données seraient donc en défaveur de la présence d'un seul mécanisme d'intérêt au niveau du superlocus, et donc de cette hypothèse de "superlocus généré par auto-stop". De plus, on s'attend à une érosion de la différenciation lorsque l'on s'éloigne d'un locus sous sélection (Chen *et al.*, 2010). Or, nos données indiquent une large

région différenciée de manière relativement homogène.

Un troisième modèle pourrait impliquer l'évolution d'un regroupement d'allèles physiquement liés, héritable en un seul bloc haplotypique, appelé supergène. Cet ensemble d'allèles adaptatifs dans un environnement donné, sont typiquement maintenus ensemble grâce à de larges inversions, empêchant la recombinaison avec les autres allèles. De nombreuses études ont détecté des supergènes, pouvant être impliqués dans une multitude de phénotypes adaptatifs (Figure 5.5, voir revue de Schwander *et al.* 2014).

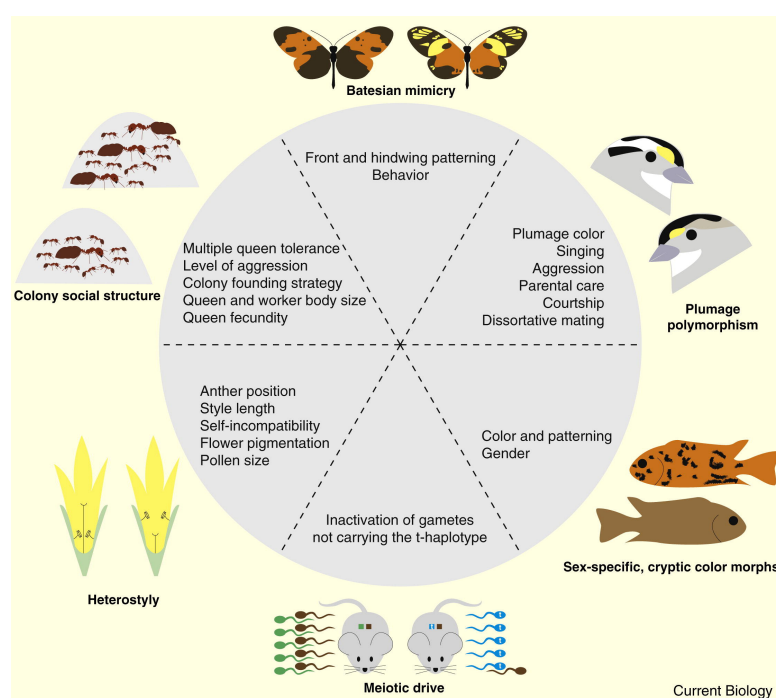


Figure 5.5 – Variété de phénotypes alternatifs pouvant être régulés par les supergènes d'après Schwander *et al.* (2014).

Les travaux de Dobzhansky and Sturtevant (1938) sur *Drosophila pseudoobscura* ont été fondateurs pour l'identification d'inversions chromosomiques impliquées dans une réduction de la recombinaison. En effet, différentes inversions coexistent à diverses positions chez cette espèce, et présentent très peu de recombinaison entre allèle sauvage et allèle inversé (Andersona *et al.*, 1991). Ainsi, ces haplotypes se sont différenciés génétiquement les uns des autres.

Il a été montré que la fréquence de ces inversions étaient corrélée à plusieurs variables environnementales (Andersona *et al.*, 1991), suggérant que ces régions constituent des "complexes de gènes co-adaptés". C'est le cas des souris sylvestres *Peromyscus maniculatus*, qui possèdent une inversion chromosomique contenant

différents allèles des gènes codant pour la taille de la queue et la couleur de la fourrure (Hager *et al.*, 2022). Cette inversion est présente à 90% chez l'écotype de forêt, possédant une longue queue et un pelage noir, et absente chez l'écotype de prairie, à la queue courte et au pelage clair. La forte corrélation observée entre phénotype et habitat traduirait donc une adaptation locale de ces différentes populations de souris : la longue queue des écotypes de forêt pourrait leur permettre de grimper et donc de se mouvoir plus facilement dans cet habitat (Hager and Hoekstra, 2021), tandis que leur pelage sombre les rend cryptiques vis-à-vis des prédateurs (Vignieri *et al.*, 2010). D'après Hager *et al.* (2022), une pression de sélection divergente maintiendrait l'inversion dans les populations de souris de forêt, puisque l'inversion empêche la recombinaison avec l'haplotype de prairie non inversé.

D'autres phénomènes adaptatifs, dont la résistance aux insecticides, ont également été associés à la présence de larges inversions présentant les caractéristiques d'un supergène. Chez le moustique *Anopheles gambiae*, où un cluster particulier de gènes candidats à la résistance aux insecticides a été identifié sur le chromosome 2L (Ranson *et al.*, 2002), une inversion de plus de 20 Mb a également été détectée sur ce même chromosome (Weetman *et al.*, 2018a).

Un polymorphisme est observé au niveau de cette inversion, avec les allèles 2La (inversé) et 2L+a (Riehle *et al.*, 2017). L'allèle 2L+a est associé à des conditions climatiques plus humides, ainsi qu'à la résistance aux pyréthriinoïdes (Weetman *et al.*, 2018a). Il semblerait en effet que la résistance soit plus forte (mortalité plus faible) lorsque l'humidité augmente (voir Figure S3 chez Weetman *et al.* 2018a). Bien que cet allèle soit sur le même chromosome que le gène *VGSC*, les variations de leurs fréquences sont décorréelées, et la résistance semble donc due à des mutations au niveau d'enzymes de détoxification (Ranson *et al.*, 2010; Matoke-Muhia *et al.*, 2016). Certaines de ces études ont également tenté de relier les fréquences de l'inversion et du nombre de moustiquaires traitées aux insecticides utilisées, mais ont obtenu des résultats contradictoires, potentiellement à cause des différentes méthodes d'échantillonnages favorisant respectivement la capture d'individus endophiles ou exophiles (Brooke *et al.*, 2002; Matoke-Muhia *et al.*, 2016). De manière similaire, une étude génomique chez *Anopheles coluzzi* a mis en évidence plusieurs régions fortement associées à la résistance, incluant une inversion sur le chromosome 2R (Ingham *et al.*, 2021b).

Le superlocus montre un enrichissement significatif en gènes potentiellement impliqués dans la résistance aux insecticides, avec la détection de cinq des sept

gènes identifiés par Fountain (Figure 5.4.B), ainsi que de deux autres GST et de gènes potentiellement impliqués dans le système nerveux et d'autres mécanismes de détoxification, pour un total de neuf gènes.

La colocalisation de gènes impliqués dans la détoxification dans des clusters est fréquemment observée chez les insectes. Chez le moustique *Anopheles gambiae*, 16 clusters de quatre gènes ou plus parmi les GST, carboxylestérases et P450s ont été détectés, la plupart étant composés de paralogues (Ranson *et al.*, 2002). Cette étude a même mis en évidence un cluster particulier contenant 23 COE, 1 GST, and 11 P450s sur le chromosome 2L. De manière similaire, une région exceptionnellement grande et peu polymorphe de 1,3 Mb a été décrite dans une population du lépidoptère *Amyelois transitella* résistante au DDT et à la bifenthrine (Calla *et al.*, 2021). Cette région contenait plusieurs gènes probablement impliqués dans la résistance aux insecticides, comme le *VGSC* et un cluster de *CYP6B*.

Néanmoins, il peut être étonnant de détecter ce regroupement de gènes impliqués dans la résistance aux pyréthrinoïdes chez la punaise de lit dans un génome de référence échantillonné antérieurement à l'utilisation de pyréthrinoïdes.

La présence de clusters de gènes impliqués dans la résistance aux insecticides à l'intérieur de génomes de référence "sensibles" peut-être due à des millions d'années d'exposition à une variété de xénobiotiques présents *in natura*. En effet, la plupart des insecticides chimiques sont similaires aux composés produits naturellement par les plantes, comme les pyréthrinoïdes, extraits de fleurs appelées pyrèthres de Dalmatie, ou les nicotinoïdes qui sont similaires à la nicotine produite par le tabac. L'herbivorie des insectes agit comme une pression sélective, et entraîne la sélection de la production de toxines par les plantes (Després *et al.*, 2007). En réponse, les insectes se nourrissant de plantes ont été sélectionnés pour résister à ces toxines végétales, souvent *via* une détoxification accrue, permettant une adaptation à l'environnement (Berenbaum, 2002). Cette co-évolution intense a conduit à une grande diversification des systèmes de détoxification des insectes, leur permettant de s'adapter rapidement à la plupart des insecticides synthétiques Wen *et al.* (2005). Des regroupements de ces gènes impliqués dans la résistance peuvent être dus à des translocations, ou à des duplications locales de gènes, ensuite soumis à une pression de sélection positive (Richly *et al.*, 2002; Ranson *et al.*, 2002).

De plus, la souche de référence de punaises de lit Harlan, sensible aux pyréthrinoïdes, avait potentiellement déjà été exposée au DDT lors de sa collecte

en 1973, dont la première utilisation remonte aux années 40. Or, le DDT, tout comme les pyréthriinoïdes, cible le *VGSC* : des mécanismes conférant une résistance croisées peuvent donc être présents dans cette souche à l'état ancestral, bien qu'ayant pu également être perdus en cas de coût trop important. Au cours de cette thèse, nous étudions donc des phénomènes relativement récents de mutations, conférant une résistance aux insecticides modernes.

Chez la punaise de lit, nos travaux suggèrent donc un enrichissement à la fois en gènes de résistance putatifs et en inversions, dans une large région génomique très différenciée et associée à la résistance aux insecticides. Il est donc tentant de spéculer qu'un supergène combinant plusieurs allèles avantageux lors de l'exposition aux insecticides ségrège dans les populations de *C. lectularius*. Néanmoins, nos données devront être complétées par d'autres méthodes permettant d'une part d'éclaircir l'organisation du génome dans cette région, et d'autre part de rechercher la présence d'une ou plusieurs inversions sur ce locus.

5.2.3 Organisation du génome

L'objectif de l'amélioration de l'assemblage du génome de la punaise de lit est double : (i) obtenir une résolution génomique à l'échelle chromosomique, permettant entre autre de pouvoir placer le scaffold portant la mutation *kdr* dans un des chromosomes, (ii) préciser l'architecture du superlocus identifié, en ouvrant la possibilité de détecter des variants structuraux (et notamment des inversions) intra-chromosomiques. Pour ce faire, deux méthodes semblent particulièrement pertinentes : le séquençage à lectures longues (long-reads) et/ou le séquençage Hi-C.

Le séquençage long-reads donne en effet la possibilité de détecter les variants structuraux (Shao *et al.*, 2018), et de résoudre les régions répétées (dans une certaine mesure). Les données Hi-C (High Chromosome Contact) quant à elles permettent de générer des cartes de contacts chromosomiques. Les fragments d'ADN spatialement proches dans le noyau sont liés ensemble, et une matrice contenant le nombre de co-occurrences entre les reads est ensuite produite. Plus deux régions génomiques sont proches, et plus cette carte recensera de contacts entre elles, permettant ainsi de générer un assemblage du génome à l'échelle chromosomique.

Une étude récente cumulant des données long-reads PacBio et Hi-C a ainsi

réassemblé les scaffolds du génome du hareng *Clupea harengus* en 26 chromosomes (Pettersson *et al.*, 2019). En parallèle, deux sous-familles de harengs (de la mer Baltique et de l'Atlantique) ont été génotypés pour 45 000 marqueurs et ont formé 26 groupes de liaison, confirmant l'assemblage chromosomiques. Le déséquilibre de liaison entre ces marqueurs a été mesuré, et révélé une région à faible recombinaison sur le chromosome 12, définie comme un potentiel supergène. Ce supergène, correspondant à une inversion de 7.8 Mb sur le chromosome 12, contient des allèles impliqués dans l'adaptation écologique du hareng à différents environnements, probablement en lien avec la température.

Si le superlocus détecté chez la punaise de lit est effectivement un supergène, on s'attend à une réduction de la recombinaison entre l'haplotype des populations sensibles (Ref) et celui portant les allèles impliqués dans la résistante (Inv). Néanmoins, ce supergène pourrait également présenter un certain polymorphisme. On peut en effet imaginer que plusieurs allèles co-existent, certains portant *kdr* (représentant potentiellement un coût en absence d'insecticide) et d'autres non. Il pourrait donc être intéressant de génotyper individuellement de nombreuses punaises de lit issues de populations différentes et présentant des phénotypes de résistance et de sensibilité distincts, afin d'identifier les différents haplotypes.

Des croisements contrôlés au sein d'une lignée résistante portant le potentiel superlocus inversé et dans laquelle la mutation *kdr* n'est pas fixée, comme London Field (fréquence de 0,44 contre 0 chez LL), peuvent permettre de sélectionner une lignée isofemelle *kdr-free*, ainsi qu'une lignée *kdr+* (ces lignées existent par exemple déjà chez *Aedes aegypti* bien que n'ayant pas encore été publiées, Bacot *et al.* 2023). Pour cela, des extractions peuvent être réalisées sur des exuvies des stades 5 avant de former les couples, afin de tester la présence ou l'absence de *kdr* par PCR, suivant une méthodologie similaire à celle décrite chez le moustique *Aedes aegypti* (Saavedra-Rodriguez *et al.*, 2007).

Un séquençage pool-seq des lignées *kdr-free* et *kdr+* pourrait ensuite permettre de mesurer la différenciation de la région du superlocus par rapport au génome de référence sensible (Figure 5.6). Cette différenciation pourrait ainsi mettre en évidence la présence de la mutation *kdr* sur le supergène (Inv-1 *kdr+*), son absence du supergène (Inv-2 *kdr+*), ou encore la présence de deux haplotypes du supergène co-existants avec ou sans cette mutation (Inv-1 *kdr-free*, Inv-1 *kdr+*). Dans un second temps, le phénotypage des lignées *kdr-free* et *kdr+* peut être réalisé, afin de mesurer un éventuel décalage entre les niveaux de résistance de ces lignées. Une absence de différence entre les phénotypes de résistance de ces souches impliquerait

que d'autres allèles que *kdr* induisent la résistance observée. Si la souche *kdr-free* avait un phénotype de sensibilité restauré par rapport à la souche *kdr+*, soit comparable à celui de la souche de référence Ref, la mutation *kdr* jouerait alors un rôle crucial dans la résistance. Enfin, si la lignée *kdr-free* présentait un niveau de résistance intermédiaire entre *kdr+* et la souche de référence, la co-existence de plusieurs haplotypes du supergène semble probable, certains portant l'allèle sensible du *VGSC*, et d'autres l'allèle résistant *kdr*. Ce niveau de résistance intermédiaire impliquerait une réponse polygénique du supergène, soit la contribution d'autres loci que *kdr* dans le phénotype de résistance.

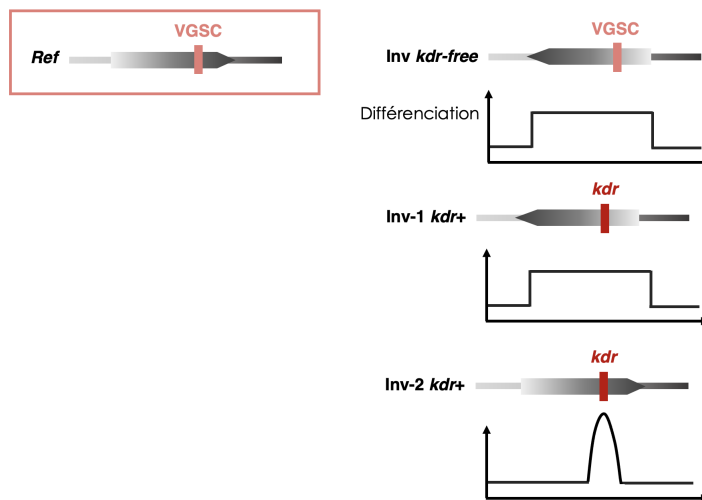


Figure 5.6 – Haplotypes putatifs (non-exhaustifs) du superlocus inversé (*Inv*, i.e. le supergène) pouvant (co-)exister par rapport au génome de référence (*Ref*), en fonction de la présence de l'allèle muté *kdr* du *VGSC*. L'allèle de référence du gène *VGSC* est coloré plus pâlement que sa version mutée conférant un avantage adaptatif, *kdr*. La mutation *kdr* peut être associée à une forte différenciation du superlocus, traduisant la présence du supergène (*Inv-1 kdr+*), ou ne pas être associée à un signal large de différenciation sur cette région par rapport à l'allèle de référence, traduisant l'absence de *kdr* du supergène (*Inv-2 kdr+*). De même pour la lignée *kdr-free* : l'absence de *kdr* peut être associée à la présence du superlocus inversé (*Inv kdr-free*) ou à son absence (pas de différenciation avec le génome de référence, soit restauration de l'allèle *Ref*).

Ce type d'expérience permet donc d'identifier la présence ou l'absence du locus pour lequel la sélection est faite (ici, *kdr*) à l'intérieur d'un supergène putatif. Elle pourrait donc également être réalisée avec d'autres loci, afin d'identifier quels allèles mutés sont associés à ce large signal de différenciation.

L'allèle *kdr* étant récessif, la détection de populations de punaises de lit ne le portant pas mais étant fortement résistantes aux insecticides (décrites plus bas en section 5.3.1) implique l'existence d'autres mécanismes de résistance, potentiellement également sous forme de supergène mais avec l'allèle de référence

sensible du *VGSC*. Ces populations de punaises résistantes et "*kdr-free*" détectées sur le terrain ayant été échantillonnées aux USA, on peut aussi imaginer que le mécanisme de supergène observé est propre aux populations échantillonnées à Londres. L'influence de la structure des populations sur la diversité des mécanismes de résistance devra donc être explorée afin d'avoir des marqueurs fiables pour pouvoir réellement suivre la propagation de la résistance sur le terrain (voir section 5.4.1).

5.3 Mécanismes de résistance aux pyréthrinoïdes chez *C. lectularius*

Un des objectifs principaux de cette thèse était de mettre en évidence les bases génétiques sous-jacentes aux mécanismes impliqués dans le phénotype de résistance aux pyréthrinoïdes chez la punaise de lit. Nous avons choisi de diviser ces mécanismes en fonction de leurs réponses transcriptionnelles et génomiques, telles que nous avons pu les observer. Ainsi, nos résultats seront discutés dans cette section par le prisme de trois grands types de variants et de leurs impacts : les SNPs non-synonymes, les SNPs ou SVs associés à une réponse transcriptionnelle, et les SNPs associés à une réponse post-transcriptionnelle (Figure 5.7).

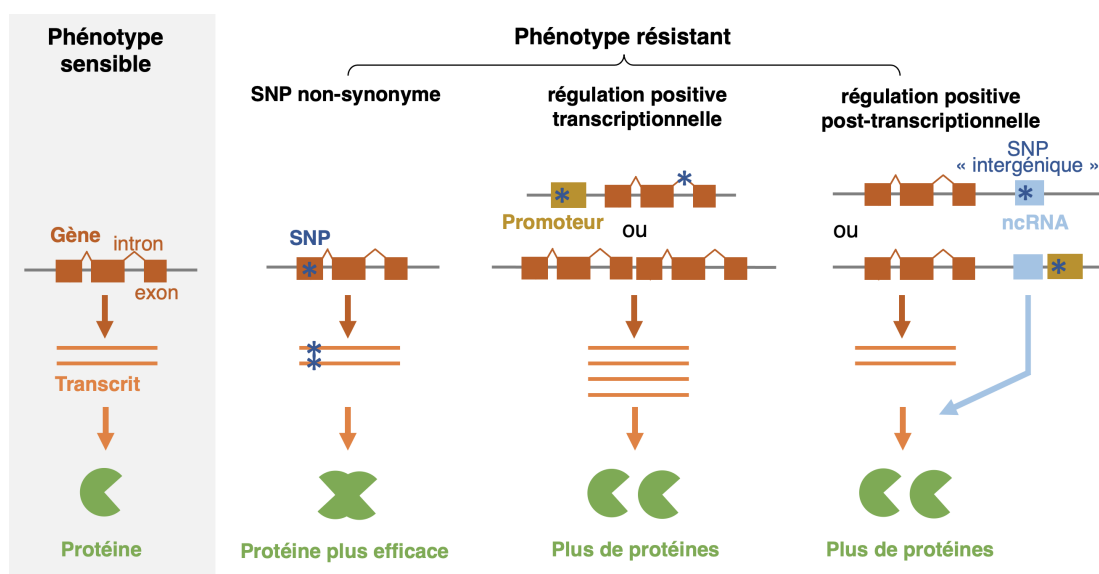


Figure 5.7 – Impact transcriptionnel des différents mécanismes potentiellement impliqués dans le phénotype de résistance aux insecticides. Chez le phénotype sensible, les transcrits (ARNm mature) sont produits suite à la transcription du gène en pré-ARNm et à l'épissage de ce dernier, éliminant les introns. Ces transcrits seront ensuite traduits en protéines, capables d'assurer une ou plusieurs fonctions. Chez le phénotype résistant, plusieurs mécanismes peuvent (co-)exister : (i) un SNP non-synonyme de la séquence codante va changer l'acide aminé produit, ce qui peut modifier la protéine en augmentant son affinité pour sa cible, la rendant plus efficace ; (ii) la présence de SNP dans la région promotrice ou intronique d'un gène, ou la duplication du gène, peut influencer de manière positive la régulation de ce gène et donc augmenter le nombre de transcrits produits, synthétisant ainsi plus de protéines ; (iii) la présence de SNP (détectés dans la région "intergénique") dans des non-coding RNA (ncRNA), ou les régions promotrices de ces derniers, peut influencer positivement la traduction de protéines à partir des transcrits du gènes, augmentant ainsi leur nombre. Si l'on observe un même gène dans tous ces cas de figure, les promoteurs ainsi que les ncRNA sont bien évidemment présents aux mêmes emplacements, mais n'ont pas été représentés lorsqu'ils n'étaient pas mutés, dans un souci de lisibilité.

5.3.1 Rares SNPs non-synonymes

Nos analyses génomiques conduites sur quatre populations aux statuts de résistance contrastés n'ont permis de mettre en évidence qu'un nombre très restreint de SNPs non-synonymes. Ainsi, parmi les 576 SNPs outliers les plus différenciés en fonction du statut de résistance des populations identifié lors de l'analyse génomique des souches, un seul était non-synonyme, et détecté dans le facteur de transcription homeobox protein one-cut. La mutation *kdr*, le variant génétique le plus recherché chez les insectes résistants aux pyréthriinoïdes, ne faisait même pas partie de ces SNPs outliers.

Importance de la mutation *kdr*

La mutation *kdr* est un variant changeant la conformation du site ciblé par les pyréthriinoïdes, empêchant ainsi leur fixation. Ce type de mutation est donc considéré comme responsable de la résistance aux pyréthriinoïdes chez divers arthropodes nuisibles (Dong *et al.*, 2014). Pourtant, comme mentionné plus haut, cette mutation n'a pas été détectée comme faisant partie des SNPs outliers détectés dans le chapitre 2 (i.e. fortement associé au phénotype de résistance par le logiciel BayPass, prenant en compte l'histoire évolutive des populations). Nous pouvons nous interroger sur la signification de ce résultat.

Tout d'abord, nous pouvons émettre l'hypothèse que cette mutation n'est pas aussi cruciale pour la résistance aux insecticides chez *C. lectularius* qu'on ne le pensait auparavant. Dans ce scénario, d'autres mécanismes tels que la détoxification, l'épaississement de la cuticule ou l'évitement comportemental seraient plus centraux dans la résistance aux insecticides des punaises. En effet, une étude menée sur de multiples populations aux États-Unis a montré que certaines de ces populations (CIN-1, CIN-3) étaient résistantes aux pyréthriinoïdes, sans pour autant porter le variant *kdr* (Zhu *et al.*, 2010). Pour autant, le ratio de résistance de ces populations par rapport à une population sensible échantillonnée à Fort Dix (la souche Harlan) n'a pas pu être déterminé (>12,765 X), la dose létale pour les souches résistantes étant en effet supérieure à la limite de solubilité de la deltaméthrine dans l'acétone (Romero *et al.*, 2007). On peut donc observer sur le terrain des populations très fortement résistantes, sans implication de la mutation *kdr*.

La résistance aux pyréthriinoïdes conférée par la mutation *kdr* a été identifiée chez de nombreuses espèces d'arthropodes, telles que *Musca domestica*

(Williamson *et al.*, 1993), la blatte germanique *Blattella germanica* (Dong and Scott, 1994), la mouche *Haematobia irritans* (Jamroz *et al.*, 1998), et de nombreuses espèces de moustiques - *A. aegypti*, *C. quinquefasciatus* ou encore *An. gambiae* (Saavedra-Rodriguez *et al.*, 2007; Silva Martins *et al.*, 2019; Lynd *et al.*, 2010). Néanmoins, cette mutation ne s'est pas (encore?) répandue dans toutes les populations de terrain pour ces différents insectes - quelques irréductibles villages résistent! C'est le cas de certaines population d'*An. gambiae* récemment échantillonnées sur l'archipel du Zanzibar, qui ont montré des niveaux de résistance élevés pour les pyréthriinoïdes, sans association avec la mutation *kdr* présente chez seulement 0.97% des individus échantillonnés (Kampango *et al.*, 2022). On peut imaginer que d'autres mécanismes plus efficace et/ou moins coûteux en l'absence d'insecticide aient évolué au sein de ces populations, ralentissant la montée en fréquence de *kdr*, qui ne présente pas à titre comparatif un avantage suffisant.

Variations des séquences codantes de gènes induits par l'exposition aux pyréthriinoïdes

Les gènes qui ne montrent pas de différence transcriptomique entre London Lab et London Field ne devraient pas être impliqués dans l'écart de phénotype de résistance observé entre ces souches, à moins d'une différence dans leur séquence codante. En effet, certaines mutations peuvent faciliter la dégradation de l'insecticide, en rendant la protéine plus efficace. Nous avons recherché de telles différences spécifiquement pour les gènes "induits" par l'exposition à la deltaméthrine, ces gènes étant plus probablement impliqués dans des mécanismes permettant à l'insecte de lutter contre l'insecticide.

Parmi les 375 gènes détectés par notre approche transcriptomique comme étant induits suite à l'exposition aux pyréthriinoïdes de manière similaire chez les souches résistantes et sensibles, quelques SNPs non-synonymes fortement différenciés entre LL et LF ont ainsi été identifiés (filtres moins stringents que pour identifier les "outliers", détaillés dans le chapitre 3). Bien qu'aucun de ces SNPs ne concernent des gènes de résistance connus, un en particulier a attiré notre attention.

Ce gène codant pour une protéine serpine B3 était en effet associé à trois SNP non-synonymes fortement différenciés, avec la fixation de l'allèle dérivé putatif chez la souche résistante (LF). De plus, ce gène était associé au total à 28 mutations non-synonymes dans notre analyse génomique précédente (Haberkorn *et al.*, 2023), faisant ainsi partie du premier percentile des gènes ayant le rapport

entre les mutations non synonymes et synonymes le plus élevé ($\pi_n/\pi_s=4,67$). De plus, cette serpine B3 était globalement très différenciée sur l'ensemble de sa séquence, puisque l'indice de différenciation calculé à l'échelle du gène entier faisait partie des 5% les plus élevés dans la comparaison entre LL et LF ($F_{ST}=0,15$). Toutes ces données suggèrent que ce gène a subi une sélection positive dans un passé récent, conduisant à une augmentation de la fréquence de l'allèle dérivé putatif. Bien qu'à notre connaissance ce gène n'ait jamais été associé à la résistance aux insecticides, le fait qu'il soit également induit lors de l'exposition à la deltaméthrine rend plausible son implication dans ce phénomène. La fonction de ce gène précis (serpine B3) est inconnue chez *C. lectularius*, mais on sait que les serpines sont impliquées dans la réponse immunitaire des insectes, et peuvent donc avoir un effet dans la résistance aux insecticides, comme chez les moustiques *Culex pipiens* (Li *et al.*, 2016). Les serpines jouent en effet un rôle régulateur dans le système immunitaire en inhibant les protéases à sérines, permettant par exemple la mélanisation (encapsulation) d'un pathogène chez le moustique ou encore la drosophile (Tang *et al.*, 2008). L'encapsulation tue le pathogène ciblé en l'enveloppant d'un complexe protéique épais et sombre, l'empêchant alors d'absorber les nutriments qui l'entourent (Kumar *et al.*, 2018). Ce processus est connu pour produire de nombreux dérivés réactifs de l'oxygène (ROS), engendrant un stress oxydatif chez l'insecte. Or, deux mécanismes de résistance aux insecticides peuvent modifier le taux de ROS dans l'organisme de l'insecte : les P450s peuvent l'augmenter, tandis que les GSTs le diminuer. On peut donc envisager que l'insecticide irait perturber la régulation des serpines à la hausse, engendrant un stress oxydatif, qui irait enclencher à son tour la sur-expression d'enzymes de détoxification pour permettre à l'insecticide de réduire ce stress, dégradant également les insecticides.

Les mutations non-synonymes ont souvent été explorées chez les gènes "cibles" des différents insecticides, comme l'acetylcholinestérase *ace1* pour les carbamates et les organophosphorés, les *VGSC* pour les pyréthrinoïdes et le DDT, ou encore les récepteurs GABA *rdl* pour les organochlorés (Weetman *et al.*, 2015; Silva Martins *et al.*, 2019; Le Goff *et al.*, 2003). Néanmoins, quelques mutations non-synonymes ont également été associées à la résistance chez des gènes de détoxification, comme GSTe2, où le changement d'acide aminé L119F confère une résistance croisée au DDT et aux pyréthrinoïdes chez *An. funestus* (Riveron *et al.*, 2014). La même étude a également montré que ce gène était régulé de manière positive chez les moustiques suite à une exposition au DDT. Des analyses structurales ont ensuite identifié que le SNP non-synonyme détecté élargissait le site de liaison au DDT chez l'allèle résistant, augmentant ainsi ses capacité de métabolisation (Riveron

et al., 2014). Néanmoins, aucune mutation de ce type n'a été détectée à l'intérieur de gènes candidats à la résistance dans nos données, laissant donc penser que la réponse est plutôt transcriptionnelle.

5.3.2 Régulation transcriptionnelle

La présence de mutations ponctuelles en dehors de la séquence codante, mais également de plus grands variants structuraux, peuvent affecter la régulation transcriptionnelle des gènes.

Les variants structuraux peuvent contribuer à l'adaptation notamment par le biais d'inversions ou de duplications. Les variations de nombre de copies de gènes (CNV) peuvent avoir un impact sur leur niveau d'expression *via* un effet gène-dosage (Kondrashov, 2012), tandis que les inversions peuvent rapprocher le gène d'éléments régulateurs influant leur expression (Hough *et al.*, 1998; Puig *et al.*, 2004).

Concernant les variants ponctuels, des mutations dans les introns ou le promoteur d'un gène peuvent affecter sa régulation. La résistance à l'insecticide spinosad chez *Plutella xylostella* a par exemple été associée à une mutation dans la région intronique du gène, entraînant des protéines tronquées (Bastarache *et al.*, 2020). Chez *Nilaparvata lugens*, la sur-expression du P450 *CYP6ER1* confère une résistance à l'imidaclopride (Liang *et al.*, 2018). Les allèles de résistances de 4 SNPs augmentent en effet l'activité du promoteur du transcrit alternatif A2 où ils sont détectés. Nous avons donc tenté de corrélérer la présence de mutations de ce type dans nos données génomiques avec celle de gènes présentant une expression différentielle dans nos données transcriptomiques. Néanmoins, il nous faut souligner qu'il n'existe encore que très peu d'études démontrant clairement l'implication de telles mutations dans la régulation transcriptionnelle de gènes de résistance (Ingham *et al.*, 2021a).

Différences constitutives d'expression entre les souches résistantes et sensibles

Les données génomiques ont révélé un enrichissement en gènes potentiellement impliqués dans la résistance cuticulaire parmi les variants structuraux les plus différenciés (top 10% des valeurs de F_{ST}). De manière concomitante, l'analyse des données transcriptomiques a permis de mettre en évidence un enrichissement en

gènes de cette même catégorie de résistance chez ceux sur-exprimés chez les individus résistants aux pyréthrinoïdes par rapport aux individus sensibles.

En effet, parmi les 20 gènes détectés comme significativement sur-exprimés chez la souche résistante London Field, six étaient des protéines cuticulaires. La résistance cuticulaire a été observée chez de nombreux insectes résistants aux pyréthrinoïdes, comme *Helicoverpa armigera*. (Ahmad *et al.*, 2006), *Anopheles gambiae* (Yahouédo *et al.*, 2017) ou *Culex pipiens* (Pan *et al.*, 2009). Chez les punaises de lit également, plusieurs résultats suggèrent que la cuticule joue un rôle central dans la résistance aux insecticides. Par exemple, une étude morphologique a montré une corrélation positive entre l'épaisseur de la cuticule et le statut de résistance des souches (Lilly *et al.*, 2016a). De plus, trois transcrits codant pour des protéines cuticulaires (C2, C10 et C13) ont été détectés comme surexprimés dans différentes populations résistantes de punaises de lit (Zhu *et al.*, 2013). D'autres études ont montré que la chitine synthase ou la protéine cuticulaire larvaire/pupale avaient un niveau de transcription plus élevé dans les souches de punaises de lit résistantes aux pyréthrinoïdes (Mamidala *et al.*, 2012; Koganemaru *et al.*, 2013).

Chez cinq des six protéines cuticulaires détectées, une duplication inversée ou des SNPs présents dans une région non-codante (située dans l'intron, le promoteur ou l'intergénique) peuvent potentiellement expliquer ce différentiel d'expression. Néanmoins, bien que la duplication inversée soit dans le top 10% des variants structuraux ayant l'indice de différenciation le plus élevé, les fréquences observées entre les populations n'était pas si discriminantes (12,3% chez LF contre 5,7% chez LL). Concernant les SNPs d'intérêt détectés, certains se trouvent dans des régions promotrices ou les introns des gènes, et présentent des F_{ST} extrêmes indiquant une différenciation génétique exceptionnelle ($pF_{ST} < 1E-03$). Les mutations de ce type, proches du gène (intergéniques) et introniques, peuvent en effet entraîner des modifications transcriptomiques en jouant un rôle de séquences régulatrices, recrutant l'ARN polymérase II au niveau du promoteur (Cannavò *et al.*, 2016). Des expériences de validation pourront être envisagées sur ces candidats cuticulaires, afin de mesurer l'implication des différents SNPs détectés dans le phénotype de résistance, comme cela a été réalisé par retard sur gel (Electrophoretic mobility shift assay) chez la cicadelle brune *Nilaparvata lugens* (Liang *et al.*, 2018).

Chez *C. lectularius*, et plus particulièrement chez les souches Londoniennes, les punaises de lit au phénotype résistant ont plusieurs protéines cuticulaires sur-exprimées par rapport à celles du phénotype sensible. Aucune interaction n'a

été mesurée avec le différentiel d'expression lié au traitement aux insecticides pour ces gènes, ce qui signifie que la sur-expression est exclusivement constitutive. Cette régulation de la production en quantité de transcrits mesurée est donc vraisemblablement liée à des différences génétiques entre les deux souches, qui pourraient ici prendre la forme de SNPs dans les introns ou encore dans les régions promotrices.

Quid du cluster de P450 détecté sur le superlocus ?

Certains allèles ont également été détectés dans nos données génomiques comme semblant particulièrement intéressants, mais sans présenter de signal transcriptionnel par la suite. C'est le cas d'un cluster de trois P450 (deux *CYP6B1* et un *CYP6B5*), détecté dans le scaffold NW_019392721.1, au milieu du superlocus LG11 chez qui nous avons identifié quatre SNPs présentant des indices de différenciation particulièrement élevés entre les souches de Londres (top 5% des F_{ST}). En outre, ce cluster était caractérisé par un nombre étonnamment élevé de mutations non-synonymes (n=19). Le ratio π_n/π_s pour les trois gènes se situait en effet dans les 5% des ratios les plus élevés à l'échelle du génome, ce qui suggère que la sélection a pu agir sur ces loci récemment. Un signal similaire a été observé pour le *CYP6B56* du lépidoptère *Amyelois transitella*, probablement impliqué dans la résistance au pyréthrianoïde bifenthrine (Calla *et al.*, 2021).

Ces gènes faisaient également partie de variants structuraux dont les extrémités étaient quasiment identiques (Figure 5.8). En observant les mapping anormaux de reads, une première inversion englobant un *CYP6B5* et un *CYP6B1* (LOC106663983 et 82) a été identifiée, suivie d'une duplication en tandem sur le second *CYP6B1* du cluster (LOC106663981).

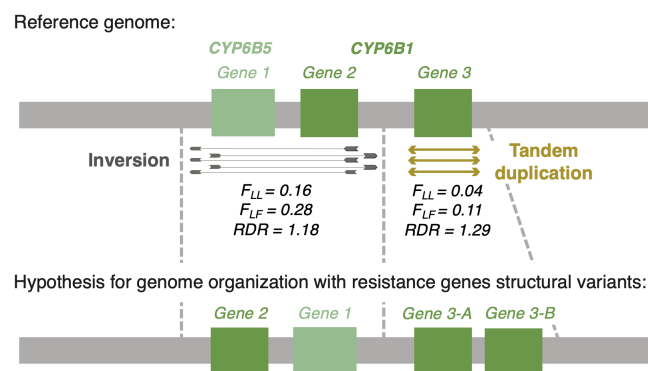


Figure 5.8 – Architecture putative du cluster de cytochromes P450 détecté. Les fréquences de détection des paires de reads supportant l'évènement sont indiquées pour chaque souche (LL, London Lab; LF, London Field), ainsi que le ratio de profondeur entre LF et LL (RDR, Read Depth Ratio).

Néanmoins, l'inversion ne semble pas être particulièrement différenciée entre les souches Londoniennes ($F_{ST} = 0.003$). Quant à la potentielle duplication en tandem, elle n'a pas passé le filtre du ratio de profondeur de couverture entre London Field et London Lab (1,29 alors que le ratio devait être supérieur au quantile 75%, soit 1,33), et n'est donc pas d'être comptabilisée en tant que duplication.

Les variants structuraux détectés autour de ce cluster de P450 semblent être en fréquence relativement faible, y compris dans la souche LF (27% pour l'inversion, 11% pour la duplication sur le pool de 30 individus), ce qui peut expliquer pourquoi aucun signal n'est observé lors de l'analyse transcriptomique de ces gènes. En effet, avec seulement 8 individus échantillonnés pour chaque condition dans l'analyse transcriptomique, nous manquons probablement de puissance pour espérer détecter une éventuelle différence transcriptomique associée.

De manière plus globale sur le superlocus, une très faible réponse transcriptomique (i.e. aucun gène différentiellement exprimé entre les souches) a été observée malgré le grand nombre de SNPs outliers détectés. Ce constat nous a amené à nous interroger sur un autre impact éventuel de ces SNPs, en dehors du nombre d'ARN messagers (transcrits) produits.

5.3.3 Régulation post-transcriptionnelle

Les mutations détectées au cours de l'étude génomique comme étant potentiellement impliquées dans la résistance n'entraînent pas, dans une large majorité des cas, de changement de la séquence protéique. De manière générale, ces mutations n'ont pas non plus été associées à un quelconque changement

transcriptomique.

Deux explications ne s'excluant pas mutuellement peuvent alors être envisagées : (i) les SNPs identifiés ne sont pas associés à la résistance aux insecticides, ou pas directement (mutations neutres entraînées par un balayage sélectif), (ii) la différence de résistance entre les souches n'est médiée ni par des changements protéiques, ni par des changements transcriptomiques, mais par d'autres mécanismes que notre analyse n'a pas mesuré.

En effet, la régulation de l'activité des gènes peut également s'effectuer au niveau post-transcriptionnel lors de leur interaction avec certains ARN non codants (ARNnc, ou ncRNA en anglais). La découverte que la quasi-totalité du génome pouvait être transcrite a tout d'abord déchaîné les passions autour de ces séquences considérées jusque là comme de l'"ARN poubelle" (Djebali *et al.*, 2012). Néanmoins, d'autres études ont rappelé que ces transcrits étant présents à des niveaux très faibles, ce qui pouvait empêcher ces séquences d'avoir une quelconque fonction (Palazzo and Gregory, 2014). L'importance de l'implication de l'ARNnc reste encore soumise à débat : bien que l'on convienne généralement qu'une petite fraction des ARNnc ont des fonctions biologiques importantes, il n'en va pas nécessairement de même pour le reste des ARNnc (Palazzo and Lee, 2015). Cet article propose donc qu'en l'absence de données suggérant le contraire, l'hypothèse nulle appropriée est qu'un ARNnc non caractérisé n'a probablement pas de fonction, plutôt que de lui attribuer un rôle putatif dans le développement et/ou l'homéostasie. Cela étant dit, au delà des ARN de transfert (tRNAs) et ribosomiques (rRNAs), il existe de nombreux exemples dans la littérature ayant associé les ARNnc à des phénomènes adaptatifs. Ces ARNnc englobent les petits ARNs (microRNAs, siRNAs, piRNAs), et les longs ARNs non-codants (lncRNAs). Les lncRNAs sont les ARNnc les plus abondants au sein du génome humain, suivie des microRNAs, et ces deux catégories sont toutes les deux impliquées dans des processus de régulation de gènes (Seal *et al.*, 2020).

Explorer l'implication des ARN non-codants dans la résistance aux insecticides

L'intérêt croissant pour les ARNnc a permis de mettre en lumière quelques exemples de leur implication dans la résistance aux insecticides, en régulant certains gènes candidats.

Une étude récente a ainsi démontré qu'un lncRNA modulait la traduction

d'une enzyme de détoxification GST en entrant en compétition avec la liaison des micro ARN (miRNA). La sur-production de protéines induite confère ainsi une résistance au cyflumétofène chez *Tetranychus cinnabarinus* (Feng *et al.*, 2020). Chez *Aphis gossypii*, des études RT-qPCR et RNAi ont évalué le rôle des lncRNA dans la régulation de l'acétyl-CoA carboxylase, site cible de l'insecticide spirotetramat (Peng *et al.*, 2021). Les miRNAs sont un autre type d'ARNnc qui peuvent affecter la résistance aux insecticides par eux-mêmes, en s'appariant avec des ARNm, réprimant la traduction de la protéine correspondante. Ce phénomène a été observé pour quatre miRNAs chez *Spodoptera frugiperda*, tous significativement sous-exprimés lors du traitement de cet insecte avec trois insecticides différents (Yang *et al.*, 2022). Nous pensons que des études futures devront examiner la présence d'ARNnc à proximité des SNPs candidats identifiés par l'étude génomique, en particulier à l'intérieur du superlocus, ainsi que leur impact putatif sur la production de protéines et, en fin de compte, sur la résistance aux insecticides.

Des ARNnc dans le supergène ?

Le superlocus identifié dans les données génomique comme enrichi en SNPs outliers a donc été associé à la résistance aux pyréthrinoïdes.

Cependant, un seul des SNPs outliers détectés au cours des analyses génomiques était non-synonyme. Les autres variants identifiés étaient répartis dans les introns, les exons, ou l'intergénique. Comme décrit précédemment, une partie de ces variants pourraient tout de même engendrer une régulation positive de gènes impliqués dans la résistance, mais cela ne semble pas être le cas ici, avec seulement une vingtaine de gènes sur-exprimés dans la souche résistante, dont aucun dans ce superlocus. De plus, les gènes exprimés de manière différentielle, que ce soit entre les souches ou entre les conditions non traitées et traitées, n'étaient pas concentrés au niveau de ce superlocus : ils étaient *a contrario* plutôt dispersés le long des chromosomes putatifs. Un seul gène de résistance putatif exprimé de manière différentielle a été trouvé à l'intérieur de cette région d'intérêt : un GST, régulé positivement chez tous les survivants traités.

L'absence d'enrichissement en gènes différentiellement exprimés dans cette région est finalement cohérent avec le faible nombre de duplications qui y ont été détectées dans les données génomiques. De prochains travaux devront donc s'attacher à identifier de potentiels ARNnc à l'intérieur du superlocus corrélativement à la présence des SNPs outliers détectés, et à comprendre leurs

effets putatifs sur les gènes de résistance identifiés. Dans un premier temps, les protéines pourraient être dosées chez les souches sensibles et résistantes aux insecticides, afin d'observer de potentielles différences de traduction entre les souches, par exemple à l'aide du kit All-in-One de BioBasic (permettant d'extraire à partir d'un même individu l'ADN, l'ARN et les protéines).

Ensemble, ces résultats ne remettent également pas en cause la théorie suivant laquelle les inversions seraient en jeu dans l'évolution du "supergène" détecté. En effet, le superlocus n'était significativement enrichi qu'en ce type de variant, ce qui n'est pas incompatible avec l'hypothèse d'ARNnc majoritairement impliqués dans la résistance dans cette région.

5.4 Comprendre pour faciliter la lutte de demain

Comme décrit lors du chapitre introductif, la punaise de lit est un insecte actuellement au cœur de nombreux débats. Au delà de l'approche très théorique de cette thèse, comment aider le grand public à comprendre le lien entre ce type d'études et l'aboutissement à des débouchés plus appliqués? De manière très concrète, qu'est ce que la recherche peut faire pour aider à lutter contre cet insecte nuisible? Deux grands axes seront explorés dans cette partie de discussion : le développement de tests diagnostics, et le potentiel des pathogènes naturels pour une lutte plus écologique.

5.4.1 Développement de tests diagnostics

D'un point de vue appliqué, la connaissance des bases génétiques des mutations impliquées dans la résistance aux insecticides est nécessaire à la conception de tests diagnostics, et donc au suivi de l'évolution de la résistance aux insecticides sur le terrain (Van Leeuwen *et al.*, 2020).

Dans le cas où le supergène détecté chez la punaise de lit serait maintenu par une large inversion, il est possible de réaliser des amorces (primers) sur les bordures de cette inversion (Figure 5.9). Ces primers pourront ensuite être testés par PCR sur les différentes populations résistantes et sensibles aux pyréthrinoïdes.

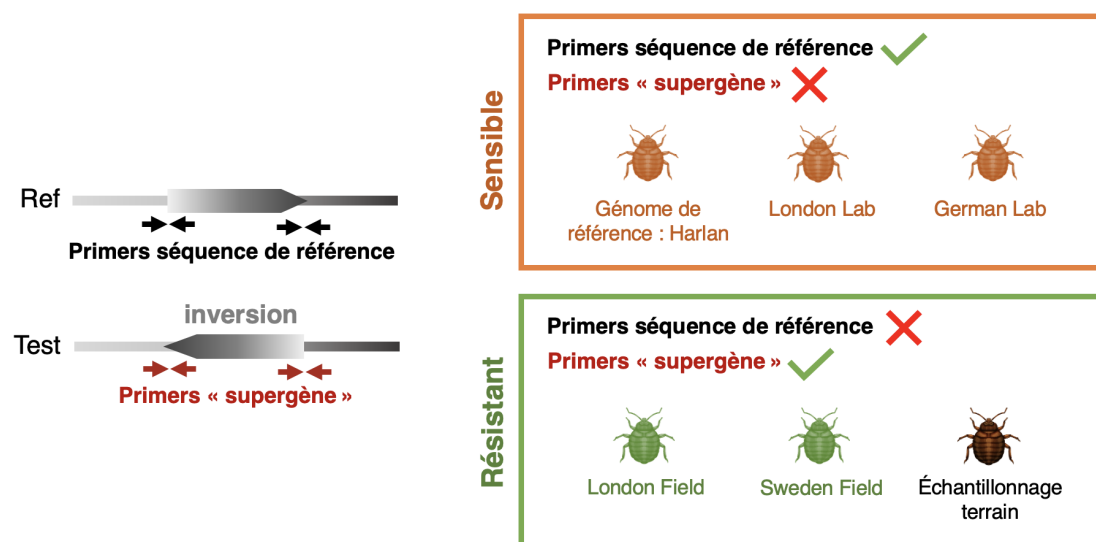


Figure 5.9 – Test diagnostic du supergène putatif. Chaque paire de flèches symbolise la position d'une paire de primers (forward et reverse) sur une bordure de la région du supergène, maintenu par une inversion génomique.

L'attendu est que les primers réalisés en bordure du supergène amplifieraient avec succès des fragments chez les homozygotes pour l'inversion (résistants) tandis que les primers de la séquence de référence ne fonctionneraient pas, et vice-versa chez les homozygotes sensibles. Il serait alors très intéressant de tester ces primers sur des individus de différentes populations de punaises de lit échantillonnées sur le terrain, à corréliser avec le phénotype de résistance aux insecticides en réalisant un test d'exposition à une dose sublétales définie préalablement pour les populations résistantes. Ainsi, il serait possible d'évaluer la répartition de ce mécanisme sur le terrain, ou de mettre en évidence sa spécificité chez les souches London Field et/ou Sweden Field, étant donné que le superlocus a été détecté en les contrastant toutes les deux avec les souches sensibles German Lab et London Lab.

Néanmoins, avec un insecte au mode de dispersion si particulier dépendant principalement des déplacements de son hôte, nous pouvons légitimement nous demander comment un tel mécanisme pourrait se propager entre populations. En effet, nous ne sommes pas dans un système avec de grandes populations panmictiques. Il s'agit plutôt de métapopulations très fragmentées, avec des flux de gènes potentiellement restreints entre elles, et subissant des goulets d'étranglement à répétition lorsqu'un hôte leur permet de coloniser un nouvel habitat. On peut toutefois imaginer qu'une population résistante se propage en moyenne plus qu'une population non-résistante, qui sera rapidement maîtrisée dès lors que la lutte sera engagée à l'aide d'insecticides (Figure 5.10).

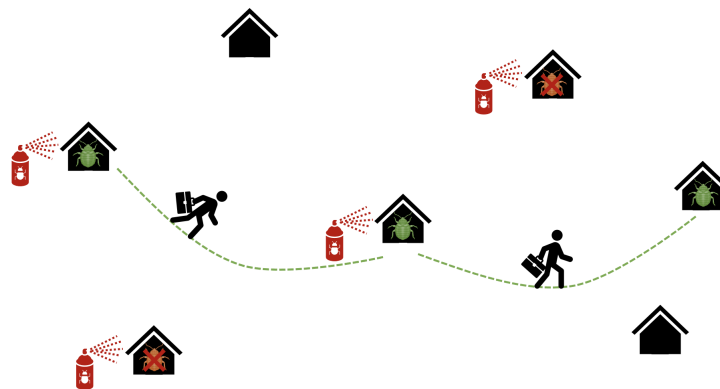


Figure 5.10 – Propagation des punaises de lit entre foyers suivant leur statut de résistance (en vert, résistant, en orange, sensible).

Toutefois, la diversité génétique inter-populations mesurée par les différentes études est très élevée (Booth *et al.*, 2012; Saenz *et al.*, 2012; Fountain *et al.*, 2014; Akhouni *et al.*, 2015). Reste donc à savoir si différents mécanismes ont évolué et se propagent en parallèle, et/ou si un unique mécanisme (comme le superlocus) est largement répandu. Il faudra ensuite définir si ce mécanisme partagé a convergé

chez plusieurs populations de manière indépendante (Karasov *et al.*, 2010), ou si l'on peut retracer une origine unique.

Nous avons vu dans le chapitre introductif qu'une seule femelle pouvait donner naissance à toute une population de punaises de lit au sein d'un nouvel habitat (Polanco *et al.*, 2011a). Pour qu'un mécanisme ait une chance élevée de se propager au sein de l'espèce, il faudrait donc qu'il soit rapidement fixé, au risque que le propagule n'ait pas cet allèle alternatif.

On peut espérer que la mise en place d'un test diagnostique à partir de nos données puisse fonctionner pour détecter des populations résistantes aux pyréthriinoïdes, bien que nous n'ayons aucune certitude quant à la prévalence du superlocus détecté dans nos données *in natura* pour les raisons évoquées ci-dessus. De plus, la résistance à cette famille d'insecticide n'est pas le seul danger qui risque d'entraver la lutte contre cet insecte.

Aujourd'hui, les formations françaises de type Certibiocide évoquées en introduction insistent sur la résistance aux pyréthriinoïdes, et préconisent de changer de famille d'insecticide entre deux passages pour traiter un logement. La plupart conseillent tout de même un produit à base de pyréthriinoïde en premier passage (Permax, K-Othrine...) suivi d'un pyrrole de type chlorfénapyr (Mythic 10SC), ou un carbamate comme le bendiocarbe (Ficam WP). Néanmoins, des niveaux de sensibilité réduits à ces produits ont d'ores et déjà été observés sur le terrain, suggérant que les populations ont déjà évolué une certaine résistance à ces insecticides (Ashbrook *et al.*, 2017; Tawatsin *et al.*, 2011). En particulier, des bio-essais réalisés à Paris en 2018 ont montré de forts taux de résistance au bendiocarbe (Candy *et al.*, 2018). La nécessité d'identifier des marqueurs de résistance à ces autres familles d'insecticide est donc forte, si l'on veut pouvoir également suivre leur propagation, et ajuster les conseils faits aux professionnel.le.s du secteur¹.

Les tests diagnostics peuvent permettre de traquer la résistance sur le terrain, mais comment être sûrs que les tests suivent la vitesse de propagation des résistances? En effet, contrairement à d'autres modèles d'études comme le rat ou le moustique, le mode de dispersion de la punaise de lit et l'absence d'échange entre les différentes sous-populations rend probable l'évolution de mécanismes

1. Dans le cadre de cette thèse, j'ai réalisé avec l'entreprise IZInnovation des formations pour sensibiliser les professionnel.le.s du secteur, ainsi qu'un podcast en partenariat avec Hamelin Info (<https://www.hamelin.info/actus/desinsectisation/podcast-punaise-de-lit-sang-larme/>).

très diversifiés. Comment savoir si un test est négatif parce que la population est sensible, ou parce qu'un autre mécanisme de résistance que celui dont la présence est testée existe?

Finalement, le meilleur moyen de gérer l'évolution de résistance et de stopper cette course aux armements est peut-être de diversifier les outils de lutte, en axant la recherche sur le développement de nouvelles solutions innovantes de bio-contrôle.

5.4.2 La recherche dans de beaux draps !

Dans le dernier chapitre de cette thèse, nous avons tenté d'identifier de potentiels agents pathogènes de la punaise de lit. En effet, outre la problématique des élevages de ce modèle d'étude, la résistance croissante aux insecticides rend nécessaire la détection d'agents de lutte spécifiques de la punaise de lit. En introduction, nous avons rapidement évoqué les pistes déjà explorées chez cet insecte en terme de bio-contrôle, et en particulier le champignon entomopathogène *Beauveria bassiana*. Le principal problème rencontré lors de l'utilisation d'un tel champignon afin de contrôler une infestation sur le terrain est la longueur du temps d'action, entre le premier contact et la mort de l'insecte.

La détection d'un agent pathogène spécifique de la punaise de lit et ayant une cinétique d'infection plus rapide pourrait donc ouvrir de nouvelles perspectives très attendues de bio-contrôle.

Néanmoins, beaucoup de travaux doivent encore être conduits une fois le potentiel pathogène identifié, avant de pouvoir l'utiliser sur le terrain. Une fois que l'on s'est assuré qu'un agent était inoffensif pour l'homme, le principal danger avec l'utilisation d'entomopathogènes est bien sûr lié à leur spécificité. Bien que le traitement des punaises de lit soit réalisé en intérieur, le risque d'une contamination extérieure est loin d'être nul, et l'impact potentiel sur d'autres insectes doit donc être évalué. Dans le cas des virus, le risque est moindre, la plupart d'entre eux étant très spécifiques de leur hôte (El-Wakeil *et al.*, 2019). Un autre risque est lié à l'évolution de résistance sous la pression de sélection exercée par le pathogène. En effet, plusieurs cas de résistance d'insectes aux virus les infectant ont été décrits. L'épaississement de la paroi de la membrane péritrophique (à l'intérieur de l'intestin) des larves du lépidoptère *Anticarsia gemmatalis* semble par exemple les rendre plus résistantes au nucleopolyhedrovirus AgMNPV (Levy *et al.*, 2011).

Pour l'heure, nous avons deux principaux candidats soupçonnés de provoquer le phénotype oxydé conduisant à la mort des punaises de lit. Le premier est un réovirus, genre potentiellement entomopathogène, comme chez le criquet migrateur *Locusta migratoria* (Xu *et al.*, 2022). Ce n'est pas la première fois qu'un réovirus est identifié dans des populations de punaises de lit : la première occurrence remonte ainsi aux années 80, lorsqu'un réovirus a été observé en microscopie électronique au sein d'une souche élevée en laboratoire (Eley *et al.*, 1987). Cette étude a détecté un grand nombre de particules virale dans le cytoplasme de cellules intestinales, suggérant que ce virus infecte bien la punaise de lit. Un réovirus a également été détecté en analysant des punaises de lit échantillonnées sur le terrain à Belgrade (Serbie) en 2016 (Stanojević *et al.*, 2020). Le second candidat est une bactérie du genre *Alcaligenes*, en revanche jamais identifiée auparavant chez la punaise de lit à notre connaissance.

Bien que la détection de ces candidats ne semble pas être limitée aux individus mourant avec un phénotype oxydé au sein des élevages, ils peuvent représenter d'intéressant candidats à la formulation future de produits de bio-contrôle, et nécessitent donc de faire l'objet d'études plus approfondies. Dans un premier temps, il serait intéressant de mettre au point un test PCR pour les détecter, puis de mesurer par PCR quantitative l'abondance de ces candidats chez des punaises saines et présentant le phénotype oxydé. Il peut ensuite être envisagé d'ajouter aux repas de sang les micro-organismes candidats, ou différents antiviraux/antibiotiques ciblés, afin de corréliser leur présence avec la mortalité.

Nous espérons que l'agent infectieux potentiellement responsable du phénotype oxydé pourra être identifié dans un futur proche, permettant ainsi d'éviter les gros épisodes de mortalité entravant l'élevage de la punaise de lit, et facilitant la réalisation des recherches sur cet insecte d'intérêt public. Les efforts de détection de ce pathogène pourront également ouvrir de nouvelles pistes de bio-contrôle de la punaise de lit, nous permettant peut-être de bien dormir dans nos draps...

5.5 Conclusion générale

Au cours de cette thèse, nous avons détecté de nombreuses mutations pouvant impacter la résistance aux pyréthrinoïdes chez la punaise de lit. Grâce au séquençage de génomes complets, nous avons pu analyser les séquences sous-jacentes aux différents phénotypes de résistance. Cela a permis d'identifier pour la première fois la présence de variants structuraux, tels que des duplications et des inversions, potentiellement impliqués dans la résistance aux pyréthrinoïdes chez cet insecte. La détection d'un superlocus, de manière cohérente avec une précédente étude, n'a néanmoins pas été corrélée avec des SNPs non-synonymes, ni avec des différentiels d'expression particuliers, mettant en évidence la nécessité d'explorer les réponses post-transcriptionnelles pouvant être en jeu dans cette région. L'hypothèse d'un supergène associé au phénotype de résistance a été formulée, mais la répartition de ce supergène dans les populations de terrain reste énigmatique. Cette thèse pose la question plus globale de l'adaptation d'un organisme parasite de l'homme se dispersant grâce à son hôte, subissant de forts goulots d'étranglement et n'ayant potentiellement aucun flux de gènes entre populations. Pour répondre à ces questions, un effort d'échantillonnage des souches de terrain devra être réalisé, ainsi qu'une amélioration de l'assemblage du génome de *C. lectularius* à l'échelle chromosomique. Enfin, la perte d'une grande partie de l'élevage à cause d'une mortalité incontrôlable pourrait se transformer en espoir de lutte contre ce nuisible, en dessinant une future opportunité de bio-contrôle.

Bibliographie

- Aak, A., Hage, M., and Rukke, B. A. 2018. Insect pathogenic fungi and bed bugs : behaviour, horizontal transfer and the potential contribution to IPM solutions. *Journal of Pest Science*, 91(2) : 823–835. *Cité page 41*
- Abd-Alla, A. M. M., Kariithi, H. M., Parker, A. G., Robinson, A. S., Kiflom, M., Bergoin, M., and Vreysen, M. J. B. 2010. Dynamics of the salivary gland hypertrophy virus in laboratory colonies of *Glossina pallidipes* (Diptera : Glossinidae). *Virus research*, 150(1-2) : 103–110. *Cité page 165*
- Abd-Alla, A. M. M., Parker, A. G., Vreysen, M. J. B., and Bergoin, M. 2011. Tsetse Salivary Gland Hypertrophy Virus : Hope or Hindrance for Tsetse Control? *PLOS Neglected Tropical Diseases*, 5(8) : e1220. *Cité page 166*
- Abd-Alla, A. M. M., Adun, H., Parker, A. G., Vreysen, M. J. B., and Bergoin, M. 2012. The Antiviral Drug Valacyclovir Successfully Suppresses Salivary Gland Hypertrophy Virus (SGHV) in Laboratory Colonies of *Glossina pallidipes*. *PLOS ONE*, 7(6) : e38417. *Cité page 166*
- Abd-Alla, A. M. M., Kariithi, H. M., Mohamed, A. H., Lapid, E., Parker, A. G., and Vreysen, M. J. B. 2013. Managing Hytrosavirus Infections in *Glossina pallidipes* Colonies : Feeding Regime Affects the Prevalence of Salivary Gland Hypertrophy Syndrome. *PLOS ONE*, 8(5) : e61875. *Cité page 165*
- Abd-Alla, A. M. M., Kariithi, M. H., and Bergoin, M. 2021. Managing pathogens in insect mass-rearing for the Sterile Insect Technique, with the tsetse fly Salivary Gland Hypertrophy Virus as an example. In *Sterile Insect Technique. Principles and Practice in Area-Wide Integrated Pest Management*, chapter 3.3, page 317–354. CRC Press, 2nd ed edition. *Cité pages 142, 166*
- Achee, N. L., Grieco, J. P., Vatandoost, H., Seixas, G., Pinto, J., Ching-Ng, L., Martins, A. J., Juntarajumnong, W., Corbel, V., Gouagna, C., David, J. P., Logan, J. G., Orsborne, J., Marois, E., Devine, G. J., and Vontas, J. 2019. Alternative strategies for mosquito-borne arbovirus control. *PLoS Neglected Tropical Diseases*, 13(1) : 1–22. *Cité page 32*
- Ahmad, M., Denholm, I., and Bromilow, R. H. 2006. Delayed cuticular penetration and enhanced metabolism of deltamethrin in pyrethroid-resistant strains of *Helicoverpa armigera* from China and Pakistan. *Pest management science*, 62(9) : 805–810. *Cité pages 49, 189*
- Akhoundi, M., Kengne, P., Cannet, A., Brengues, C., Berenger, J. M., Izri, A., Marty, P., Simard, F., Fontenille, D., and Delaunay, P. 2015. Spatial genetic structure and restricted gene flow in bed bugs (*Cimex lectularius*) populations in France. *Infection, Genetics and Evolution*, 34 : 236–243. *Cité pages 31, 54, 175, 196*
- Akhoundi, M., Bruel, C., and Izri, A. 2019. Harmful Effects of Bed Bug-Killing Method of Diatomaceous Earth on Human Health. *Journal of Insect Science*, 19(5) : 13. *Cité page 40*
- Amezian, D., Nauen, R., and Le Goff, G. 2021. Transcriptional regulation of xenobiotic detoxification genes in insects - An overview. *Pesticide Biochemistry and Physiology*, 174 : 104822. *Cité pages 51, 54, 55*

- Andersona, W. W., Arnolda, J., Baldwinb, D. G., Beckenbachc, A. T., Browna, C. J., Bryantd, S. H., Coyne, J. A., Harshman, L. G., Heedb, W. B., Jefferyg, D. E., Klaczkoh, L. B., Moore, B. C., Portera, J. M., Powellj, J. R., Proutf, T., Schaefferk, S. W., Stephens, J. C., Taylorm, C. E., Turnern, M. E., Williams, G., and Moore, J. A. 1991. Four decades of inversion polymorphism in *Drosophilapseudoobscura*. *Proc Natl Acad Sci USA*, 88(November) : 1–5.
Cité page 177
- Anses 2014. Avis de l'Anses Relatif aux substances actives biocides pouvant être utilisées dans le cadre de la prévention d'une épidémie de Chikungunya en Guyane. Technical Report 0, Anses.
Cité page 43
- Arouri, R., Le Goff, G., Hemden, H., Navarro-Llopis, V., M'saad, M., Castañera, P., Feyereisen, R., Hernández-Crespo, P., and Ortego, F. 2015. Resistance to lambda-cyhalothrin in Spanish field populations of *Ceratitis capitata* and metabolic resistance mediated by P450 in a resistant strain. *Pest management science*, 71(9) : 1281–1291.
Cité page 51
- Ashbrook, A. R., Scharf, M. E., Bennett, G. W., and Gondhalekar, A. D. 2017. Detection of Reduced Susceptibility to Chlorfenapyr- and Bifenthrin-Containing Products in Field Populations of the Bed Bug (Hemiptera : Cimicidae). *Journal of Economic Entomology*, 110(3) : 1195–1202.
Cité page 197
- Ashford, R. 2000. Parasites as indicators of human biology and evolution. *Journal of Medical Microbiology*, 49(9) : 771–772.
Cité page 28
- Assogba, B. S., Djogbénu, L. S., Milesi, P., Berthomieu, A., Perez, J., Ayala, D., Chandre, F., Makoutodé, M., Labbé, P., and Weill, M. 2015. An ace-1 gene duplication resorbs the fitness cost associated with resistance in *Anopheles gambiae*, the main malaria mosquito. *Scientific Reports*, 5(May) : 19–21.
Cité page 65
- Awori, R. M. 2022. Nematophilic bacteria associated with entomopathogenic nematodes and drug development of their biomolecules. *Frontiers in Microbiology*, 13(September) : 1–11.
Cité page 147
- Ayala, D., Ullastres, A., and González, J. 2014. Adaptation through chromosomal inversions in *Anopheles*. *Frontiers in Genetics*, 5(MAY) : 1–10.
Cité page 64
- Bacot, T., Haberkorn, C. ., Guilliet, J. ., Cattel, J. , Keffi, M., Filee, J., Cumer, T., Boyer, F., Gaude, T., Laporte, F., Tutagata, J., Nadalin, L., Vontas, J., Dusfour, L., Bonneville, J. ., and David, J. . 2023. Duplication of a P450 gene cluster confers insecticide resistance in the dengue mosquito *Aedes aegypti*. *in prep.*
Cité page 181
- Balabanidou, V., Kampouraki, A., MacLean, M., Blomquist, G. J., Tittiger, C., Juárez, M. P. g., Mijailovsky, S. J., Chalepakis, G., Anthousi, A., Lynd, A., Antoine, S., Hemingway, J., Ranson, H., Lycett, G. J., and Vontas, J. 2016. Cytochrome P450 associated with insecticide resistance catalyzes cuticular hydrocarbon production in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences*, 113(33) : 9268–9273.
Cité page 50
- Balabanidou, V., Grigoraki, L., and Vontas, J. 2018. Insect cuticle : a critical determinant of insecticide resistance. *Current Opinion in Insect Science*, 27 : 68–74.
Cité pages 49, 50
- Balachandran, P. and Beck, C. R. 2020. Structural variant identification and characterization. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 28(1) : 31–47.
Cité page 63

- Balmand, S., Lohs, C., Aksoy, S., and Heddi, A. 2013. Tissue distribution and transmission routes for the tsetse fly endosymbionts. *Journal of invertebrate pathology*, 112 : S116–S122. *Cité page 156*
- Balvín, O. and Booth, W. 2018. Distribution and frequency of pyrethroid resistance-associated mutations in host lineages of the bed bug (Hemiptera : Cimicidae) across Europe. *Journal of Medical Entomology*, 55(4) : 923–928. *Cité pages 54, 175*
- Barbarin, A. M., Jenkins, N. E., Rajotte, E. G., and Thomas, M. B. 2012. A preliminary evaluation of the potential of *Beauveria bassiana* for bed bug control. *Journal of Invertebrate Pathology*, 111(1) : 82–85. *Cité page 41*
- Barbarin, A. M., Bellicanta, G. S., Osborne, J. A., Schal, C., and Jenkins, N. E. 2017. Susceptibility of insecticide-resistant bed bugs (*Cimex lectularius*) to infection by fungal biopesticide. *Pest Management Science*, 73(8) : 1568–1573. *Cité page 41*
- Bariami, V., Jones, C. M., Poupardin, R., Vontas, J., and Ranson, H. 2012. Gene Amplification, ABC Transporters and Cytochrome P450s : Unraveling the Molecular Basis of Pyrethroid Resistance in the Dengue Vector, *Aedes aegypti*. *PLOS Neglected Tropical Diseases*, 6(6) : 1–10. *Cité page 52*
- Barrès, B., Micoud, A., Corio-Costet, M.-F., Debieu, D., Fillinger, S., Walker, A.-S., Délye, C., Grosman, J., and Siegwart, M. 2016. Trends and Challenges in Pesticide Resistance Detection. *Trends in Plant Science*, 21(10) : 834–853. *Cité pages 46, 50*
- Bass, C. and Field, L. M. 2011. Gene amplification and insecticide resistance. *Pest Management Science*, 67(8) : 886–890. *Cité page 52*
- Bastarache, P., Wajnberg, G., Dumas, P., Chacko, S., Lacroix, J., Crapoulet, N., Moffat, C. E., and Morin, P. 2020. Transcriptomics-Based Approach Identifies Spinosad-Associated Targets in the Colorado Potato Beetle, *Leptinotarsa decemlineata*. *Cité page 188*
- Behjati, S. and Tarpey, P. S. 2013. What is next generation sequencing? *Arch Dis Child Educ Pract Ed.*, 98(6) : 236–238. *Cité page 61*
- Benoit, J. B., Adelman, Z. N., Reinhardt, K., Dolan, A., Poelchau, M., Jennings, E. C., Szuter, E. M., Hagan, R. W., Gujar, H., Shukla, J. N., Zhu, F., Mohan, M., Nelson, D. R., Rosendale, A. J., Derst, C., Resnik, V., Wernig, S., Menegazzi, P., Wegener, C., Peschel, N., Hendershot, J. M., Blenau, W., Predel, R., Johnston, P. R., Ioannidis, P., Waterhouse, R. M., Nauen, R., Schorn, C., Ott, M. C., Maiwald, F., Johnston, J. S., Gondhalekar, A. D., Scharf, M. E., Peterson, B. F., Raje, K. R., Hottel, B. A., Armisén, D., Crumière, A. J. J., Refki, P. N., Santos, M. E., Sghaier, E., Viala, S., Khila, A., Ahn, S. J., Childers, C., Lee, C. Y., Lin, H., Hughes, D. S., Duncan, E. J., Murali, S. C., Qu, J., Dugan, S., Lee, S. L., Chao, H., Dinh, H., Han, Y., Doddapaneni, H., Worley, K. C., Muzny, D. M., Wheeler, D., Panfilio, K. A., Vargas Jentzsch, I. M., Vargo, E. L., Booth, W., Friedrich, M., Weirauch, M. T., Anderson, M. A., Jones, J. W., Mittapalli, O., Zhao, C., Zhou, J. J., Evans, J. D., Attardo, G. M., Robertson, H. M., Zdobnov, E. M., Ribeiro, J. M., Gibbs, R. A., Werren, J. H., Palli, S. R., Schal, C., and Richards, S. 2016. Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature Communications*, 7 : 1–10. *Cité page 36*
- Berenbaum, M. R. 2002. Postgenomic chemical ecology : From genetic code to ecological interactions. *Journal of Chemical Ecology*, 28(5) : 873–896. *Cité page 179*

- Bøgh, C., Pedersen, E. M., Mukoko, D. A., and Ouma, J. H. 1998. Permethrin-impregnated bednet effects on resting and feeding behaviour of lymphatic filariasis vector mosquitoes in Kenya. *Medical and Veterinary Entomology*, 12(1) : 52–59. *Cité page 49*
- Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic : A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15) : 2114–2120. *Cité page 144*
- Booth, W., Saenz, V. L., Santangelo, R. G., Wang, C., Schal, C., and Vargo, E. L. 2012. Molecular markers reveal infestation dynamics of the bed bug (Hemiptera : Cimicidae) within apartment buildings. *Journal of medical entomology*, 49(3) : 535–546. *Cité pages 30, 31, 196*
- Breitwieser, F. P. and Salzberg, S. L. 2020. Pavian : interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics*, 36(4) : 1303–1304. *Cité page 153*
- Brooke, B. D., Hunt, R. H., Chandre, F., Carnevale, P., and Coetzee, M. 2002. Stable Chromosomal Inversion Polymorphisms and Insecticide Resistance in the Malaria Vector Mosquito *Anopheles gambiae* (Diptera : Culicidae) . *Journal of Medical Entomology*, 39(4) : 568–573. *Cité page 178*
- Brüssow, H. and Desiere, F. 2001. Comparative phage genomics and the evolution of Siphoviridae : insights from dairy phages. *Molecular Microbiology*, 39(2) : 213–223. *Cité page 148*
- Burrows, S., Perron, S., and Susser, S. 2013. Suicide following an infestation of bed bugs. *The American journal of case reports*, 14 : 176–178. *Cité page 33*
- Busvine, J. R. 1958. Insecticide-resistance in bed-bugs. *Bulletin of the World Health Organization*, 19(6) : 1041–1052. *Cité page 56*
- Cáceres, M., Santo-Orihuela, P. L., and Vassena, C. V. 2019. Evaluation of Resistance to Different Insecticides and Metabolic Detoxification Mechanism by Use of Synergist in the Common Bed Bug (Heteroptera : Cimicidae). *Journal of medical entomology*, 56(5) : 1324–1330. *Cité page 46*
- Caffrey, M. and Cherezov, V. 2009. Crystallizing membrane proteins using lipidic mesophases. *Nature protocols*, 4(5) : 706–731. *Cité page 47*
- Calla, B., Demkovich, M., Siegel, J. P., Viana, J. P. G., Walden, K. K. O., Robertson, H. M., and Berenbaum, M. R. 2021. Selective Sweeps in a Nutshell : The Genomic Footprint of Rapid Insecticide Resistance Evolution in the Almond Agroecosystem. *Genome Biology and Evolution*, 13(1) : evaa234. *Cité pages 179, 190*
- Campbell, P. M., Newcomb, R. D., Russell, R. J., and Oakeshott, J. G. 1998. Two different amino acid substitutions in the *ali*-esterase, E3, confer alternative types of organophosphorus insecticide resistance in the sheep blowfly, *Lucilia cuprina*. *Insect Biochemistry and Molecular Biology*, 28(3) : 139–150. *Cité page 52*
- Candy, K., Akhoundi, M., Bruel, C., and Izri, A. 2018. Ineffectiveness of Insecticide Bendiocarb Against a *Cimex lectularius* (Hemiptera : Cimicidae) Population in Paris, France. *Journal of Medical Entomology*, 55(6) : 1648–1650. *Cité page 197*
- Cannavò, E., Khoueiry, P., Garfield, D., Geeleher, P., Zichner, T., Gustafson, E., Ciglar, L., Korbel, J., and Furlong, E. 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. *Current Biology*, 26(1) : 38–51. *Cité page 189*

- Cannet, A., Akhouni, M., Berenger, J. M., Michel, G., Marty, P., and Delaunay, P. 2015. A review of data on laboratory colonies of bed bugs (Cimicidae), an insect of emerging medical relevance. *Parasite*, 22. *Cité page 37*
- Carrasco, D., Lefèvre, T., Moiroux, N., Penner, C., Chandre, F., and Cohuet, A. 2019. Behavioural adaptations of mosquito vectors to insecticide control. *Current Opinion in Insect Science*, 34 : 48–54. *Cité page 48*
- Carson, R. 1962. *Silent Spring*. Houghton Mifflin, Boston. *Cité page 43*
- Carvalho, R. A., Azeredo-Espin, A. M. L., and Torres, T. T. 2010. Deep sequencing of New World screw-worm transcripts to discover genes involved in insecticide resistance. *BMC Genomics*, 11(1) : 695. *Cité page 63*
- Cattell, J., Faucon, F., Le Péron, B., Sherpa, S., Monchal, M., Grillet, L., Gaude, T., Laporte, F., Dusfour, I., Reynaud, S., and David, J. P. 2020. Combining genetic crosses and pool targeted DNA-seq for untangling genomic variations associated with resistance to multiple insecticides in the mosquito *Aedes aegypti*. *Evolutionary Applications*, 13(April) : 303–317. *Cité pages 65, 66*
- Cattell, J., Haberkorn, C., Laporte, F., Gaude, T., Cumer, T., Renaud, J., Sutherland, I. W., Hertz, J. C., Bonneville, J. M., Arnaud, V., Fustec, B., Boyer, S., Marcombe, S., and David, J. P. 2021. A genomic amplification affecting a carboxylesterase gene cluster confers organophosphate resistance in the mosquito *Aedes aegypti* : From genomic characterization to high-throughput field detection. *Evolutionary Applications*, 14(4) : 1009–1022. *Cité page 65*
- Centre National d'Expertise sur les Vecteurs (Institution) 2015. Punaises de lit en France : état des lieux et recommandations. Technical report, Anses. *Cité page 43*
- Champion, C. J. and Xu, J. 2018. Redox state affects fecundity and insecticide susceptibility in *Anopheles gambiae*. *Scientific Reports*, 8(1) : 1–11. *Cité pages 54, 55*
- Chang, K. P. and Musgrave, A. J. 1973. Morphology, histochemistry, and ultrastructure of mycetome and its rickettsial symbiotes in *Cimex lectularius* L. *Canadian Journal of Microbiology*, 19(9) : 1075–1081. *Cité page 150*
- Charkowski, A. O. 2015. Biology and control of *Pectobacterium* in potato. *American Journal of Potato Research*, 92(2) : 223–229. *Cité page 155*
- Chen, H., Patterson, N., and Reich, D. 2010. Population differentiation as a test for selective sweeps. *Genome Research*, 20(3) : 393–402. *Cité page 176*
- Chiapello, M., Bosco, L., Ciuffo, M., Ottati, S., Salem, N., Rosa, C., Tavella, L., and Turina, M. 2021. Complexity and Local Specificity of the Virome Associated with Tospovirus-Transmitting Thrips Species. *Journal of Virology*, 95(21) : 00597–21. *Cité page 158*
- Chrudimský, T., Husník, F., Nováková, E., and Hypša, V. 2012. *Candidatus Sodalis melophagi* sp. nov. : Phylogenetically Independent Comparative Model to the Tsetse Fly Symbiont *Sodalis glossinidius*. *PLOS ONE*, 7(7) : e40354. *Cité page 156*
- Conrad, D. F. and Hurles, M. E. 2007. The population genetics of structural variation. *Nature Genetics*, 39(7) : S30–S36. *Cité page 62*
- Crawley, S. E. and Borden, J. H. 2021. Detection and monitoring of bed bugs (Hemiptera : Cimicidae) : review of the underlying science, existing products and future prospects. *Pest Management Science*, 77(12) : 5334–5346. *Cité page 40*

- Crossley, M. S., Chen, Y. H., Groves, R. L., and Schoville, S. D. 2017. Landscape genomics of Colorado potato beetle provides evidence of polygenic adaptation to insecticides. *Molecular Ecology*, 26(22) : 6284–6300. *Cité page 57*
- Dang, K., Toi, C. S., Lilly, D. G., Bu, W., and Doggett, S. L. 2015. Detection of knockdown resistance mutations in the common bed bug, *Cimex lectularius* (Hemiptera : Cimicidae), in Australia. *Pest Management Science*, 71(7) : 914–922. *Cité pages 54, 175*
- Dang, K., Doggett, S. L., Veera Singham, G., and Lee, C. Y. 2017. Insecticide resistance and resistance mechanisms in bed bugs, *Cimex* spp. (Hemiptera : Cimicidae). *Parasites and Vectors*, 10(1) : 1–31. *Cité pages 49, 176*
- Dang, K. A. I., Toi, C. S., Lilly, D. G., Lee, C.-y., Naylor, R., Tawatsin, A., Thavara, U., Bu, W., and Doggett, S. L. 2014. Detection of knockdown resistance (kdr) in *Cimex lectularius* and *Cimex hemipterus* (Hemiptera : Cimicidae). *The Eighth International Conference on Urban Pests*, 71(February 2018) : 914–922. *Cité pages 54, 175*
- Darwin, C. 1866. *De l'origine des espèces : par sélection naturelle ou des lois de transformation des êtres organisés ...* Guillaumin et cie. *Cité page 66*
- Davies, T. G., Field, L. M., and Williamson, M. S. 2012. The re-emergence of the bed bug as a nuisance pest : Implications of resistance to the pyrethroid insecticides. *Medical and Veterinary Entomology*, 26(3) : 241–254. *Cité page 173*
- Davison, H. R., Pilgrim, J., Wybouw, N., Parker, J., Pirro, S., Hunter-Barnett, S., Campbell, P. M., Blow, F., Darby, A. C., Hurst, G. D. D., and Siozios, S. 2022. Genomic diversity across the *Rickettsia* and 'Candidatus Megaira' genera and proposal of genus status for the *Torix* group. *Nature Communications*, 13(1) : 2630. *Cité page 148*
- Daïter, A. B. 1963. [Experimental infection of arthropods with human Q fever]. *Trudy Leningradskogo instituta epidemiologii i mikrobiologii im. Pastera*, 25 : 92–100. *Cité page 32*
- Degnan, P. H., Bittleston, L. S., Hansen, A. K., Sabree, Z. L., Moran, N. A., and Almeida, R. P. P. 2011. Origin and Examination of a Leafhopper Facultative Endosymbiont. *Current Microbiology*, 62(5) : 1565–1572. *Cité page 157*
- Delaunay, P., Berenger, J.-M., Izri, A., and Chosidow, O. 2010. Les punaises de lit. *Riviera Scientifique*, pages 91–106. *Cité page 32*
- Delaunay, P., Blanc, V., Del Giudice, P., Levy-Bencheton, A., Chosidow, O., Marty, P., and Brouqui, P. 2011. Bedbugs and infectious diseases. *Clinical Infectious Diseases*, 52(2) : 200–210. *Cité page 32*
- Després, L., David, J. P., and Gallet, C. 2007. The evolutionary ecology of insect resistance to plant chemicals. *Trends in Ecology and Evolution*, 22(6) : 298–307. *Cité pages 50, 179*
- DeVries, Z. C., Reid, W. R., Kells, S. A., and Appel, A. G. 2015. Effects of starvation on deltamethrin tolerance in bed bugs, *Cimex lectularius* L. (Hemiptera : Cimicidae). *Insects*, 6(1) : 102–111. *Cité page 48*
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., Alioto, T., Antoshechkin, I., Baer, M. T., Bar, N. S., Batut, P., Bell, K., Bell, I., Chakraborty, S., Chen, X., Chrast, J., Curado, J.,

- Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M. J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Luo, O. J., Park, E., Persaud, K., Preall, J. B., Ribeca, P., Risk, B., Robyr, D., Sammeth, M., Schaffer, L., See, L.-H., Shahab, A., Skancke, J., Suzuki, A. M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S. E., Hannon, G., Giddings, M. C., Ruan, Y., Wold, B., Carninci, P., Guigó, R., and Gingeras, T. R. 2012. Landscape of transcription in human cells. *Nature*, 489(7414) : 101–108. *Cité pages 62, 192*
- Dobzhansky, T. and Sturtevant, A. H. 1938. Inversions in the chromosomes of *Drosophila pseudoobscura*. *Genetics*, 23(1) : 28. *Cité page 177*
- Doggett, S. L. and Russell, R. C. 2009. Emerging Challenges in Bed Bug Management Emerging Pest Management Challenges. Technical report, Department of Medical Entomology, University of Sydney and ICPMR, Westmead Hospital, Westmead, NSW 2145. *Cité page 41*
- Doggett, S. L., Geary, M. J., and Russell, R. C. 2004. The resurgence of bed bugs in Australia : with notes on their ecology and control. *Environmental health*, 4(2) : 30–38. *Cité page 29*
- Dong, K., Du, Y., Rinkevich, F., Nomura, Y., Xu, P., Wang, L., Silver, K., and Zhorov, B. S. 2014. Molecular biology of insect sodium channels and pyrethroid resistance. *Insect Biochemistry and Molecular Biology*, 50 : 1–17. *Cité page 185*
- Dong, K. E. and Scott, J. G. 1994. Linkage of kdr-type resistance and the para-homologous sodium channel gene in German cockroaches (*Blattella germanica*). *Insect biochemistry and molecular biology*, 24(7) : 647–654. *Cité pages 54, 186*
- Durand, R., Cannet, A., Berdjane, Z., Bruel, C., Haouchine, D., Delaunay, P., and Izri, A. 2012. Infestation by pyrethroids resistant bed bugs in the suburb of Paris, France. *Parasite (Paris, France)*, 19(4) : 381–387. *Cité page 54*
- Eberle, K. E., Sayed, S., Rezapanah, M., Shojai-Estabragh, S., and Jehle, J. A. 2009. Diversity and evolution of the *Cydia pomonella* granulovirus. *Journal of General Virology*, 90(3) : 662–671. *Cité page 156*
- El-Wakeil, N., Saleh, M., and Abu-hashim, M. 2019. *Cottage Industry of Biocontrol Agents and Their Applications : Practical Aspects to Deal Biologically with Pests and Stresses Facing Strategic Crops*. Springer. *Cité page 198*
- Eley, S. M., Gardner, R., Molyneux, D. H., and Moore, N. F. 1987. A reovirus from the bedbug, *Cimex lectularius*. *Journal of General Virology*, 68(1) : 195–199. *Cité pages 36, 199*
- Engel, M. S. 2008. A stem-group cimicid in mid-Cretaceous amber from Myanmar (Hemiptera : Cimicoidea). *Alavesia*, 2 : 233–237. *Cité page 26*
- FAO/IAEA 1987. Activities of the entomology unit. Technical report, Food and Agriculture Organization of the United Nations/International Atomic Energy Agency, Report No. 2. *Cité page 142*
- Faucon, F., Dusfour, I., Gaude, T., Navratil, V., Boyer, F., Chandre, F., Sirisopa, P., Thanispong, K., Juntarajumnong, W., Poupardin, R., Chareonviriyaphap, T., Girod, R., Corbel, V., Reynaud, S., and David, J.-p. 2015. Unravelling genomic changes associated with insecticide resistance in the dengue mosquito *Aedes aegypti* by deep targeted sequencing. *Genome Research*, pages 1–13. *Cité page 65*

- Feldlaufer, M. F., O'Connor, L.-L., and Ulrich, K. R. 2018. Bed Bug Laboratory Maintenance. In *Advances in the Biology and Management of Modern Bed Bugs*, chapter 21, pages 199–208. John Wiley & Sons, Ltd. *Cité page 38*
- Feng, K., Liu, J., Wei, P., Ou, S., Wen, X., Shen, G., Xu, Z., Xu, Q., and He, L. 2020. lincRNA_Tc13743.2-miR-133-5p-TcGSTm02 regulation pathway mediates cyflumetofen resistance in *Tetranychus cinnabarinus*. *Insect Biochemistry and Molecular Biology*, 123 : 103413. *Cité page 193*
- Ffrench-Constant, R. H., Roush, R. T., Mortlock, D., and Dively, G. P. 1990. Isolation of Dieldrin Resistance from Field Populations of *Drosophila melanogaster* (Diptera : Drosophilidae). *Journal of Economic Entomology*, 83(5) : 1733–1737. *Cité page 57*
- Ffrench-Constant, R. H., Daborn, P. J., and Le Goff, G. 2004. The genetics and genomics of insecticide resistance. *Trends in Genetics*, 20(3) : 163–170. *Cité page 57*
- Fischer, K., Beatty, W. L., Weil, G. J., and Fischer, P. U. 2014. High Pressure Freezing/Freeze Substitution Fixation Improves the Ultrastructural Assessment of Wolbachia Endosymbiont – Filarial Nematode Host Interaction. *PLOS ONE*, 9(1) : e86383. *Cité page 150*
- Fisher, M. L., Levine, J. F., Guy, J. S., Mochizuki, H., Breen, M., Schal, C., and Watson, D. W. 2019. Lack of influence by endosymbiont *Wolbachia* on virus titer in the common bed bug, *Cimex lectularius*. *Parasites & vectors*, 12(1) : 436. *Cité page 160*
- Foster, S. P., Young, S., Williamson, M. S., Duce, I., Denholm, I., and Devine, G. J. 2003. Analogous pleiotropic effects of insecticide resistance genotypes in peach–potato aphids and houseflies. *Heredity*, 91(2) : 98–106. *Cité page 59*
- Fountain, T., Duvaux, L., Horsburgh, G., Reinhardt, K., and Butlin, R. K. 2014. Human-facilitated metapopulation dynamics in an emerging pest species, *Cimex lectularius*. *Molecular Ecology*, 23(5) : 1071–1084. *Cité pages 31, 196*
- Fountain, T., Ravinet, M., Naylor, R., Reinhardt, K., and Butlin, R. K. 2016. A linkage map and QTL analysis for pyrethroid resistance in the bed bug *cimex lectularius*. *G3 : Genes, Genomes, Genetics*, 6(12) : 4059–4066. *Cité pages 12, 174, 175*
- Freeman, J. C., Smith, L. B., Silva, J. J., Fan, Y., Sun, H., and Scott, J. G. 2021. Fitness studies of insecticide resistant strains : lessons learned and future directions. *Pest Management Science*, 77(9) : 3847–3856. *Cité page 176*
- Garrity, G., Staley, J. T., Boone, D. R., De Vos, P., Goodfellow, M., Rainey, F. A., Schleifer, K. H., Brenner, D. J., and Krieg, N. R. 2006. *Bergey's manual® of systematic bacteriology : The Proteobacteria. Springer Science & Business Media*, 2. *Cité page 154*
- Georghiou, G. P. 1972. The Evolution of Resistance to Pesticides. *Annual Review of Ecology and Systematics*, 3(1) : 133–168. *Cité page 57*
- Ghoda, M. K. and Shah, R. A. 2005. A prospective epidemiological study to see if mosquito bite could be responsible for spread of hepatitis B virus infection. *Tropical gastroenterology : official journal of the Digestive Diseases Foundation*, 26(1) : 29–30. *Cité page 32*

- Gimenez, S., Abdelgaffar, H., Goff, G. L., Hilliou, F., Blanco, C. A., Hänniger, S., Bretaudeau, A., Legeai, F., Nègre, N., Jurat-Fuentes, J. L., d'Alençon, E., and Nam, K. 2020. Adaptation by copy number variation increases insecticide resistance in the fall armyworm. *Communications Biology*, 3(1) : 664. *Cité page 65*
- Goddard, J. and de Shazo, R. 2012. Psychological effects of bed bug attacks (*Cimex lectularius* L.). *The American journal of medicine*, 125(1) : 101–103. *Cité page 33*
- Gonzalez-Morales, M. A. and Romero, A. 2019. Effect of Synergists on Deltamethrin Resistance in the Common Bed Bug (Hemiptera : Cimicidae). *Journal of Economic Entomology*, 112(2) : 786–791. *Cité page 176*
- Gordon, J. R., Potter, M. F., and Haynes, K. F. 2015. Insecticide resistance in the bed bug comes with a cost. *Scientific Reports*, 5 : 1–7. *Cité pages 59, 176*
- Gott, R. C., Kunkel, G. R., Zobel, E. S., Lovett, B. R., and Hawthorne, D. J. 2017. Implicating ABC Transporters in Insecticide Resistance : Research Strategies and a Decision Framework. *Journal of Economic Entomology*, 110(2) : 667–677. *Cité page 52*
- Gries, R., Britton, R., Holmes, M., Zhai, H., Draper, J., and Gries, G. 2015. Bed bug aggregation pheromone finally identified. *Angewandte Chemie - International Edition*, 54(4) : 1135–1138. *Cité page 40*
- Grozeva, S., Kuznetsova, V., and Anokhin, B. 2010. Bed bug cytogenetics : Karyotype, sex chromosome system, FISH mapping of 18S rDNA, and male meiosis in *Cimex lectularius* Linnaeus, 1758 (Heteroptera : Cimicidae). *Comparative Cytogenetics*, 4(2) : 151–160. *Cité page 37*
- Grzywacz, D. 2017. Basic and Applied Research : Baculovirus. In L. A. B. T. M. C. o. I. Lacey and M. Pests, editors, *Microbial Control of Insect and Mite Pests - From Theory to Practice*, chapter 3, pages 27–46. Academic Press. *Cité page 156*
- Gunning, R. V., Moores, G. D., and Devonshire, A. L. 1996. Esterases and esfenvalerate resistance in Australian *Helicoverpa armigera* (Hübner) Lepidoptera : Noctuidae. *Pesticide Biochemistry and Physiology*, 54(1) : 12–23. *Cité page 52*
- Haberkorn, C., David, J.-P., Henri, H., Delpuech, J.-M., Lasseur, R., Vavre, F., and Varaldi, J. 2023. Identification of new candidates for insecticide resistance in bed bugs through genome-wide analysis. *in prep*, xx(xx) : xx. *Cité page 186*
- Hager, E. R. and Hoekstra, H. E. 2021. Tail Length Evolution in Deer Mice : Linking Morphology, Behavior, and Function. *Integrative and comparative biology*, 61(2) : 385–397. *Cité page 178*
- Hager, E. R., Harringmeyer, O. S., Wooldridge, T. B., Theingi, S., Gable, J. T., McFadden, S., Neugeboren, B., Turner, K. M., Jensen, J. D., and Hoekstra, H. E. 2022. A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science*, 377(6604) : 399–405. *Cité pages 68, 178*
- Harrison, R. L. and Lynn, D. E. 2007. Genomic sequence analysis of a nucleopolyhedrovirus isolated from the diamondback moth, *Plutella xylostella*. *Virus Genes*, 35(3) : 857–873. *Cité page 156*
- Hawkes, N. J. and Hemingway, J. 2002. Analysis of the promoters for the beta-esterase genes associated with insecticide resistance in the mosquito *Culex quinquefasciatus*. *Biochimica et biophysica acta*, 1574(1) : 51–62. *Cité page 63*

- Hawkins, N. J., Bass, C., Dixon, A., and Neve, P. 2019. The evolutionary origins of pesticide resistance. *Biological Reviews*, 94(1) : 135–155. *Cité page 66*
- Hedrick, P. W. 2013. Adaptive introgression in animals : Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, 22(18) : 4606–4618. *Cité page 67*
- Hermisson, J. and Pennings, P. S. 2005. Soft Sweeps : Molecular Population Genetics of Adaptation From Standing Genetic Variation. *Genetics*, 169(4) : 2335–2352. *Cité page 67*
- Hingamp, P., Grimsley, N., Acinas, S. G., Clerissi, C., Subirana, L., Poulain, J., Ferrera, I., Sarmiento, H., Villar, E., Lima-Mendez, G., Faust, K., Sunagawa, S., Claverie, J.-M., Moreau, H., Desdevises, Y., Bork, P., Raes, J., de Vargas, C., Karsenti, E., Kandels-Lewis, S., Jaillon, O., Not, F., Pesant, S., Wincker, P., and Ogata, H. 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *The ISME Journal*, 7(9) : 1678–1695. *Cité page 153*
- Hockenyos, G. L. 1940. Sulfur dioxide fumigations. *Pests*, 8(11) : 23–25. *Cité page 42*
- Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X. Y., and Fukatsu, T. 2010. Wolbachia as a bacteriocyte-associated nutritional mutualist. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2) : 769–774. *Cité pages 35, 147, 154, 155, 160*
- Hough, R. B., Lengeling, A., Bedian, V., Lo, C., and Bućan, M. 1998. Rump white inversion in the mouse disrupts dipeptidyl aminopeptidase-like protein 6 and causes dysregulation of Kit expression. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23) : 13800–13805. *Cité page 188*
- How, Y.-F. and Lee, C.-Y. 2010. Survey of bed bugs in infested premises in Malaysia and Singapore. *Journal of Vector Ecology*, 35(1) : 89–94. *Cité page 31*
- Hu, B., He, H., Hu, S., Ren, M., Wei, Q., Tian, X., Elzaki, M. E. A., Bass, C., Su, J., and Palli, S. R. 2021. Changes in both trans- and cis-regulatory elements mediate insecticide resistance in a lepidopteron pest , *Spodoptera exigua*. *PLoS Genetics*, 17(3) : 1–22. *Cité page 55*
- Hwang, S. W., Svoboda, T. J., De Jong, I. J., Kabasele, K. J., and Gogosis, E. 2005. Bed bug infestations in an urban environment. *Emerging infectious diseases*, 11(4) : 533–538. *Cité page 29*
- Hypša, V. and Aksoy, S. 1997. Phylogenetic characterization of two transovarially transmitted endosymbionts of the bedbug *Cimex lectularius* (Heteroptera : Cimicidae). *Insect Molecular Biology*, 6(3) : 301–304. *Cité page 155*
- Ingham, V. A., Elg, S., Nagi, S. C., and Dondelinger, F. 2021a. Capturing the transcription factor interactome in response to sub-lethal insecticide exposure. *Current Research in Insect Science*, 1 : 100018. *Cité page 188*
- Ingham, V. A., Tennessen, J. A., Lucas, E. R., Elg, S., Yates, H. C., Carson, J., Guelbeogo, W. M., Sagnon, N., Hughes, G. L., Heinz, E., Neafsey, D. E., and Ranson, H. 2021b. Integration of whole genome sequencing and transcriptomics reveals a complex picture of the reestablishment of insecticide resistance in the major malaria vector *Anopheles coluzzii*. *PLOS Genetics*, 17(12) : e1009970. *Cité page 178*
- Itokawa, K., Komagata, O., Kasai, S., Ogawa, K., and Tomita, T. 2016. Testing the causality between CYP9M10 and pyrethroid resistance using the TALEN and CRISPR/Cas9 technologies. *Scientific Reports*, 6(January) : 1–10. *Cité page 65*

- Iturriza-Gomara, M., Green, J., Brown, D. W., Desselberger, U., and Gray, J. J. 1999. Comparison of specific and random priming in the reverse transcriptase polymerase chain reaction for genotyping group A rotaviruses. *Journal of virological methods*, 78(1-2) : 93–103. *Cité page 159*
- Jaenson, T. G. T. 1986. Sex ratio distortion and reduced lifespan of *Glossina pallidipes* infected with the virus causing salivary gland hyperplasia. *Entomologia Experimentalis et Applicata*, 41(3) : 265–271. *Cité page 142*
- Jamroz, R. C., Guerrero, F. D., Kammlah, D. M., and Kunz, S. E. 1998. Role of the *kdr* and super-*kdr* sodium channel mutations in pyrethroid resistance : correlation of allelic frequency to resistance level in wild and laboratory populations of horn flies (*Haematobia irritans*). *Insect Biochemistry and Molecular Biology*, 28(12) : 1031–1037. *Cité pages 54, 186*
- Jaramillo-O., N., Fonseca-González, I., and Chaverra-Rodríguez, D. 2014. Geometric Morphometrics of Nine Field Isolates of *Aedes aegypti* with Different Resistance Levels to Lambda-Cyhalothrin and Relative Fitness of One Artificially Selected for Resistance. *PLOS ONE*, 9(5) : 1–15. *Cité page 58*
- Johnson, C. G. 1940. Development, hatching and mortality of the eggs of *Cimex lectularius* L. (Hemiptera) in relation to climate, with observations on the effects of preconditioning to temperature. *Parasitology*, 32(2) : 127–173. *Cité page 34*
- Johnson, M. and Hill, A. 1948. Partial resistance of a strain of bedbugs to DDT residuals. *Medical News Letter*, 12(1) : 26–28. *Cité page 43*
- Joußen, N., Heckel, D. G., Haas, M., Schuphan, I., and Schmidt, B. 2008. Metabolism of imidacloprid and DDT by P450 CYP6G1 expressed in cell cultures of *Nicotiana tabacum* suggests detoxification of these insecticides in Cyp6g1-overexpressing strains of *Drosophila melanogaster*, leading to resistance. *Pest Management Science*, 64(1) : 65–73. *Cité page 56*
- Kaiwa, N., Hosokawa, T., Kikuchi, Y., Nikoh, N., Meng Xian, Y., Kimura, N., Ito, M., and Fukatsu, T. 2010. Primary Gut Symbiont and Secondary, Sodalised-Allied Symbiont of the Scutellerid Stinkbug *Cantao ocellatus*. *Applied and Environmental Microbiology*, 76(11) : 3486–3494. *Cité page 156*
- Kampango, A., Hocke, E. F., Hansson, H., Furu, P., Haji, K. A., David, J.-P., Konradsen, F., Saleh, F., Weldon, C. W., Schiøler, K. L., and Alifrangis, M. 2022. High DDT resistance without apparent association to *kdr* and Glutathione-S-transferase (GST) gene mutations in *Aedes aegypti* population at hotel compounds in Zanzibar. *PLOS Neglected Tropical Diseases*, 16(5) : e0010355. *Cité page 186*
- Kämpfer, P., Martin, E., Lodders, N., Jäckel, U., Huber, B. E., Schumann, P., Langer, S., Busse, H. J., and Scholz, H. 2010. *Paenochrobactrum gallinarii* gen. nov., sp. nov., isolated from air of a duck barn, and reclassification of *Pseudochrobactrum glaciei* as *Paenochrobactrum glaciei* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 60(7) : 1493–1498. *Cité page 146*
- Karasov, T., Messer, P. W., and Petrov, D. A. 2010. Evidence that Adaptation in *Drosophila* Is Not Limited by Mutation at Single Sites. *PLOS Genetics*, 6(6) : 1–10. *Cité page 197*
- Kay, T., Helleu, Q., and Keller, L. 2022. Iterative evolution of supergene-based social polymorphism in ants. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 377(1856). *Cité page 68*

- Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., and Fukatsu, T. 2012. Symbiont-mediated insecticide resistance. *Proceedings of the National Academy of Sciences*, 109(22) : 8618–8622. *Cité page 54*
- Kinsley, C. 1893. *The Circle of Useful Knowledge; for the use of Farmers, Mechanics, Merchants, Manufacturers, Surveyors, Housekeepers, Professional Men, etc.* Kinsley Publishing. *Cité page 42*
- Kliot, A. and Ghanim, M. 2012. Fitness costs associated with insecticide resistance. *Pest Management Science*, 68(11) : 1431–1437. *Cité page 58*
- Koga, R. and Moran, N. A. 2014. Swapping symbionts in spittlebugs : evolutionary replacement of a reduced genome symbiont. *The ISME journal*, 8(6) : 1237–1246. *Cité page 156*
- Koga, R., Moriyama, M., Onodera-Tanifuji, N., Ishii, Y., Takai, H., Mizutani, M., Oguchi, K., Okura, R., Suzuki, S., Gotoh, Y., Hayashi, T., Seki, M., Suzuki, Y., Nishide, Y., Hosokawa, T., Wakamoto, Y., Furusawa, C., and Fukatsu, T. 2022. Single mutation makes *Escherichia coli* an insect mutualist. *Nature Microbiology*, 7(8) : 1141–1150. *Cité page 155*
- Koganemaru, R., Miller, D. M., and Adelman, Z. N. 2013. Robust cuticular penetration resistance in the common bed bug (*Cimex lectularius* L.) correlates with increased steady-state transcript levels of CPR-type cuticle protein genes. *Pesticide Biochemistry and Physiology*, 106(3) : 190–197. *Cité pages 50, 189*
- Kokwaro, E. D. 2006. Virus particles in male accessory reproductive glands of tsetse, *Glossina morsitans morsitans* (Diptera : Glossinidae) and associated tissue changes. *International Journal of Tropical Insect Science*, 26(4) : 266–272. *Cité page 142*
- Kondrashov, F. A. 2012. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B : Biological Sciences*, 279(1749) : 5048–5057. *Cité pages 65, 188*
- Kumar, A., Srivastava, P., Sirisena, P. D., Dubey, S. K., Kumar, R., Shrinet, J., and Sunil, S. 2018. Mosquito innate immunity. *Insects*, 9(3). *Cité page 187*
- Labbé, P., Berthomieu, A., Berticat, C., Alout, H., Raymond, M., Lenormand, T., and Weill, M. 2007. Independent duplications of the acetylcholinesterase gene conferring insecticide resistance in the mosquito *Culex pipiens*. *Molecular Biology and Evolution*, 24(4) : 1056–1067. *Cité page 65*
- Labbé, P., Sidos, N., Raymond, M., and Lenormand, T. 2009. Resistance Gene Replacement in the Mosquito *Culex pipiens* : Fitness Estimation From Long-Term Cline Series. *Genetics*, 182(1) : 303–312. *Cité page 60*
- Labbé, P., Milesi, P., Yébakima, A., Pasteur, N., Weill, M., and Lenormand, T. 2014. Gene-dosage effects on fitness in recent adaptative duplications : *ace-1* in the mosquito *Culex pipiens*. *Evolution*, 68(7) : 2092–2101. *Cité pages 58, 60*
- Laffargue, S., Leininger, A., and Dabaghi, N. 2021. Les Français face aux punaises de lit. Technical report, IPSOS. *Cité page 29*
- Laidoudi, Y., Lévasseur, A., Medkour, H., Maaloum, M., Ben Khedher, M., Sambou, M., Bassene, H., Davoust, B., Fenollar, F., Raoult, D., and Mediannikov, O. 2020. An Earliest Endosymbiont, *Wolbachia massiliensis* sp. nov., Strain PL13 from the Bed Bug (*Cimex hemipterus*), Type Strain of a New Supergroup T. *International journal of molecular sciences*, 21(21). *Cité page 147*

- Langrand, J., Boucaud-Maitre, D., Creusat, G., Delcourt, N., Jegou, F., Kammerer, M., Marcotullio, E., Nisse, P., Penouil Pucheu, F., Pineau, X., and Saponi, J.-M. 2018. Expositions à des produits phytopharmaceutiques à base de substances actives non autorisées. Technical report, ANSES. *Cité page 43*
- Le Corre, V. and Kremer, A. 1998. Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, 11(4) : 495–512. *Cité page 31*
- Le Goff, G. and Hilliou, F. 2017. Resistance evolution in *Drosophila* : the case of CYP6G1. *Pest Management Science*, 73(3) : 493–499. *Cité page 56*
- Le Goff, G., Boundy, S., Daborn, P. J., Yen, J. L., Sofer, L., Lind, R., Sabourault, C., Madi-Ravazzi, L., and Ffrench-Constant, R. H. 2003. Microarray analysis of cytochrome P450 mediated insecticide resistance in *Drosophila*. *Insect Biochemistry and Molecular Biology*, 33(7) : 701–708. *Cité page 187*
- Le Monde 2022. Les néonicotinoïdes à nouveau autorisés temporairement dans les champs de betteraves sucrières. *Cité page 43*
- Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V., and Joron, M. 2014. Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature Communications*, 5(1) : 5644. *Cité page 68*
- Lenormand, T., Guillemaud, T., Bourguet, D., and Raymond, M. 1998. Appearance and sweep of a gene duplication : Adaptive response and potential for new functions in the mosquito *Culex pipiens*. *Evolution*, 52(6) : 1705–1712. *Cité page 60*
- Levy, S. M., Falleiros, M. F., Moscardi, F., and Gregório, E. A. 2011. The role of peritrophic membrane in the resistance of *Anticarsia gemmatalis* larvae (Lepidoptera : Noctuidae) during the infection by its nucleopolyhedrovirus (AgMNPV). *Arthropod structure & development*, 40 5 : 429–434. *Cité page 198*
- Lewontin, R. C. and Krakauer, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1) : 175–195. *Cité page 66*
- Li, C.-x., Guo, X.-x., Zhang, Y.-m., Dong, Y.-d., Xing, D., Yan, T., Wang, G., Zhang, H.-d., and Zhao, T.-y. 2016. Identification of genes involved in pyrethroid-, propoxur-, and dichlorvos- insecticides resistance in the mosquitoes, *Culex pipiens* complex (Diptera : Culicidae). *Acta Tropica*, 157 : 84–95. *Cité page 187*
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. 2015. MEGAHIT : an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford, England)*, 31(10) : 1674–1676. *Cité page 145*
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16) : 2078–2079. *Cité page 145*
- Li, X., Ma, L., Sun, L., and Zhu, C. 2002. Biotic characteristics in the deltamethrin-susceptible and resistant strains of *Culex pipiens pallens* (Diptera : Culicidae) in China. *Applied Entomology and Zoology*, 37(2) : 305–308. *Cité page 58*
- Liang, Z.-K., Pang, R., Dong, Y., Sun, Z.-X., Ling, Y., and Zhang, W.-Q. 2018. Identification of SNPs involved in regulating a novel alternative transcript of P450 CYP6ER1 in the brown planthopper. *Insect Science*, 25(5) : 726–738. *Cité pages 188, 189*

- Lilly, D. G., Latham, S. L., Webb, C. E., and Doggett, S. L. 2016a. Cuticle thickening in a pyrethroid-resistant strain of the common bed bug, *Cimex lectularius* L. (hemiptera : Cimicidae). *PLoS ONE*, 11(4) : 6–16. *Cité pages 50, 189*
- Lilly, D. G., Dang, K., Webb, C. E., and Doggett, S. L. 2016b. Evidence for metabolic pyrethroid resistance in the common bed bug (Hemiptera : Cimicidae). *Journal of Economic Entomology*, 109(3) : 1364–1368. *Cité page 52*
- Lin, M., Whitmire, S., Chen, J., Farrel, A., Shi, X., and Guo, J.-T. 2017. Effects of short indels on protein structure and function in human genomes. *Scientific reports*, 7(1) : 9313. *Cité page 62*
- Liu, N., Wang, Y., Li, T., and Feng, X. 2021. G-protein coupled receptors (Gpcrs) : Signaling pathways, characterization, and functions in insect physiology and toxicology. *International Journal of Molecular Sciences*, 22(10) : 14–18. *Cité page 54*
- Lopez, O. and Fernandez-Bolanos, J. 2011. Pyrethrins and Pyrethroid Insecticides. In *Green Trends in Insect Control*, ISSN, chapter 3, pages 99–124. Royal Society of Chemistry. *Cité page 45*
- Lucas, E. R., Miles, A., Harding, N. J., Clarkson, C. S., Lawniczak, M. K., Kwiatkowski, D. P., Weetman, D., and Donnelly, M. J. 2019. Whole-genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes. *Genome Research*, 29(8) : 1250–1261. *Cité page 65*
- Lynd, A., Weetman, D., Barbosa, S., Egyir Yawson, A., Mitchell, S., Pinto, J., Hastings, I., and Donnelly, M. J. 2010. Field, Genetic, and Modeling Approaches Show Strong Positive Selection Acting upon an Insecticide Resistance Mutation in *Anopheles gambiae* s.s. *Molecular Biology and Evolution*, 27(5) : 1117–1125. *Cité pages 67, 186*
- Mallis, A. 1954. *Handbook of Pest Control*. Mac Nair-Dorland Compagny, New York. *Cité page 43*
- Mamidala, P., Jones, S. C., and Mittapalli, O. 2011. Metabolic resistance in bed bugs. *Insects*, 2(1) : 36–48. *Cité pages 52, 57*
- Mamidala, P., Wijeratne, A. J., Wijeratne, S., Kornacker, K., Sudhamalla, B., Rivera-Vega, L. J., Hoelmer, A., Meulia, T., Jones, S. C., and Mittapalli, O. 2012. RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug. *BMC Genomics*, 13(1). *Cité pages 50, 189*
- Martins, A. J., Ribeiro, C. D. e. M., Bellinato, D. F., Peixoto, A. A., Valle, D., and Lima, J. B. P. 2012. Effect of Insecticide Resistance on Development, Longevity and Reproduction of Field or Laboratory Selected *Aedes aegypti* Populations. *PLOS ONE*, 7(3) : 1–9. *Cité page 58*
- Matoke-Muhia, D., Gimnig, J. E., Kamau, L., Shililu, J., Bayoh, M. N., and Walker, E. D. 2016. Decline in frequency of the 2La chromosomal inversion in *Anopheles gambiae* (s.s.) in Western Kenya : correlation with increase in ownership of insecticide-treated bed nets. *Parasites & Vectors*, 9(1) : 334. *Cité page 178*
- Matsuo, N. 2019. Discovery and development of pyrethroid insecticides. *Proceedings of the Japan Academy Series B : Physical and Biological Sciences*, 95(7) : 378–400. *Cité page 41*

- Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York. *Cité page 57*
- Mazzarri, M. B. and Georghiou, G. P. 1995. Characterization of resistance to organophosphate, carbamate, and pyrethroid insecticides in field populations of *Aedes aegypti* from Venezuela. *Journal of the American Mosquito Control Association*, 11(3) : 315–322. *Cité page 47*
- McCoy, R. C. and Akey, J. M. 2017. Selection plays the hand it was dealt : Evidence that human adaptation commonly targets standing genetic variation. *Genome Biology*, 18(1) : 1–5. *Cité page 68*
- McDonald, J. H. and Kreitman, M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, 351(6328) : 652–654. *Cité page 66*
- Meriweather, M., Matthews, S., Rio, R., and Baucom, R. S. 2013. A 454 Survey Reveals the Community Composition and Core Microbiome of the Common Bed Bug (*Cimex lectularius*) across an Urban Landscape. *PLoS ONE*, 8(4). *Cité pages 36, 148, 155*
- Moiroux, N., Gomez, M. B., Penetier, C., Elanga, E., Djènontin, A., Chandre, F., Djègbé, I., Guis, H., and Corbel, V. 2012. Changes in *Anopheles funestus* biting behavior following universal coverage of long-lasting insecticidal nets in Benin. *The Journal of infectious diseases*, 206(10) : 1622–1629. *Cité page 49*
- Montes, C., Cuadrillero, C., and Vilella, D. 2009. Maintenance of a Laboratory Colony of *Cimex lectularius* (Hemiptera : Cimicidae) Using an Artificial Feeding Technique. *Journal of Medical Entomology*, 39(4) : 675–679. *Cité page 38*
- Mullaney, J. M., Mills, R. E., Pittard, W. S., and Devine, S. E. 2010. Small insertions and deletions (INDELS) in human genomes. *Human Molecular Genetics*, 19(R2) : R131–R136. *Cité page 63*
- Musiba, R. M., Tarimo, B. B., Monroe, A., Msaky, D., Ngowo, H., Mihayo, K., Limwagu, A., Chilla, G. T., Shubis, G. K., Ibrahim, A., Greer, G., Mcha, J. H., Haji, K. A., Abbas, F. B., Ali, A., Okumu, F. O., and Kiware, S. S. 2022. Outdoor biting and pyrethroid resistance as potential drivers of persistent malaria transmission in Zanzibar. *Malaria Journal*, 21(1) : 172. *Cité page 49*
- Naushad, H. S., Lee, B., and Gupta, R. S. 2014. Conserved signature indels and signature proteins as novel tools for understanding microbial phylogeny and systematics : identification of molecular signatures that are specific for the phytopathogenic genera *Dickeya*, *Pectobacterium* and *Brenneria*. *International Journal of Systematic and Evolutionary Microbiology*, 64(Pt_2) : 366–383. *Cité page 157*
- Naylor, R., Balvín, O., Delaunay, P., and Akhondi, M. 2018. The Bed Bug Resurgence in Europe and Russia. In *Advances in the Biology and Management of Modern Bed Bugs*, pages 59–68. John Wiley & Sons Ltd. *Cité page 29*
- Ndenga, B. A., Mulaya, N. L., Musaki, S. K., Shiroko, J. N., Dongus, S., and Fillinger, U. 2016. Malaria vectors and their blood-meal sources in an area of high bed net ownership in the western Kenya highlands. *Malaria Journal*, 15(1) : 76. *Cité page 49*
- North, H. L., Caminade, P., Severac, D., Belkhir, K., and Smadja, C. M. 2020. The role of copy-number variation in the reinforcement of sexual isolation between the two European subspecies of the house mouse : CNV and reinforcement in the house mouse. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 375(1806). *Cité page 64*

- Palazzo, A. F. and Gregory, T. R. 2014. The Case for Junk DNA. *PLOS Genetics*, 10(5) : e1004351. *Cité page 192*
- Palazzo, A. F. and Lee, E. S. 2015. Non-coding RNA : what is functional and what is junk? *Frontiers in genetics*, 6 : 2. *Cité page 192*
- Pan, C., Zhou, Y., and Mo, J. 2009. The clone of laccase gene and its potential function in cuticular penetration resistance of *Culex pipiens pallens* to fenvalerate. *Pesticide Biochemistry and Physiology*, 93(3) : 105–111. *Cité pages 50, 189*
- Panagiotakopulu, E. and Buckland, P. C. 1999. *Cimex lectularius* L., the common bed bug from pharaonic Egypt. *Antiquity*, 73(282) : 908–911. *Cité page 28*
- Pang, R., Li, Y., Dong, Y., Liang, Z., Zhang, Y., and Zhang, W. 2014. Identification of promoter polymorphisms in the cytochrome P450 CYP6AY1 linked with insecticide resistance in the brown planthopper, *Nilaparvata lugens*. *Insect Molecular Biology*, 23(6) : 768–778. *Cité page 52*
- Paulke-Korinek, M., Széll, M., Laferl, H., Auer, H., and Wenisch, C. 2012. Bed bugs can cause severe anaemia in adults. *Parasitology Research*, 110(6) : 2577–2579. *Cité page 32*
- Pearson, W. R. 2013. An introduction to sequence similarity ("homology") searching. *Current protocols in bioinformatics*, Chapter 3 : 1–3. *Cité page 146*
- Peng, T., Pan, Y., Tian, F., Xu, H., Yang, F., Chen, X., Gao, X., Li, J., Wang, H., and Shang, Q. 2021. Identification and the potential roles of long non-coding RNAs in regulating acetyl-CoA carboxylase ACC transcription in spirotetramat-resistant *Aphis gossypii*. *Pesticide Biochemistry and Physiology*, 179 : 104972. *Cité page 193*
- Pettersson, M. E., Rochus, C. M., Han, F., Chen, J., Hill, J., Wallerman, O., Fan, G., Hong, X., Xu, Q., Zhang, H., Liu, S., Liu, X., Haggerty, L., Hunt, T., Martin, F. J., Flicek, P., Bunikis, I., Folkvord, A., and Andersson, L. 2019. A chromosome-level assembly of the Atlantic herring genome-detection of a supergene and other signals of selection. *Genome Research*, 29(11) : 1919–1928. *Cité page 181*
- Pfiester, M., Koehler, P. G., and Pereira, R. M. 2008. Ability of Bed Bug-Detecting Canines to Locate Live Bed Bugs and Viable Bed Bug Eggs. *Journal of Economic Entomology*, 101(4) : 1389–1396. *Cité page 39*
- Pietri, J. E. and Liang, D. 2018. Virulence of entomopathogenic bacteria in the bed bug, *Cimex lectularius*. *Journal of Invertebrate Pathology*, 151(October) : 1–6. *Cité page 41*
- Polanco, A. M., Brewster, C. C., and Miller, D. M. 2011a. Population growth potential of the bed bug, *Cimex lectularius* L. : A life table analysis. *Insects*, 2(2) : 173–185. *Cité page 197*
- Polanco, A. M., Miller, D. M., and Brewster, C. C. 2011b. Survivorship During Starvation for *Cimex lectularius* L. *Insects*, 2(2) : 232–242. *Cité page 34*
- Potter, M. F. 2011. The history of bed bug management- With lessons from the past. *American Entomologist*, 57(1) : 14–25. *Cité pages 28, 42*
- Potts, R., Molina, I., Sheele, J. M., and Pietri, J. E. 2020. Molecular detection of *Rickettsia* infection in field-collected bed bugs. *New Microbes and New Infections*, 34 : 100646. *Cité pages 36, 148*

- Puig, M., Cáceres, M., and Ruiz, A. 2004. Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24) : 9013–9018. *Cité page 188*
- Quiroz-Castañeda, R. E., Mendoza-Mejía, A., Obregón-Barboza, V., Martínez-Ocampo, F., Hernández-Mendoza, A., Martínez-Garduño, F., Guillén-Solís, G., Sánchez-Rodríguez, F., Peña-Chora, G., Ortíz-Hernández, L., Gaytán-Colín, P., and Dantán-González, E. 2015. Identification of a New *Alcaligenes faecalis* Strain MOR02 and Assessment of Its Toxicity and Pathogenicity to Insects. *BioMed Research International*, 2015. *Cité page 147*
- Ranson, H., Claudianos, C., Ortell, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H., and Feyereisen, R. 2002. Evolution of Supergene Families Associated with Insecticide Resistance. *Science*, 298(5591) : 179–181. *Cité pages 178, 179*
- Ranson, H., Burhani, J., Lumjuan, N., and Black, W. C. I. 2010. Insecticide resistance in dengue vectors - Review. *TropIKA.net*, pages 1–12. *Cité page 178*
- Raymond, M., Berticat, C., Weill, M., Pasteur, N., and Chevillon, C. 2001. Insecticide resistance in the mosquito *Culex pipiens* : what have we learned about adaptation? *Genetica*, 112-113 : 287–296. *Cité page 60*
- Richly, E., Kurth, J., and Leister, D. 2002. Mode of Amplification and Reorganization of Resistance Genes During Recent *Arabidopsis thaliana* Evolution. *Molecular Biology and Evolution*, 19(1) : 76–84. *Cité page 179*
- Riehle, M. M., Bukhari, T., Gneme, A., Guelbeogo, W. M., Coulibaly, B., Fofana, A., Pain, A., Bischoff, E., Renaud, F., Beavogui, A. H., Traore, S. F., Sagnon, N., and Vernick, K. D. 2017. The *Anopheles gambiae* 2La chromosome inversion is associated with susceptibility to *Plasmodium falciparum* in Africa. *eLife*, 6 : e25813. *Cité page 178*
- Riveron, J. M., Yunta, C., Ibrahim, S. S., Djouaka, R., Irving, H., Menze, B. D., Ismail, H. M., Hemingway, J., Ranson, H., Albert, A., and Wondji, C. S. 2014. A single mutation in the *GSTe2* gene allows tracking of metabolically based insecticide resistance in a major malaria vector. *Genome Biology*, 15(2) : 1–20. *Cité page 187*
- Romero, A., Potter, M. F., Potter, D. A., and Haynes, K. F. 2007. Insecticide Resistance in the Bed Bug : A Factor in the Pest's Sudden Resurgence? *Journal of Medical Entomology*, 44(2) : 175–178. *Cité pages 29, 46, 185*
- Romero, A., Potter, M. F., and Haynes, K. F. 2009a. Behavioral responses of the bed bug to insecticide residues. *Journal of Medical Entomology*, 46(1) : 51–57. *Cité page 49*
- Romero, A., Potter, M. F., and Haynes, K. F. 2009b. Evaluation of Piperonyl Butoxide as a Deltamethrin Synergist for Pyrethroid-Resistant Bed Bugs. *Journal of Economic Entomology*, 102(6) : 2310–2315. *Cité pages 52, 176*
- Rosenfeld, J. A., Reeves, D., Brugler, M. R., Narechania, A., Simon, S., Durrett, R., Foox, J., Shianna, K., Schatz, M. C., Gandara, J., Afshinnekoo, E., Lam, E. T., Hastie, A. R., Chan, S., Cao, H., Saghbini, M., Kentsis, A., Planet, P. J., Kholodovych, V., Tessler, M., Baker, R., DeSalle, R., Sorkin, L. N., Kolokotronis, S. O., Siddall, M. E., Amato, G., and Mason, C. E. 2016. Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nature Communications*, 7. *Cité page 36*

- Roth, S., Balvín, O., Siva-Jothy, M. T., Di Iorio, O., Benda, P., Calva, O., Faundez, E. I., Anwarali Khan, F. A., McFadzen, M., Lehnert, M. P., Naylor, R., Simov, N., Morrow, E. H., Willassen, E., and Reinhardt, K. 2019. Bedbugs Evolved before Their Bat Hosts and Did Not Co-speciate with Ancient Humans. *Current Biology*, 29(11) : 1847–1853. *Cité pages 26, 27, 28*
- Roush, R. T. and McKenzie, J. A. 1987. Ecological genetics of insecticide and acaricide resistance. *Annual review of entomology*, 32(1) : 361–380. *Cité page 57*
- Ruaux, N. 2013. Les résistances aux insecticides, antiparasitaires, antibiotiques... *anses-01568720*. *Cité page 42*
- Saavedra-Rodriguez, K., Urdaneta-Marquez, L., Rajatileka, S., Moulton, M., Flores, A. E., Fernandez-Salas, I., Bisset, J., Rodriguez, M., McCall, P. J., Donnelly, M. J., Ranson, H., Hemingway, J., and Black IV, W. C. 2007. A mutation in the voltage-gated sodium channel gene associated with pyrethroid resistance in Latin American *Aedes aegypti*. *Insect Molecular Biology*, 16(6) : 785–798. *Cité pages 54, 181, 186*
- Sadilek, D., Štáhlavský, F., Vilímová, J., and Zima, J. 2013. Extensive fragmentation of the X chromosome in the bed bug *Cimex lectularius* Linnaeus, 1758 (Heteroptera, Cimicidae) : A survey across Europe. *Comparative Cytogenetics*, 7(4) : 253–269. *Cité page 37*
- Saenz, V. L., Booth, W., Schal, C., and Vargo, E. L. 2012. Genetic Analysis of Bed Bug Populations Reveals Small Propagule Size Within Individual Infestations but High Genetic Diversity Across Infestations From the Eastern United States. *Journal of Medical Entomology*, 49(4) : 865–875. *Cité pages 31, 196*
- Sassera, D., Epis, S., Pajoro, M., and Bandi, C. 2013. Microbial symbiosis and the control of vector-borne pathogens in tsetse flies, human lice, and triatomine bugs. *Pathogens and Global Health*, 107(6) : 285–292. *Cité page 156*
- Sato, Y., Jang, S., Takeshita, K., Itoh, H., Koike, H., Tago, K., Hayatsu, M., Hori, T., and Kikuchi, Y. 2021. Insecticide resistance by a host-symbiont reciprocal detoxification. *Nature Communications*, 12(1) : 1–8. *Cité page 55*
- Schlötterer, C., Tobler, R., Kofler, R., and Nolte, V. 2014. Sequencing pools of individuals-mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, 15(11) : 749–763. *Cité page 62*
- Schmidt, J. M., Good, R. T., Appleton, B., Sherrard, J., Raymant, G. C., Bogwitz, M. R., Martin, J., Daborn, P. J., Goddard, M. E., Batterham, P., and Robin, C. 2010. Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. *PLoS Genetics*, 6(6) : 1–11. *Cité page 65*
- Schwander, T., Libbrecht, R., and Keller, L. 2014. Supergenes and Complex Phenotypes. *Current Biology*, 24(7) : R288–R294. *Cité pages 12, 68, 69, 177*
- Seal, R. L., Chen, L.-L., Griffiths-Jones, S., Lowe, T. M., Mathews, M. B., O'Reilly, D., Pierce, A. J., Stadler, P. F., Ulitsky, I., Wolin, S. L., and Bruford, E. A. 2020. A guide to naming human non-coding RNA genes. *The EMBO Journal*, 39(6) : e103777. *Cité page 192*
- Seong, K. M., Lee, D.-Y., Yoon, K. S., Kwon, D. H., Kim, H. C., Klein, T. A., Clark, J. M., and Lee, S. H. 2010. Establishment of Quantitative Sequencing and Filter Contact Vial Bioassay for Monitoring Pyrethroid Resistance in the Common Bed Bug, *Cimex lectularius*. *Journal of Medical Entomology*, 47(4) : 592–599. *Cité page 47*

- Shao, H., Ganesamoorthy, D., Duarte, T., Cao, M. D., Hoggart, C. J., and Coin, L. J. M. 2018. npInv : accurate detection and genotyping of inversions using long read sub-alignment. *BMC Bioinformatics*, 19(1) : 261. *Cité page 180*
- Shatkin, A. J., Sipe, J. D., and Loh, P. 1968. Separation of ten reovirus genome segments by polyacrylamide gel electrophoresis. *Journal of Virology*, 2(10) : 986–991. *Cité page 151*
- Sheele, J. M., Pritt, B. S., Libertin, C. R., and Wysokinska, E. M. 2021. Bed bugs are associated with anemia. *The American Journal of Emergency Medicine*, 46 : 482–488. *Cité page 32*
- Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., and Zhang, Y.-Z. 2016. Redefining the invertebrate RNA virosphere. *Nature*, 540(7634) : 539–543. *Cité page 158*
- Silva Martins, W. F., Silva Pereira, B. N., Vieira Alves, A. T., Murphy, A., Silva Martins, P. G., Weetman, D., Wilding, C. S., and Donnelly, M. J. 2019. Development and application of a tri-allelic PCR assay for screening Vgsc-L1014F kdr mutations associated with pyrethroid and organochlorine resistance in the mosquito *Culex quinquefasciatus*. *Parasites & Vectors*, 12(1) : 232. *Cité pages 186, 187*
- Simmons, N. B., Seymour, K. L., Habersetzer, J., and Gunnell, G. F. 2008. Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature*, 451(7180) : 818–821. *Cité page 26*
- Singh, N., Wang, C., and Cooper, R. 2013. Effect of Trap Design, Chemical Lure, Carbon Dioxide Release Rate, and Source of Carbon Dioxide on Efficacy of Bed Bug Monitors. *Journal of Economic Entomology*, 106(4) : 1802–1811. *Cité page 40*
- Smith, J. M. and Haigh, J. 1974. The hitch-hiking effect of a favourable gene. *Genetics Research*, 23(1) : 23–35. *Cité pages 67, 176*
- Soderlund, D. M. 2012. Molecular mechanisms of pyrethroid insecticide neurotoxicity : Recent advances. *Archives of Toxicology*, 86(2) : 165–181. *Cité pages 45, 46*
- Soh, L. S. and Veera Singham, G. 2022. Bacterial symbionts influence host susceptibility to fenitrothion and imidacloprid in the obligate hematophagous bed bug, *Cimex hemipterus*. *Scientific Reports*, 12(1) : 1–16. *Cité page 56*
- Southall, J. 1730. *A treatise of bugs*. Nineteenth Century Collections Online : Science, Technology, and Medicine : 1780-1925, Part II. J. Roberts, near the Oxford-Arms in Warwick-Lane. *Cité page 28*
- Stanojević, M., Li, K., Stamenković, G., Ilić, B., Paunović, M., Pešić, B., Maslovara, I., Šiljić, M., Ćirković, V., and Zhang, Y. 2020. Depicting the RNA Virome of Hematophagous Arthropods from Belgrade, Serbia. *Viruses*, 12(9) : 1–11. *Cité pages 36, 199*
- Steinegger, M. and Söding, J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11) : 1026–1028. *Cité page 145*
- Stutt, A. D. and Siva-Jothy, M. T. 2001. Traumatic insemination and sexual conflict in the bed bug *Cimex lectularius*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10) : 5683–5687. *Cité page 34*

- Tan, L., Zhang, J., Li, Y., Li, Y., Jiang, H., Cao, X., and Hu, Y. 2008. The complete nucleotide sequence of the type 5 *Helicoverpa armigera* cytoplasmic polyhedrosis virus genome. *Virus Genes*, 36(3) : 587–593. *Cité page 158*
- Tanada, Y. and Omi, E. M. 1974. Epizootiology of virus diseases in three lepidopterous insect species of alfalfa. *Population Ecology*, 16(1) : 59–68. *Cité page 142*
- Tang, H., Kambris, Z., Lemaitre, B., and Hashimoto, C. 2008. A serpin that regulates immune melanization in the respiratory system of *Drosophila*. *Developmental cell*, 15(4) : 617–626. *Cité page 187*
- Tawatsin, A., Thavara, U., Chompoonsri, J., Phusup, Y., Jonjang, N., Khumsawads, C., Bhakdeenuan, P., Sawanpanyalert, P., Asavadachanukorn, P., Mulla, M. S., Siritiyasatien, P., and Debboun, M. 2011. Insecticide Resistance in Bedbugs in Thailand and Laboratory Evaluation of Insecticides for the Control of *Cimex hemipterus* and *Cimex lectularius* (Hemiptera : Cimicidae). *Journal of Medical Entomology*, 48(5) : 1023–1030. *Cité page 197*
- Tchouakui, M., Riveron Miranda, J., Mugenzi, L. M. J., Djonabaye, D., Wondji, M. J., Tchoupo, M., Tchappa, W., Njiokou, F., and Wondji, C. S. 2020. Cytochrome P450 metabolic resistance (CYP6P9a) to pyrethroids imposes a fitness cost in the major African malaria vector *Anopheles funestus*. *Heredity*, 124(5) : 621–632. *Cité page 59*
- Thomas, G. W., Dohmen, E., Hughes, D. S., Murali, S. C., Poelchau, M., Glastad, K., Anstead, C. A., Ayoub, N. A., Batterham, P., Bellair, M., Binford, G. J., Chao, H., Chen, Y. H., Childers, C., Dinh, H., Doddapaneni, H. V., Duan, J. J., Dugan, S., Esposito, L. A., Friedrich, M., Garb, J., Gasser, R. B., Goodisman, M. A., Gundersen-Rindal, D. E., Han, Y., Handler, A. M., Hatakeyama, M., Hering, L., Hunter, W. B., Ioannidis, P., Jayaseelan, J. C., Kalra, D., Khila, A., Korhonen, P. K., Lee, C. E., Lee, S. L., Li, Y., Lindsey, A. R., Mayer, G., McGregor, A. P., McKenna, D. D., Misof, B., Munidasa, M., Munoz-Torres, M., Muzny, D. M., Niehuis, O., Osuji-Lacy, N., Palli, S. R., Panfilio, K. A., Pechmann, M., Perry, T., Peters, R. S., Poynton, H. C., Prpic, N. M., Qu, J., Rotenberg, D., Schal, C., Schoville, S. D., Scully, E. D., Skinner, E., Sloan, D. B., Stouthamer, R., Strand, M. R., Szucsich, N. U., Wijeratne, A., Young, N. D., Zattara, E. E., Benoit, J. B., Zdobnov, E. M., Pfrender, M. E., Hackett, K. J., Werren, J. H., Worley, K. C., Gibbs, R. A., Chipman, A. D., Waterhouse, R. M., Bornberg-Bauer, E., Hahn, M. W., and Richards, S. 2020. Gene content evolution in the arthropods. *Genome Biology*, 21(1) : 1–14. *Cité pages 36, 62*
- Thongprem, P., Evison, S. E., Hurst, G. D., and Otti, O. 2020. Transmission, Tropism, and Biological Impacts of *Torix rickettsia* in the Common Bed Bug *Cimex lectularius* (Hemiptera : Cimicidae). *Frontiers in Microbiology*, 11(December) : 1–16. *Cité page 36*
- Toh, H., Weiss, B. L., Perkin, S. A. H., Yamashita, A., Oshima, K., Hattori, M., and Aksoy, S. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome research*, 16(2) : 149–156. *Cité page 156*
- Trdan, S. 2016. *Insecticides Resistance*. IntechOpen. *Cité page 58*
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., and Posthuma, D. 2021. Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1) : 59. *Cité page 61*
- Usinger, R. L. 1966. *Monograph of Cimicidae*, volume VII. ESA. *Cité pages 25, 28*

- Vainio, E. J., Chiba, S., Ghabrial, S. A., Maiss, E., Roossinck, M., Sabanadzovic, S., Suzuki, N., Xie, J., Nibert, M., and Consortium, I. R. 2018. ICTV Virus Taxonomy Profile : Partitiviridae. *Journal of General Virology*, 99(1) : 17–18. *Cité page 158*
- Van Leeuwen, T., Dermauw, W., Mavridis, K., and Vontas, J. 2020. Significance and interpretation of molecular diagnostics for insecticide resistance management of agricultural pests. *Current opinion in insect science*, 39 : 69–76. *Cité page 195*
- Vaughan, A., Hawkes, N., and Hemingway, J. 1997. Co-amplification explains linkage disequilibrium of two mosquito esterase genes in insecticide-resistant *Culex quinquefasciatus*. *Biochemical Journal*, 325(2) : 359–365. *Cité page 63*
- Vignieri, S. N., Larson, J. G., and Hoekstra, H. E. 2010. The selective advantage of crypsis in mice. *Evolution ; international journal of organic evolution*, 64(7) : 2153–2158. *Cité page 178*
- Villagrán, M. E., Marín, C., Hurtado, A., Sánchez-Moreno, M., and de Diego, J. A. 2008. Natural infection and distribution of triatomines (Hemiptera : Reduviidae) in the state of Querétaro, Mexico. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102(8) : 833–838. *Cité page 32*
- Villmoare, B., Kimbel, W. H., Seyoum, C., Campisano, C. J., DiMaggio, E. N., Rowan, J., Braun, D. R., Arrowsmith, J. R., and Reed, K. E. 2015. Early Homo at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia. *Science*, 347(6228) : 1352–1355. *Cité page 28*
- Vontas, J. G., Small, G. J., and Hemingway, J. 2001. Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in *Nilaparvata lugens*. *Biochemical Journal*, 357(1) : 65–72. *Cité page 52*
- Wang, C., Singh, N., and Cooper, R. 2014. Efficacy of an essential oil-based pesticide for controlling bed bug (*Cimex lectularius*) infestations in apartment buildings. *Insects*, 5(4) : 849–859. *Cité page 41*
- Wang, K., Zhao, J. N., Bai, J. Y., Shang, Y. Z., Zhang, S. Q., Hou, Y. F., Chen, M. H., and Han, Z. J. 2021. Pyrethroid Resistance and Fitness Cost Conferred by the super-kdr Mutation M918L in *Rhopalosiphum padi* (Hemiptera : Aphididae). *Journal of Economic Entomology*, 114(4) : 1789–1795. *Cité page 176*
- Wang, X., Martínez, M.-a., Dai, M., Chen, D., Ares, I., Romero, A., Castellano, V., Martínez, M., Luis, J., Martínez-larrañaga, M.-r., Anadón, A., and Yuan, Z. 2016. Permethrin-induced oxidative stress and toxicity and metabolism . A review. *Environmental Research*, 149 : 86–104. *Cité page 55*
- Weedall, G. D., Mugenzi, L. M. J., Menze, B. D., Tchouakui, M., Ibrahim, S. S., Amvongo-Adjia, N., Irving, H., Wondji, M. J., Tchoupo, M., Djouaka, R., Riveron, J. M., and Wondji, C. S. 2019. A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. *Science translational medicine*, 11(484). *Cité page 59*
- Weetman, D., Mitchell, S. N., Wilding, C. S., Birks, D. P., Yawson, A. E., Essandoh, J., Mawejje, H. D., Djogbenou, L. S., Steen, K., Rippon, E. J., Clarkson, C. S., Field, S. G., Rigden, D. J., and Donnelly, M. J. 2015. Contemporary evolution of resistance at the major insecticide target site gene *Ace-1* by mutation and copy number variation in the malaria mosquito *Anopheles gambiae*. *Molecular ecology*, 24(11) : 2656–2672. *Cité pages 67, 187*

- Weetman, D., Wilding, C. S., Neafsey, D. E., Müller, P., Ochomo, E., Isaacs, A. T., Steen, K., Rippon, E. J., Morgan, J. C., Mawejje, H. D., Rigden, D. J., Okedi, L. M., and Donnelly, M. J. 2018a. Candidate-gene based GWAS identifies reproducible DNA markers for metabolic pyrethroid resistance from standing genetic variation in East African *Anopheles gambiae*. *Scientific Reports*, 8(1) : 2920. *Cité page 178*
- Weetman, D., Djogbenou, L. S., and Lucas, E. 2018b. Copy number variation (CNV) and insecticide resistance in mosquitoes : evolving knowledge or an evolving problem? *Current Opinion in Insect Science*, 27 : 82–88. *Cité page 65*
- Weill, M., Fort, P., Berthomieu, A., Dubois, M. P., Pasteur, N., and Raymond, M. 2002. A novel acetylcholinesterase gene in mosquitoes codes for the insecticide target and is non-homologous to the ace gene in *Drosophila*. *Hungarian Quarterly*, 49(191) : 2007–2016. *Cité page 60*
- Weill, M., Berthomieu, A., Berticat, C., Lutfalla, G., Nègre, V., Pasteur, N., Philips, A., Leonetti, J.-p., Fort, P., and Raymond, M. 2004. Insecticide resistance : a silent base. *Current Biology*, 14(14) : 552–553. *Cité page 67*
- Weiner, R. and Crow, J. F. 1951. The resistance of DDT-resistant *Drosophila* to other insecticides. *Science (New York, N.Y.)*, 113(2937) : 403–404. *Cité page 56*
- Weirauch, C., Schuh, R. T., Cassis, G., and Wheeler, W. C. 2019. Revisiting habitat and lifestyle transitions in Heteroptera (Insecta : Hemiptera) : insights from a combined morphological and molecular phylogeny. *Cladistics*, 35(1) : 67–105. *Cité pages 26, 50*
- Wen, Z., Baudry, J., Berenbaum, M. R., and Schuler, M. A. 2005. Ile115Leu mutation in the SRSI region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Engineering, Design and Selection*, 18(4) : 191–199. *Cité page 179*
- Wilding, C. S. 2018. Regulating resistance : CncC :Maf, antioxidant response elements and the overexpression of detoxification genes in insecticide resistance. *Current Opinion in Insect Science*, 27 : 89–96. *Cité pages 52, 54*
- Williamson, M. S., Denholm, I., Bell, C. A., and Devonshire, A. L. 1993. Knockdown resistance (kdr) to DDT and pyrethroid insecticides maps to a sodium channel gene locus in the housefly (*Musca domestica*). *Molecular and General Genetics MGG*, 240(1) : 17–22. *Cité pages 53, 62, 186*
- Xu, C., Li, C. Y.-T., and Kong, A.-N. T. 2005. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Archives of Pharmacal Research*, 28(3) : 249–268. *Cité page 51*
- Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J., and Chen, S. 2012. FastUniq : A Fast De Novo Duplicates Removal Tool for Paired Short Reads. *PLoS ONE*, 7(12) : 1–6. *Cité page 144*
- Xu, Y., Jiang, J., Wang, Z., Tian, J., Shi, W., and Cao, C. 2022. Characterization of a Novel Pathogenic Reovirus in Grasshoppers. *Viruses*, 14(12) : 1–11. *Cité pages 158, 199*
- Yahouédo, G. A., Chandre, F., Rossignol, M., Ginibre, C., Balabanidou, V., Mendez, N. G. A., Pigeon, O., Vontas, J., and Cornelie, S. 2017. Contributions of cuticle permeability and enzyme detoxification to pyrethroid resistance in the major malaria vector *Anopheles gambiae*. *Scientific reports*, 7(1) : 11091. *Cité pages 50, 189*

- Yang, X., Deng, S., Wei, X., Yang, J., Zhao, Q., Yin, C., Du, T., Guo, Z., Xia, J., Yang, Z., Xie, W., Wang, S., Wu, Q., Yang, F., Zhou, X., Nauen, R., Bass, C., and Zhang, Y. 2020. MAPK-directed activation of the whitefly transcription factor CREB leads to P450-mediated imidacloprid resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 117(19) : 10246–10253. *Cité page 55*
- Yang, Y., Zhang, Y., Wang, A., Duan, A., Xue, C., Wang, K., Zhao, M., and Zhang, J. 2022. Four MicroRNAs, miR-13b-3p, miR-278-5p, miR-10483-5p, and miR-10485-5p, Mediate Insecticide Tolerance in *Spodoptera frugiperda*. *Frontiers in Genetics*, 12(January) : 1–11. *Cité page 193*
- Yeaman, S. and Whitlock, M. C. 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution; international journal of organic evolution*, 65(7) : 1897–1911. *Cité page 68*
- Zhang, H.-Y., He, H., Chen, L.-B., Li, L., Liang, M.-Z., Wang, X.-F., Liu, X.-G., He, G.-M., Chen, R.-S., Ma, L.-G., and Deng, X. W. 2008. A Genome-Wide Transcription Analysis Reveals a Close Correlation of Promoter INDEL Polymorphism and Heterotic Gene Expression in Rice Hybrids. *Molecular Plant*, 1(5) : 720–731. *Cité page 63*
- Zhang, Y., Park, C., Bennett, C., Thornton, M., and Kim, D. 2021. Rapid and accurate alignment of nucleotide conversion sequencing reads with HISAT-3N. *Genome research*, 31(7) : 1290–1295. *Cité page 144*
- Zhou, C., Yang, H., Wang, Z., Long, G. Y., and Jin, D. C. 2019. Protective and detoxifying enzyme activity and ABCG subfamily gene expression in *Sogatella furcifera* under insecticide stress. *Frontiers in Physiology*, 10(JAN) : 1–12. *Cité page 52*
- Zhu, F., Wigginton, J., Romero, A., Moore, A., Ferguson, K., Palli, R., Potter, M. F., Haynes, K. F., and Palli, S. R. 2010. Widespread distribution of knockdown resistance mutations in the bed bug, *Cimex lectularius* (Hemiptera : Cimicidae), populations in the United States. *Archives of Insect Biochemistry and Physiology*, 73(4) : 245–257. *Cité pages 29, 54, 185*
- Zhu, F., Gujar, H., Gordon, J. R., Haynes, K. F., Potter, M. F., and Palli, S. R. 2013. Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Scientific Reports*, 3 : 1–8. *Cité pages 52, 54, 189*
- Zhu, K. Y. 2008. Insecticide Resistance. In J. L. Capinera, editor, *Encyclopedia of Entomology*, pages 1979–1981. Springer Netherlands, Dordrecht. *Cité page 48*