



HAL
open science

Bases génétiques et mécanismes cytologiques à l'origine de la diversité de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens*

Manon Bonneau

► To cite this version:

Manon Bonneau. Bases génétiques et mécanismes cytologiques à l'origine de la diversité de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens*. Sciences agricoles. Université Montpellier, 2018. Français. NNT : 2018MONTG053 . tel-02010568

HAL Id: tel-02010568

<https://theses.hal.science/tel-02010568>

Submitted on 7 Feb 2019

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

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

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

en Évolution des systèmes infectieux

École doctorale GAIA

Institut des Sciences de l'Évolution-Montpellier (ISE-M)

Bases génétiques et mécanismes cytologiques à l'origine de la diversité de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens*

Présentée par Manon BONNEAU

Le 22 Novembre 2018

Sous la direction de Mathieu SICARD
et Mylène WEILL

Devant le jury composé de

Richard CORDAUX, Directeur de recherche, CNRS Poitiers

Marylène POIRIE, Professeure des universités, Université côte d'Azur

Franck DEDEINE, Maître de conférences, Université de Tours

Isabelle MORLAIS, Directrice de recherche, IRD Montpellier (Présidente du Jury)

Mathieu SICARD, Professeur des universités, Université de Montpellier

Mylène WEILL, Directrice de recherche, CNRS Montpellier

Rapporteur

Rapporteuse

Examineur

Examinatrice

Directeur de thèse

co-Directrice de thèse



UNIVERSITÉ
DE MONTPELLIER

« Look deep into nature,
and then you will understand everything better. »

Albert Einstein

A ma famille

A Corentin

Remerciements

Je souhaiterais commencer ces remerciements par un grand merci à [Richard Cordaux](#), [Marylène Poirié](#), [Franck Dedeine](#) et [Isabelle Morlais](#) d'avoir accepté de faire partie de mon jury de thèse.

Best Team Ever :

Je souhaite dire un millier de mercis à tous les membres actuels et passés de l'équipe EVAS :

[Sandra](#), pour tout ce que tu as fait pour moi, pour tous ces bons moments passés à faire les souches ensemble, à fumer nos clopes, à faire le ménage en BM, à faire les backcross et à râler contre le monde !

[Pierrick](#), parce que sans toi les stats ça aurait été vraiment très dur et mon aquarium bien vide !

[Haoues](#), pour m'avoir écouté quand ça n'allait pas.

[Monsieur Patou](#), pour avoir passé ton temps à te payer ma tête et me laisser me moquer de toi sans jamais de représailles (ou alors tu attends la fin pour te venger de tout) !

[Nicole](#), pour avoir corrigé les articles et m'avoir fait profiter de tes connaissances.

[Marco](#), pour l'immense aide que tu m'as apportée pour les croisements.

[Arnaud](#), pour m'avoir montré que la qPCR n'a jamais tué personne.

[Pascal](#), pour m'avoir prévenu que j'allais souffrir.

[Célestine](#), pour avoir donné toute cette énergie pendant ta thèse ce qui m'a permis de travailler sur un jeu de données de rêve.

[Alex](#), pour ton cynisme et puis voilà.

[Rudy](#), pour avoir été une bouffée d'air magnifique dans ma thèse, mon partenaire de babyfoot, le kluut de l'ISEM !

Je voudrais aussi remercier les stagiaires qui m'ont aidé pendant ma thèse : [Mathilde](#), [Aude](#), [Victoria](#), [Éric](#) et [Julie](#) sans lesquels je n'aurais pas eu autant de données et ne me serais pas rendu compte à quel point encadrer des étudiants c'est cool (quoiqu'un peu stressant). Merci à vous cinq.

Enfin je souhaiterais faire une mention spéciale à [Faby](#) car même si tu ne fais pas officiellement partie de l'équipe EVAS, pour moi tu fais partie de notre petite famille. Merci Faby pour tout ce que tu as fait pour moi pendant ces trois ans. Merci pour tout ce que tu m'as appris en BM, pour avoir accepté de travailler avec moi, pour ces longues heures passées en bactério à étaler des boîtes de pétri, pour ces 800 kilos de verveine, pour m'avoir appris à faire des savons, pour avoir pété des câbles avec moi contre l'électroporateur et pour tout le reste !

Best Lab Ever

Je souhaiterais maintenant remercier les personnes de l'ISEM qui ont énormément compté pour moi au cours de ces cinq années passées à l'ISEM d'abord en tant que stagiaire de master I, II puis doctorante :

Merci à [Emmanuel Douzery](#) pour m'avoir fait découvrir et aimer la phylogénie au point de venir faire mon stage de master I dans ton équipe. Sans toi je n'aurais probablement pas découvert l'ISEM aussi rapidement.

Merci à [Emeric Figuet](#) et [Nicolas Galtier](#) pour m'avoir encadré en stage de master I et m'avoir donné goût à la bioinfo.

Merci à [Julien Veysier](#) et [Khalid Belkhir](#) pour m'avoir tant aidé sur l'assemblage du génome de *Wolbachia*, j'ai appris énormément de choses en bioinfo grâce à vous.

Merci à [Thomas Cantinelli](#), [Erick Desmarais](#) et [Frédérique Cerqueira](#) pour m'avoir aidé dans la mise au point et dans la réalisation du séquençage Sanger des milliers de clones que nous avons séquencé pendant ma thèse !

Merci à [Philippe Clair](#) pour m'avoir tout appris en qPCR, pour avoir pris le temps de me former, de m'expliquer, d'analyser mes résultats, de réfléchir à de nouvelles méthodes, de répondre à mes centaines de questions, à me rassurer sur la qualité de mes résultats ... merci pour tout.

Merci à [Laurent Marivaux](#) et [Catherine Girard](#) pour m'avoir permis d'utiliser leurs loupes binoculaires pour prendre mes milliers de photos de pontes de moustiques.

Merci à [Julie Perez](#) pour m'avoir formé à la microscopie optique et à épifluorescence.

Merci [Anna-Sophie Fiston-Lavier](#) pour l'aide que tu m'as apportée en bio-informatique.

Merci à [Dimitri Montviloff](#) et [Yannick Frontera](#), toujours prêts à se foutre de moi mais toujours là pour régler mes problèmes d'ordinateur (sans trop me prendre pour une quiche).

Merci à [Florence Saidani](#) pour avoir eu la patience de m'aider dans mes ordres de mission et autres sujets administratifs.

Merci à l'équipe [PhylEvolMol](#) de m'avoir adopté le midi (enfin plutôt 11h30) même si je ne parle pas de science à table.

Merci à [Frédéric Delsuc](#) pour avoir été mon associé de relais lors du marathon et pour m'avoir aidé à faire des phylogénies en réseau.

Merci à [Mélanie Debiais-Thibaud](#) pour m'avoir aidé dans mon monitorat et prit le temps de m'expliquer l'induction mésodermique. Merci aussi à [Hayet Bousseliou](#), [Colin Durand](#), [Olivier Guirado](#), [Nicolas Moureau](#) et [Thierry Vindolet](#) du bâtiment 4 pour avoir été si accueillants pendant ces trois années d'enseignement.

Merci à [François Rousset](#) pour avoir éclairé mon chemin dans la compréhension des obscurs GLMMs.

Merci à [Claire Gougat-Barbera](#) pour les petites sessions gossip avec Sandra et Faby !

Merci à [Fadela Tamoune](#), pour son énergie et tous les super trucs qu'elle a organisé pendant mes 5 années à l'ISEM : les journées des doctorants, les galettes des rois, le tournoi de babyfoot, les séminaires du lundi ...

Enfin, merci à [Agnès Mignot](#) de m'avoir si chaleureusement accueilli à l'ISEM dès mon arrivée en master, d'avoir permis une si bonne ambiance entre doctorants en nous laissant organiser le weekend d'intégration des nouveaux doctorants. Tu es la Best Big Boss ever

Best Collaborators Ever

Merci à [Frédéric Landmann](#) du CRBM de Montpellier pour m'avoir formé en cytologie et avoir passé de nombreuses heures à regarder mes lames au confocal.

Merci aussi à [Kevin Terretaz](#) du CRBM de Montpellier pour avoir pris le temps de m'expliquer ses expérimentations en cellules de nos variants *cidA* et *cidB*.

Merci à [Fabrice Vavre](#) du LBBE de Lyon et [Sophie Gaudriault](#) du DGIMI de Montpellier pour avoir accepté de faire partie de mon comité de thèse et avoir permis des discussions intenses et enrichissantes.

Merci à [Sylvain Charlat](#) du LBBE de Lyon, pour avoir m'avoir fait partager toutes tes connaissances sur l'IC au travers de discussions très fertiles qui ont mis mon cerveau à rude épreuve.

Merci à [Beniamino Caputo](#) et [Daniele Arnoldi](#) pour nous avoir si chaleureusement accueillis à Trento et nous avoir aidé à collecter des moustiques dans ces lieux paradisiaques que sont les caniveaux de San Michele, Roveré Della Luna, Mezzocorona et Zambana. Mais bon, les Dolomites ça valait vraiment le coup !

Merci à [John Beckmann](#), [Mark Hochstrasser](#), [Hervé Merçot](#) et [Denis Poinsot](#) pour ce brain storming autour de notre Review en cours d'écriture dans Trends in Genetics !

Best Family Ever

Parce que la thèse ce n'est pas un long fleuve tranquille et quand ça ne va pas, la famille est toujours là !

Merci à ma maman [Sylvie](#) et mon papa [Éric](#) pour avoir fait l'effort d'essayer de comprendre mon sujet de thèse, pour avoir corrigé les fautes d'orthographe de mon manuscrit, pour m'avoir toujours soutenu toute ma vie dans mes choix, pour avoir été là quand ça n'allait pas, pour être les meilleurs parents du monde !

Merci à mon frère [Titouan](#) même s'il ne comprend pas l'intérêt de faire de la science théorique qui ne rapporte pas directement d'argent.

Merci à mes grands-parents [Mamie Monique](#), [Mamie Nicole](#), [Papy Yves](#) et [Martine](#) d'avoir cru en moi et de m'avoir soutenu !

Merci à mon pâté [Corentin](#) qui a vécu ces trois ans de thèse au moins aussi intensément que moi. Merci de m'avoir supporté dans cette voie même si ça n'a pas été facile du fait de mes sautes d'humeur, du stress, de la vie de none (couchée 20h30 levée 6h45), des nombreuses heures et week-end passé au labo... Sans toi je n'y serais jamais arrivée. Cette thèse c'est aussi un peu la tienne car après toutes ces heures passées à discuter de mes résultats, des manips qui ne marchent pas ... tu connais parfaitement *Wolbachia* !

Merci au reste de ma famille : [Edna](#), [Lylou](#), [Sissi](#), [Dorine](#), [Tonton Hervé](#), [Lina](#), [Pascal](#), [Maria](#), [Mémé Mimine](#), [Loulou](#), [Thomas](#), [Manu](#) et [Tiphaine](#) pour me rappeler qu'il n'y a pas que le travail dans la vie.

Enfin, je souhaiterais remercier tout particulièrement la **M⁴ Team** car si ma thèse est telle qu'elle est aujourd'hui c'est grâce à vous !

Mathieu, tu as passé cette thèse à me dire que j'étais chanceuse, eh bien oui j'ai été très chanceuse de vous avoir tous les trois avec moi pendant cette aventure !

Tout d'abord, dix mille mercis à **Mine**. Merci d'avoir partagé ces trois ans de thèse avec moi. Merci pour tous ces merveilleux moments à jouer du ukulélé, à faire de la course à pied, à fumer des clopes, à réfléchir sur les stats, à râler contre les qPCRs, à faire la fête, à disséquer des ovaires de moustiques, à faire les souches ... Je ne saurais comment te remercier et te montrer à quel point ton soutien dans les moments durs m'a permis de garder le cap et ne pas tout lâcher. Merci pour tout !

Un grand merci à **Mylène** de m'avoir encadré durant cette thèse et de m'avoir fait grandir tant personnellement que scientifiquement même si ça c'est un peu fait dans la douleur au début. Merci d'avoir su me guider au travers de la thèse, d'avoir cru en moi et d'avoir partagé avec moi toutes tes connaissances, de m'avoir secoué quand je perdais pied, de m'avoir entraîné dans tes folies de croisements... Merci de m'avoir montré que la science c'est merveilleux même si ce n'est pas le monde des Bisounours. Merci, Merci Merci !

Enfin, un immense merci à **Mathieu** pour m'avoir recueilli en stage de master II. Pour avoir vu en moi des qualités bien enfouies derrière mon immaturité et mon impertinence de jeune fêtarde de 22 ans. Merci d'avoir voulu continuer l'aventure avec moi pendant trois ans de thèse et de m'avoir donné l'opportunité d'en faire une sur un sujet aussi passionnant avec des gens tellement géniaux. Merci de m'avoir fait évoluer de la jeune étudiante en master à la presque chercheuse que je suis maintenant. Même si ça n'a pas toujours été facile entre nous, nous avons réussi à mettre nos rancœurs et défauts de côté et à travailler ensemble pour aboutir à une thèse dont je suis fière. Merci pour tout.

PS : pour les copains, il y a une surprise à la fin !

Sommaire

Introduction générale « symbioses et parasites du sexe »	1
I. Evolution des systèmes symbiotiques	3
a. Le concept de symbiose englobe une très grande variété d'interactions durables.....	3
b. Les symbioses source d'innovation et moteur de l'évolution.....	4
c. Importance du mode de transmission dans l'évolution des systèmes symbiotiques	5
II. Le parasitisme du sexe	8
a. La transmission uniparentale à l'origine de la manipulation du sexe	8
b. Les différentes stratégies de manipulations du sexe.....	9
i. Féminisation.....	9
ii. « Male killing ».....	11
iii. Passage de la parthénogenèse arrhénotoque à la parthénogenèse thélytoque	13
iv. Incompatibilité Cytoplasmique	15
c. Les conflits génétiques entre hôtes et manipulateurs.....	15
d. Les conséquences évolutives du parasitisme sexuel.....	17
i. Modification du déterminisme du sexe	17
ii. Modification des patrons de sélection sexuelle et de choix du partenaire.....	17
iii. Evolution vers l'asexualité	18
iv. Extinction.....	19
III. <i>Wolbachia</i> reine de la manipulation	20
a. <i>Wolbachia</i> l'endosymbiote le plus répandu chez les arthropodes	20
b. Diversification des <i>Wolbachia</i> : la conséquence de transferts horizontaux ?.....	21
c. Les génomes des <i>Wolbachia</i> : des génomes très dynamiques.....	23
i. Impact du mode de vie intracellulaire sur la taille et l'évolution des génomes	23
ii. Des génomes riches en éléments génétiques répétés mobiles et en protéines à domaine ankyrine.....	25
iii. Implications des éléments génétiques mobiles et des protéines à domaine ankyrine dans l'évolution des génomes.....	27
d. <i>Wolbachia</i> peut même devenir mutualiste ou pas.....	28

IV. <i>Wolbachia</i> et l'incompatibilité cytoplasmique	29
a. Direction et diversité de l'IC	29
i. Direction de l'IC	29
ii. Diversité des phénotypes d'IC : <i>C. pipiens</i> mène la danse	31
b. Pénétrance de l'IC	35
c. Dynamique de l'infection sous l'influence de l'IC	36
d. Conséquences évolutives de l'IC	38
e. Conceptualisation des mécanismes l'IC	38
f. Les mécanismes cytologiques sous-jacents à l'IC	41
g. Les bases génétiques et moléculaires sous-jacentes à l'IC	46

Partie I Les bases génétiques de l'IC et de sa diversité chez *Culex pipiens* ... 51

I. Introduction..... 53

Article n°1 La diversité des phénotypes d'IC observée chez le moustique *Culex pipiens* est gouvernée par un opéron amplifié et polymorphe présent dans le génome de *Wolbachia*..... 57

Article n°2 La fonction *mod* de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens* est gouvernée par le gène *cidB* et non par le gène *cidA*..... 87

II. Résumé des résultats & Discussion..... 113

a. L'opéron *cidA/cidB* est amplifié et polymorphe dans les génomes des wPips..... 113

b. La variabilité de l'opéron *cidA/cidB* est très probablement responsable de la diversité des phénotypes d'IC chez *C. pipiens*..... 113

c. Le groupe wPipIV lève le voile sur les rôles de *cidA* et *cidB* dans l'IC chez *C. pipiens*..... 114

Partie II Les mécanismes cytologiques sous-jacents à l'IC et sa diversité chez *Culex pipiens*..... 121

I. Introduction..... 123

Article n°3 Les phénotypes cellulaires de l'incompatibilité cytoplasmique chez *Culex pipiens* au regard de la diversité de *cidB*..... 127

II. Résumé des résultats & Discussion	167
a. Un seul phénotype cellulaire à l'origine de l'IC chez <i>C. pipiens</i>	167
b. L'exclusion complète de la chromatine paternelle à l'origine du développement haploïde.....	168
c. L'intensité cellulaire de l'IC sous la dépendance des <i>Wolbachia</i> présentes chez les mâles et de la diversité de <i>cidB</i>	168
d. L'absence de <i>Wolbachia</i> chez les femelles empêche tout développement haploïde de se produire.....	170

Discussion & Perspectives173

I. Les phages au cœur de l'IC	176
a. Une origine phagique de l'IC : une vieille histoire.....	176
b. Un système toxine-antidote à la base de l'IC.....	177
c. Les phages impliqués dans la diversité de l'IC : le rôle des transferts horizontaux.....	179
d. Les phages, à l'origine de l'IC si particulière de <i>C. pipiens</i> ?.....	180
II. Lien entre diversité des variants <i>cidA-cidB</i> et diversité des phénotypes d'IC chez <i>C. pipiens</i>	181
a. A l'échelle cellulaire.....	181
b. Prédire et maîtriser les phénotypes d'IC : des analyses à différentes échelles.....	183
c. Nos données soutiennent l'hypothèse d'un système toxine-antidote.....	185
III. Sur quel déterminisme génétique est basée la diversité de l'IC chez les drosophiles ?	187
IV. Quid de l'IC induite par d'autres endosymbiotes : une toute autre histoire ?	190
V. L'IC dans la lutte antivectorielle	192

Bibliographie.....195

Annexes.....221

Introduction générale
« Symbioses et parasites du sexe »

I. Evolution des systèmes symbiotiques

a. Le concept de symbiose englobe une très grande variété d'interactions durables

Tous les organismes sur terre se retrouvent tout au long de leur vie en interaction avec d'autres organismes qu'ils appartiennent à la même espèce (interactions intraspécifiques) ou à des espèces différentes (interactions interspécifiques ou hétérospécifiques). Ces interactions entre organismes ont joué et jouent toujours un rôle central dans le fonctionnement du vivant tant d'un point de vue écologique qu'évolutif. Chaque organisme peut interagir ponctuellement avec d'autres organismes soit de manière antagoniste (prédation, compétition ...) soit de manière synergique (coopération). En revanche, certains organismes peuvent interagir plus durablement avec d'autres organismes sur tout ou une partie de leur cycle de vie (Combes, 2001; Combes *et al.*, 2018). Ces interactions durables sont aussi qualifiables de symbioses qui vient du Grec « vivre ensemble » et qui traduit la capacité d'organismes à interagir durablement ensemble. L'un des organismes, le plus gros, étant appelé l'hôte quand les autres organismes plus petits sont qualifiés de symbiotes. En fonction des étapes de son cycle de vie, un organisme peut changer de rôle et passer de l'état d'hôte à l'état de symbiote. Un organisme peut aussi être à la fois l'hôte d'un symbiote tout en étant le symbiote d'un autre hôte, ce qui crée des niveaux d'interaction emboîtés et complexifie grandement les interactions. Les nombreux organismes qui s'associent durablement appartiennent à des clades taxonomiques très différents et leurs associations revêtent une infinité de modalités distinctes. Cependant, trois paramètres permettent de qualifier et de décrire les symbioses.

Le premier paramètre concerne la résultante globale en termes de coûts et bénéfices sur la valeur sélective de chacun des partenaires de l'interaction (Cheng, 1991). Ainsi, si l'interaction entraîne au final un coût pour l'hôte et un bénéfice pour le symbiote, le symbiote est considéré comme un parasite. Si l'interaction est neutre (si de tels cas existent réellement), c'est-à-dire que l'interaction entre les deux partenaires n'entraîne ni coût ni bénéfice, alors les deux partenaires sont dit commensaux. Enfin, si l'interaction entraîne des bénéfices pour les deux protagonistes, on parle alors de mutualisme. Parasitisme et mutualisme sont les deux situations extrêmes d'un continuum d'interactions dont l'impact sur l'hôte peut varier en fonction de l'état de l'hôte ou de l'environnement (Bronstein, 1994). Par exemple, le passage du mutualisme au parasitisme entre deux protagonistes est bien illustré dans l'interaction durable entre la bactérie *Hamiltonella defensa* et le puceron *Acyrtosiphon pisum* où la bactérie confère au puceron une résistance contre le parasitoïde *Aphidus ervi*. Dans le cas d'une forte pression parasitaire exercée par le parasitoïde, la bactérie confère un avantage au puceron car les pucerons infectés ont une survie accrue par rapport aux non infectés (Oliver *et al.*, 2003). En revanche, en l'absence du parasitoïde, le coût que confère l'infection par *Hamiltonella* n'est plus compensé par le bénéfice de la protection et les individus infectés sont alors moins compétitifs que les individus asymbiotiques (Oliver *et al.*, 2003, 2008). Cette transition du mutualisme au parasitisme révèle que

l'avantage conféré par un symbiote dans certaines situations peut compenser le coût qu'il induit dans d'autres situations.

Le deuxième paramètre permettant de décrire les symbioses est leur caractère obligatoire *versus* facultatif (Brian, 1966) qui traduit la capacité ou non d'un organisme à se développer, survivre et se reproduire sans son partenaire (Smith & Douglas, 1987). Ainsi, dans les interactions où le symbiote agit comme un parasite pour son hôte, l'interaction est bien souvent obligatoire pour le symbiote alors qu'elle est complètement facultative pour l'hôte. L'obligation réciproque d'interaction pour les deux partenaires concerne davantage les symbioses mutualistes. En effet, dans de telles symbioses, les deux partenaires ont besoin l'un de l'autre pour se développer, se protéger, se nourrir, se reproduire etc.

Enfin, le dernier paramètre permettant de décrire les symbioses est celui du niveau d'association physique entre les partenaires (Smith & Douglas, 1987). Celui-ci varie le long d'un continuum allant d'un simple contact entre deux organismes (possibilité de dissocier facilement les partenaires) comme c'est le cas par exemple pour les pollinisateurs et les fleurs, à une association physique totale (un des partenaires vit exclusivement à l'intérieur de l'autre) comme par exemple les endosymbiotes des insectes qui vivent dans les cellules de leurs hôtes.

Ces trois paramètres constituent autant de continuums le long desquels les interactions durables entre les organismes peuvent « se déplacer » résultant ainsi en une immense diversité de symbioses.

b. Les symbioses source d'innovation et moteur de l'évolution

Il apparaît que dans le cas des interactions durables, les valeurs sélectives des partenaires sont plus ou moins étroitement liées. Ainsi, la valeur sélective d'un organisme jouant le rôle d'hôte dépend du ratio coût/bénéfice associé à la présence de symbiotes. Les génomes des deux partenaires en plus de déterminer leurs phénotypes propres vont aussi s'influencer mutuellement pour créer ce que Richard Dawkins a défini en 1982 comme le phénotype étendu. Ce phénotype étendu constitue un nouveau niveau sur lequel la sélection naturelle va pouvoir agir et entraîner de multiples conséquences sur l'évolution des deux partenaires. Parmi ces conséquences on peut citer l'apparition de nouvelles fonctions comme par exemple la transformation de l'azote atmosphérique en ammonium par les bactéries *Sinorhizobium meliloti* dans les nodosités de la légumineuse *Medicago sativa* alors que ces dernières sont incapables de le faire hors des nodosités (Sigaud *et al.*, 1999) ; l'apparition de nouveaux organes comme les bactériomes chez les arthropodes qui hébergent leurs endosymbiotes ; mais aussi la perte d'organes et de fonctions comme c'est souvent le cas de la locomotion et de la digestion lors du passage chez les métazoaires de la vie libre à la vie parasitaire (Combes *et al.*, 2018) ; la réduction de la taille des génomes des symbiotes du fait de redondances métaboliques entre l'hôte et le symbiote (Moran *et al.*, 2008) ou encore l'apparition de nouvelles niches écologiques comme pour les punaises *Megacopta*

chez lesquelles l'acquisition du symbiote *Ishikawaella capsulata* a permis la colonisation d'une nouvelle plante hôte (Hosokawa *et al.*, 2006, 2007).

Les interactions durables, en plus d'être source d'innovations, vont représenter un moteur à l'évolution. Les études de modélisation ont surtout mis en avant les interactions antagonistes entre organismes (comme la prédation, la compétition ou le parasitisme) comme moteur de l'évolution. En effet, dans de telles interactions, les intérêts des protagonistes sont opposés et il est facile de conceptualiser comment les conflits permanents existant entre ces organismes peuvent constituer des forces évolutives puissantes. Ainsi, dans les symbioses parasitaires, la sélection naturelle va mener à la mise en place i) chez les parasites de mécanismes parfois très complexes leur permettant d'exploiter leur hôte et ii) chez les hôtes de mécanismes pour éliminer ou tout au moins limiter les coûts induit par les parasites. En revanche, dans les symbioses mutualistes les organismes ne présentent pas de conflits apparents puisque leurs intérêts semblent globalement alignés. Il est ainsi difficile de conceptualiser comment ce type d'interaction pourrait être un moteur de l'évolution. En réalité, dans de telles symbioses, il existe un antagonisme sous-jacent car chacun des organismes afin de maximiser sa valeur sélective pourrait « tricher », soit en minimisant son investissement dans la symbiose, soit en surexploitant l'autre partenaire. Cette sélection de génotypes tricheurs peut mener à des conflits. De tels comportements de « tricheurs » peuvent conduire à l'évolution de caractères parfois extrêmes (Bronstein, 2001). Ainsi, dans certaines interactions plantes-pollinisateurs, des fleurs attirent les pollinisateurs avec une substance dépourvue de vraie ressource nutritive afin de se faire polliniser sans donner de ressource en contrepartie (car produire du nectar est coûteux, Bell 1986; Gilbert *et al.* 1991). La réciproque existe aussi : des pollinisateurs qui « volent » le nectar d'une fleur sans la polliniser (car transporter du pollen est coûteux, Irwin & Brody 1999; Bronstein *et al.* 2017). Cette course peut conduire comme chez les colibris et les fleurs qu'ils convoitent à l'apparition de tailles de becs et de formes de fleurs exubérants.

c. Importance du mode de transmission dans l'évolution des systèmes symbiotiques

La maintenance de l'interaction symbiotique à l'échelle évolutive ne peut avoir lieu que s'il y a transmission des symbiotes d'hôtes en hôtes, génération après génération. Deux grandes voies de transmission existent : la transmission verticale où le symbiote est transmis à la descendance principalement par la mère et la transmission horizontale où le symbiote est transmis d'un hôte à un autre hôte n'ayant pas forcément de lien de parenté *via* l'environnement ou la transmission sexuelle (Fine, 1975). Ces deux voies ne sont pas mutuellement exclusives car certains endosymbiotes comme *Hamiltonella defensa* qui infectent le puceron *Acyrtosiphon pisum* sont capables de se transmettre par les deux voies : on parle alors de transmission mixte (Moran & Dunbar, 2006; Ebert, 2013). Ces modes

de transmission vont avoir des conséquences sur la trajectoire évolutive des interactions hôtes-symbiotes.

Dans le cas où la transmission horizontale est majoritaire, les valeurs sélectives des partenaires ne sont que partiellement associées (*i.e.* la reproduction du parasite ne dépend pas de celle de son hôte contrairement à la transmission verticale) et l'on peut prédire que le système symbiotique va évoluer plus facilement vers une interaction hôte-parasite. La virulence d'un parasite (*i.e.* l'ensemble des effets néfastes sur la survie et/ou la fécondité induits par la présence du parasite) et son évolution au sein des symbiotes transmis horizontalement sont très variables. Ceci est dû à l'existence d'un compromis pour le symbiote entre sa virulence et sa transmission. En effet, une forte virulence accroît la capacité de transmission car elle est souvent la conséquence d'une forte reproduction du parasite à l'intérieur de son hôte (Mackinnon & Read, 1999, 2003). Cependant, une trop forte virulence peut entraîner la mort de l'hôte et donc diminuer considérablement les chances de propagation du parasite. Il y a alors sélection d'une virulence optimale modérée réalisant le « meilleur » compromis entre virulence et transmission (Ewald & De Leo, 2002; Alizon *et al.*, 2009). Même si la transmission horizontale favorise théoriquement l'évolution vers le parasitisme, il existe de nombreux exemples de transmission horizontale dans des symbioses mutualistes. Par exemple, la transmission horizontale est essentielle à la symbiose mutualiste entre la seiche *Euprymna scolopes* et la bactérie *Vibrio fischeri*. Celle-ci permet aux bactéries de coloniser des seiches juvéniles asymbiotiques afin d'établir la symbiose permettant à la seiche d'acquérir de la bioluminescence favorisant son camouflage et aux bactéries de se multiplier dans un environnement favorable (Nyholm & McFall-Ngai, 2004). Cependant, la maintenance du mutualisme au cours de l'évolution lorsque la transmission est horizontale apparaît plus chaotique du fait de « l'imprédictibilité » de la présence des deux partenaires ou de l'apparition de tricheurs (Bronstein, 2001). Plusieurs explications existent pour expliquer le maintien du mutualisme malgré ce mode de transmission. Premièrement, la notion de « fidélité » au partenaire propose que des interactions régulières ou sur le long terme entre deux espèces doivent favoriser les comportements de coopération car si un des partenaires meurt suite à une surexploitation par l'autre partenaire, alors le partenaire restant devra payer le coût d'être sans partenaire et d'en trouver un nouveau (Bull & Rice, 1991). La fidélité au partenaire est particulièrement forte lorsqu'un des partenaires est rare dans l'environnement. Deuxièmement, le mutualisme peut être maintenu grâce à la mise en place par l'hôte de sanctions envers des symbiotes tricheurs permettant leur exclusion et la sélection des symbiotes les plus coopératifs (Bull & Rice, 1991; Simms *et al.*, 2006; Regus *et al.*, 2014). Enfin, un fort degré de parenté entre les symbiotes tendrait à favoriser le mutualisme (Frank, 1996). En effet, un fort degré de parenté réduirait la compétition entre les symbiotes dans l'hôte. Cette diminution de compétition sélectionnerait pour une transmission horizontale induisant un coût minimum pour l'hôte. C'est ce qui est observé chez les termites champignoniste du groupe des Macrotermitinae qui transmettent horizontalement un seul

variant du champignon *Termitomyces* malgré la présence d'une multitude de variants dans les cultures de la colonie (Aanen *et al.*, 2009).

La transmission verticale de symbiotes peut être une « solution » pour maintenir le mutualisme. Celle-ci peut se faire soit avant la fécondation, c'est-à-dire que les symbiotes sont déjà présents dans les gamètes comme c'est le cas pour les endosymbiotes *Wolbachia* (Dobson *et al.*, 1999) et *Buchnera* (P, 1965), soit après la fécondation comme les bactéries du tube digestif de *Megacopta punctatissima* déposées sur les œufs dans des capsules qui seront ingérées par les nouveau-nés dès l'éclosion (Fukatsu & Hosokawa, 2002). Les symbiotes qui sont transmis verticalement, sont majoritairement transmis de manière maternelle de la mère à ses descendants par les ovocytes. Cette asymétrie de transmission vers les femelles serait en partie due à l'anisogamie (asymétrie de taille des gamètes) observée chez beaucoup d'espèces. Ainsi, les ovocytes qui contiennent une quantité de cytoplasme beaucoup plus conséquente que les spermatozoïdes constitueraient un milieu plus favorable pour les symbiotes (Thomas *et al.*, 2010). Il existe de rares cas où la transmission peut être uniquement paternelle comme chez les termites champignonnistes *Macrotermes bellicosus* où le champignon qu'elle cultive *Termitomyces* est transmis par les mâles. On trouve aussi de rares cas de transmission verticale biparentale comme chez *Steinernema*, un nématode entomopathogène dont les entérobactéries du genre *Xenorhabdus* sont transmises par les deux parents (Sicard *et al.*, 2004).

Dans des cas de transmission strictement verticale transovocytaire, il n'y a pas d'étape où l'hôte se retrouve sans le symbiote et où le symbiote se retrouve libre dans le milieu. C'est pourquoi, le nombre de descendants d'un partenaire dépend de l'autre et réciproquement. Par conséquent, les intérêts du symbiote et de l'hôte convergent et leurs valeurs sélectives sont indissociables. Du fait de ce mode de transmission, les symbiotes conférant un avantage sélectif à leur hôte vont être sélectionnés car dans le cas contraire, la lignée hôte dont ils dépendent serait complètement désavantagée par rapport à des hôtes non infectés (Fine, 1975). Ainsi, de nombreux symbiotes transmis verticalement sont bénéfiques à leurs hôtes. Le bénéfice conféré par les symbiotes le plus observé est procuré *via* l'alimentation en palliant au régime alimentaire déséquilibré de leur hôte. C'est notamment le cas des bactéries *Sulcia muelleri* et *Buchnera aphidicola* qui supplémentent en acides aminés essentiels l'alimentation des insectes suceurs de sèves, respectivement de l'ordre des *Auchenorrhyncha* et de nombreuses espèces de pucerons dont la plus étudiée est *Acyrtosiphon pisum* (Douglas, 1998; McCutcheon & Moran, 2007; McCutcheon *et al.*, 2009; Akman Gündüz & Douglas, 2012). Les symbiotes peuvent par exemple protéger leurs hôtes contre des pathogènes et parasites qui viendraient les menacer. On peut citer certains symbiotes facultatifs (aussi appelés symbiotes secondaires) des pucerons *Acyrtosiphon pisum* qui les protègent partiellement des guêpes parasitoïdes (Oliver *et al.* 2003; voir Poirié & Coustau 2011 pour revue; Schmitz *et al.* 2012), *Wolbachia* qui confère une protection contre les virus à ARN chez son hôte *Drosophila melanogaster* (Chrostek *et al.*, 2013) ou encore la bactérie *Regiella insecticola* qui protège son hôte puceron *A. pisum* du champignon pathogène *Pandora neoaphidis* (Scarborough *et al.*, 2005).

Bien que la transmission verticale doive théoriquement mener l'interaction à évoluer de façon plus probable vers une relation mutualiste, il existe des cas où les intérêts des hôtes et des symbiotes ne sont pas toujours convergents. De telles situations peuvent entraîner des conflits génétiques entre hôtes et symbiotes et l'apparition de manipulation des hôtes par les symbiotes afin de favoriser leur transmission : le parasitisme du sexe (Hurst & Werren, 2001).

II. Le parasitisme du sexe

a. La transmission uniparentale à l'origine de la manipulation du sexe

Lorsqu'il y a transmission verticale des symbiotes par les deux sexes, la coexistence dans la descendance de génotypes symbiotiques différents peut mener à de la compétition entre eux pour l'espace et les ressources. Cette compétition entre symbiotes va sélectionner les symbiotes plus compétitifs (Frank, 1996). Ces symbiotes plus compétitifs peuvent surexploiter leur hôte et lui être désavantageux. Ainsi, vont être sélectionnés chez l'hôte des mécanismes permettant de réduire la compétition entre les symbiotes. Comme décrit précédemment, la compétition entre symbiote doit en théorie diminuer avec le degré de parenté. Les symbiotes provenant du même parent sont plus proches entre eux qu'avec ceux de l'autre parent. Ainsi, transmettre à la descendance uniquement les symbiotes d'un des deux parents constitue un bon moyen d'augmenter la parenté entre symbiotes. Avec la transmission uniparentale, les symbiotes se retrouvent associés uniquement à une lignée hôte et la sélection de parentèle peut donc favoriser l'optimisation d'exploitation de la ressource par les symbiotes (*i.e.* l'hôte). Cette théorie peut expliquer pourquoi la transmission verticale monoparentale a si souvent été sélectionnée pour les symbiotes et à plus large échelle pour les organites de la cellule eucaryote (Birky, 1995; Frank, 1996).

Cependant, lorsque la transmission d'un symbiote n'est faite que par un sexe, cela induit une asymétrie dans les pressions de sélection s'exerçant sur le symbiote et le génome de l'hôte. En effet, les gènes de l'hôte étant transmis de manière biparentale et ceux du symbiote de manière uniparentale, tout gène du symbiote permettant un biais de sex-ratio en faveur du sexe transmetteur sera sélectionné et des conflits pour la transmission entre génomes de l'hôte et du symbiote vont alors apparaître (Hurst & Werren, 2001). Ce sont ces conflits génomiques de transmission qui seraient à l'origine de la sélection chez les symbiotes de gènes favorisant leur transmission et responsables de la manipulation de la reproduction de leurs hôtes (Cosmides & Tooby, 1981; Hurst & Werren, 2001). Du fait de la transmission maternelle quasi exclusive des symbiotes, les mâles sont manipulés par les parasites de la reproduction, ce qui permet d'augmenter la survie et/ou la production de femelles infectées et du même coup la transmission des symbiotes (voir Engelstädter & Hurst 2009 pour revue).

b. Les différentes stratégies de manipulations du sexe

Du fait des différences biologiques et écologiques de leurs hôtes, de nombreuses stratégies de manipulation de la reproduction ont été sélectionnées chez les symbiotes. Certaines font intervenir un biais de sex-ratio vers le sexe transmetteur afin d'augmenter dans les populations les individus capables de transmettre le symbiote à la génération suivante (féminisation, « male-killing » et parthénogenèse thélytoque). Une autre, l'incompatibilité cytoplasmique, ne modifie pas le sex-ratio mais permet aux femelles infectées de produire en moyenne plus de descendants que les femelles non infectées. Quelle que soit la stratégie, ces différentes manipulations de la reproduction des hôtes par les symbiotes vont avoir des conséquences évolutives sur les populations hôtes différentes selon les stratégies.

i. Féminisation

La première stratégie de manipulation de la reproduction de leurs hôtes par les symbiotes s'appelle la féminisation (F). Elle a été décrite chez des espèces hôtes appartenant au clade des malacostracés et des insectes et est induite par des symbiotes bactériens mais aussi des microsporidies (Bandi *et al.*, 2001). Ces symbiotes vont transformer, au cours du développement, les mâles génétiques en femelles phénotypiques totalement fonctionnelles qui vont pouvoir les transmettre à la génération suivante (Figure 1). La féminisation est induite par *Wolbachia* chez les isopodes terrestre *Armadillidium vulgare*, *Porcellionides pruinosus*, ou encore *Oniscus asellus* (Rigaud *et al.*, 1991, 1997a, 1999; Cordaux *et al.*, 2004) et par les microsporidies *Nosema granulosis* et *Microsporidium sp.* chez l'amphipode *Gammarus duebeni* (Terry *et al.*, 1998; Ironside *et al.*, 2003). Chez les malacostracés, dont font partie tous ces hôtes, la différenciation en mâle est gouvernée principalement par une hormone extragonadale : l'hormone androgène (AGH) sécrétée par la glande androgène (Martin *et al.*, 1999). Cette hormone empêche le développement des ovaires, stimule la spermatogénèse et régule l'expression des caractères sexuels primaire et secondaire (Okuno *et al.*, 1997). Ainsi, chez les mâles génétiques en l'absence de cette hormone, seuls les caractères sexuels femelles se développent (Katakura, 1989). Durant l'embryogenèse, *Wolbachia* et *N. granulosis* induisent respectivement chez des isopodes terrestres comme *A. vulgare* ou chez les amphipodes comme *G. duebeni* la féminisation des mâles génétiques en modulant le contrôle endocrinien de la différenciation sexuelle via l'empêchement de la production de l'hormone AGH (Rigaud, 1997; Rodgers-Gray *et al.*, 2004).

Chez les insectes, les mécanismes moléculaires par lesquels les symbiotes induisent la féminisation sont moins clairs que chez les malacostracés en particulier car le déterminisme du sexe est codé génétiquement dans chaque cellule. Chez *Eurema hecabe*, il a été mis en évidence que *Wolbachia* chez les mâles génétiques interagit avec le système de mise en place du sexe de manière continue durant

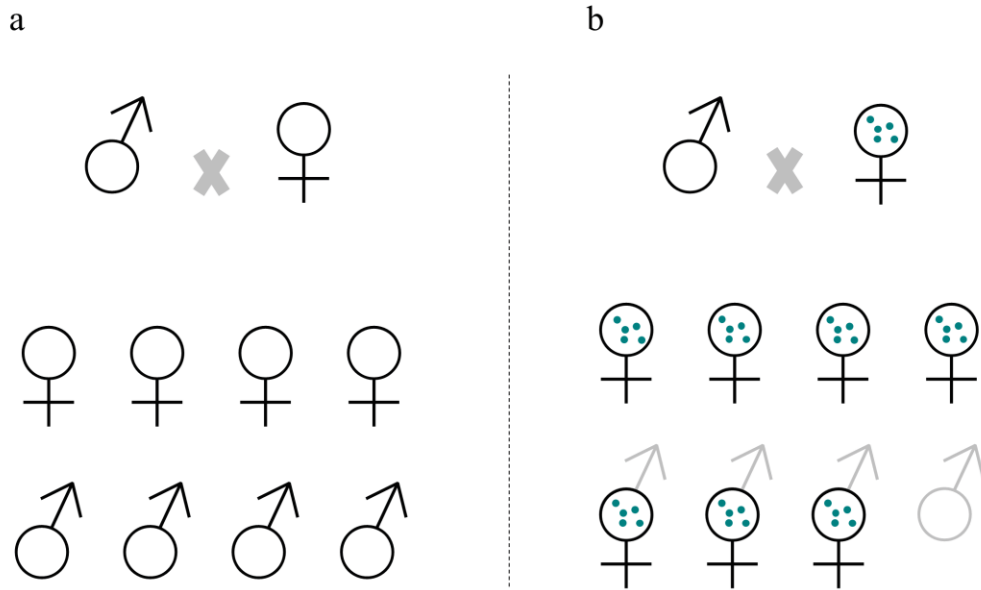


Figure 1: Féminisation.

Les points bleus représentent les *Wolbachia*.

a. Descendance attendue dans le cas où mâles et femelles ne sont pas infectés par *Wolbachia*.

b. Lorsque les femelles sont infectées par *Wolbachia* mais pas les mâles, les mâles de la descendance sont transformés en femelles.

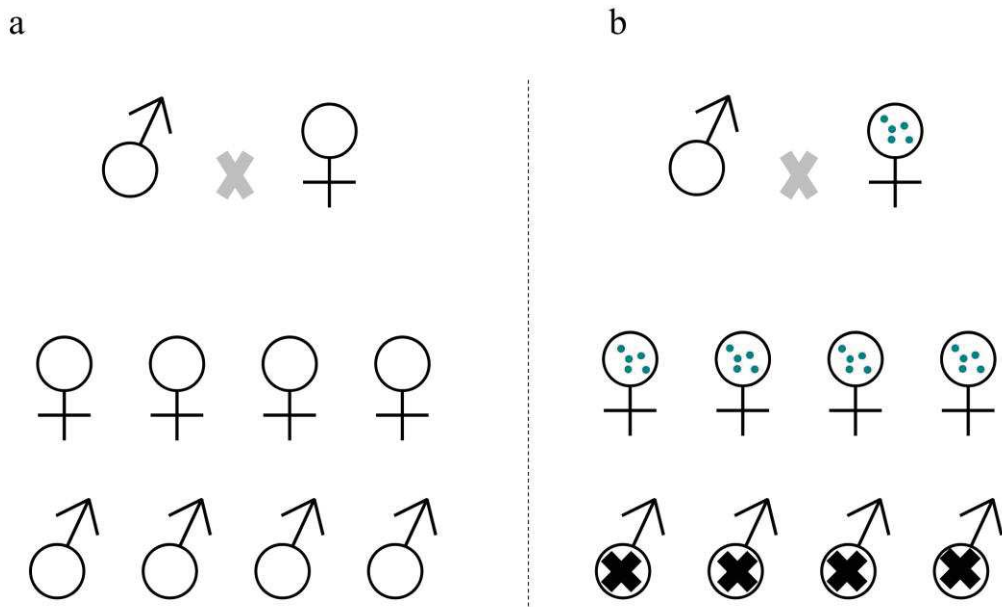


Figure 2 : Male-killing.

Les points bleus représentent les *Wolbachia*.

a. Descendance attendue dans le cas où mâle et femelle ne sont pas infectés par *Wolbachia*.

b. Lorsque la femelle est infectée par *Wolbachia*, les mâles de la descendance vont mourir.

l'embryogenèse afin de produire des femelles phénotypiques totalement fonctionnelles, mais les mécanismes impliqués sont inconnus (Narita *et al.*, 2007).

ii. « Male killing »

La deuxième stratégie, qui s'appelle le « male killing » (MK, *i.e.* la mort des mâles), est induite par de nombreux symbiotes manipulateurs de la reproduction chez de nombreuses espèces hôtes. Ainsi, *Arsenophonus nasoniae* (Gherna *et al.*, 1991), *Flavobacterium sp.* (Hurst *et al.*, 1999), *Spiroplasma ixodetis* (Tinsley & Majerus, 2006), *Spiroplasma poulsonii* (Williamson *et al.*, 1999), ou encore *Wolbachia* (Charlat *et al.*, 2006) induisent le MK respectivement chez *Nasonia vitripennis*, *Adonia variegata*, *Anisosticta novemdecimpunctata*, *Drosophila willistoni*, et enfin *Hypolimnas bolina*. Du fait de l'incapacité des mâles à les transmettre, les symbiotes capables d'induire le MK entraînent la mort des descendants mâles produits par des femelle infectées et seules les descendantes femelles capables, elles, de transmettre les symbiotes à la génération suivante vont survivre (Figure 2). Deux types de MK ont été décrits : i) le MK précoce qui a lieu lors de l'embryogenèse et le MK tardif qui a lieu à un stade plus tardif de développement (typiquement le quatrième stade larvaire) et qui est uniquement décrit chez les moustiques où il est induit par une microsporidie (Hurst 1991; Dunn & Smith 2001). Chez plusieurs espèces de Drosophile, *Spiroplasma* induit la mortalité des embryons mâles via des phénomènes d'apoptoses et de désorganisation neurale lors de l'embryogenèse qui ciblent uniquement les embryons de sexe mâle (Bentley *et al.*, 2007; Martin *et al.*, 2013; Harumoto *et al.*, 2014).

Contrairement à la féminisation où le symbiote ne diminue pas la quantité de descendants d'une femelle infectée et détourne les mâles à son avantage, dans le cas du MK, il y a une réduction de la moitié des descendants et dans ce cas-là les bénéfices pour le symbiote sont un peu moins clairs. Trois avantages en termes de fitness pour le symbiote ont été proposés afin d'expliquer la persistance de cette stratégie (Hurst & Majerus, 1993). Premièrement, la mort des descendants mâles pourrait augmenter la quantité de ressources disponibles pour les sœurs de ces mâles qui portent des symbiotes similaires à ceux-ci. Cette augmentation de ressources peut se faire i) de manière directe comme chez les coccinelles *Adalia bipunctata*, où les cadavres des frères morts à cause du MK induit par *Wolbachia* servent de nourriture aux larves sœurs qui sont cannibales ou ii) de manière indirecte en diminuant la compétition entre les descendants d'une même ponte (Hurst *et al.*, 1993). Le second avantage proviendrait de la diminution du risque de consanguinité. Par exemple, chez *D. willistoni*, *Spiroplasma* permettrait aux sœurs via l'induction de MK d'avoir moins de frères, voire aucun, donc moins de chance de se reproduire avec eux limitant ainsi la consanguinité (Ebbert, 1991). Enfin, la mort des frères pourrait constituer un mécanisme favorisant les transferts horizontaux depuis les cadavres des mâles vers d'autres hôtes potentiels (Hurst & Majerus, 1993).

En fonction du type de MK, ce ne sont pas les mêmes avantages qui seraient procurés aux symbiotes. L'augmentation des ressources et la diminution du risque de consanguinité seraient

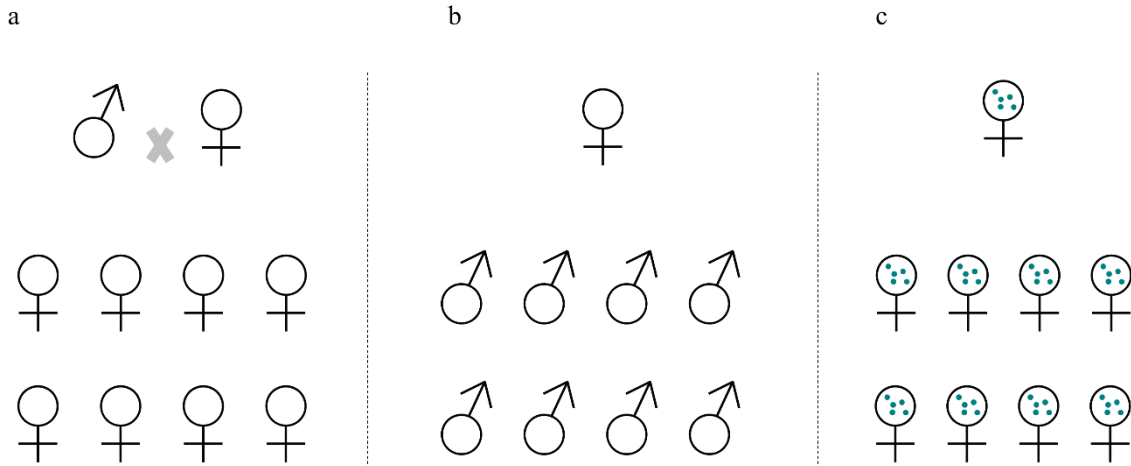


Figure 3 : Parthénogenèse thélytoque chez des organismes haplodiploïdes.

Les points bleus représentent les *Wolbachia*.

a. Dans le cas où mâles et femelles ne sont pas infectés et que les femelles sont fécondées par des mâles, des embryons diploïdes sont produits et donneront des femelles.

b. Lorsque les femelles ne sont pas infectées et non fécondées par des mâles, des embryons haploïdes seront produits et donneront des mâles.

c. Lorsque les femelles sont infectées par *Wolbachia*, les mâles ne sont plus nécessaires pour produire des femelles car *Wolbachia* induit la diploïdisation des œufs.

principalement associés au MK précoce alors que les transferts horizontaux seraient associés au MK tardif (Hurst & Majerus, 1993). Le type de MK observé dépend de l'écologie des hôtes et des symbiotes. Par exemple, du fait de la nature cannibale des larves de coccinelles ou de la limitation de la quantité de ressource pour les descendants des guêpes parasitoïdes, il apparaît qu'entraîner la mort précoce des mâles est clairement avantageux pour la transmission du symbiote (Hurst, 1991b). Inversement, tuer les hôtes mâles à un stade tardif est avantageux pour les microsporidies qui possèdent un mode de transmission mixte. En effet, au stade larvaire durant lequel les mâles meurent, le nombre de spores qu'ils contiennent est maximal, ce qui permet une infection optimale par transmission horizontale (Hurst, 1991b; Dunn & Smith, 2001).

iii. Passage de la parthénogenèse arrhénotoque à la parthénogenèse thélytoque

La troisième stratégie de manipulation de la reproduction utilisée par les symbiotes *Cardinium*, *Rickettsia* et *Wolbachia*, est la sélection vers la parthénogenèse thélytoque (PT) (Arakaki *et al.*, 2001; Provencher *et al.*, 2005; Hagimori *et al.*, 2006). Cette manipulation a, à l'heure actuelle, uniquement été décrite chez des arthropodes haplodiploïdes qui présentent naturellement une parthénogenèse de type arrhénotoque (les mâles sont produits à partir d'ovocytes non fécondés) tels que les acarins (Weeks & Breeuwer, 2001), les guêpes parasitoïdes (Pannebakker *et al.*, 2004) ou encore les thrips (Arakaki *et al.*, 2001). Chez ces espèces hôtes, les symbiotes détournent l'arrhénotoquie naturelle en l'inversant afin que seules des femelles soient produites dans la descendance (Figure 3). Pour ce faire, les embryons non fécondés (haploïdes) devant donner des mâles sont « rendus diploïdes » et produisent ainsi des femelles. Cette diploïdisation peut avoir lieu avant (modification méiotique) ou après la fécondation (modification post-méiotique, voir Cordaux *et al.* 2011 pour revue). Par exemple, chez la guêpe parasitoïde *Neochrysocharsis formosa*, *Rickettsia sp.* induit la parthénogenèse thélytoque lors de l'ovogenèse en empêchant de se produire la deuxième division méiotique (la division réductionnelle permettant la production de gamètes haploïdes), ce qui entraîne la production d'un ovocyte diploïde (Adachi-Hagimori *et al.*, 2008). Ce mécanisme particulier est nommé parthénogenèse apomictique et est aussi utilisé par *Wolbachia* chez l'acarien *Bryobia praetiosa* (Weeks & Breeuwer, 2001). A l'inverse, chez *Leptopilina clavipes*, *Wolbachia* induit la diploïdisation via le phénomène appelé la duplication de gamètes (Pannebakker *et al.*, 2004). Pendant la première division embryonnaire, l'anaphase est avortée ce qui produit un noyau diploïde à la place de deux noyaux haploïdes. Une autre méthode employée par *Wolbachia* pour induire la diploïdisation d'un embryon non fécondé passe par la fusion après la première division embryonnaire des deux noyaux haploïdes en un noyau diploïde comme c'est le cas chez la guêpe parasitoïde *Muscidifurax uniraptor* (Gottlieb *et al.*, 2002). Ce mécanisme est nommé parthénogenèse automictique (Adachi-Hagimori *et al.*, 2008).

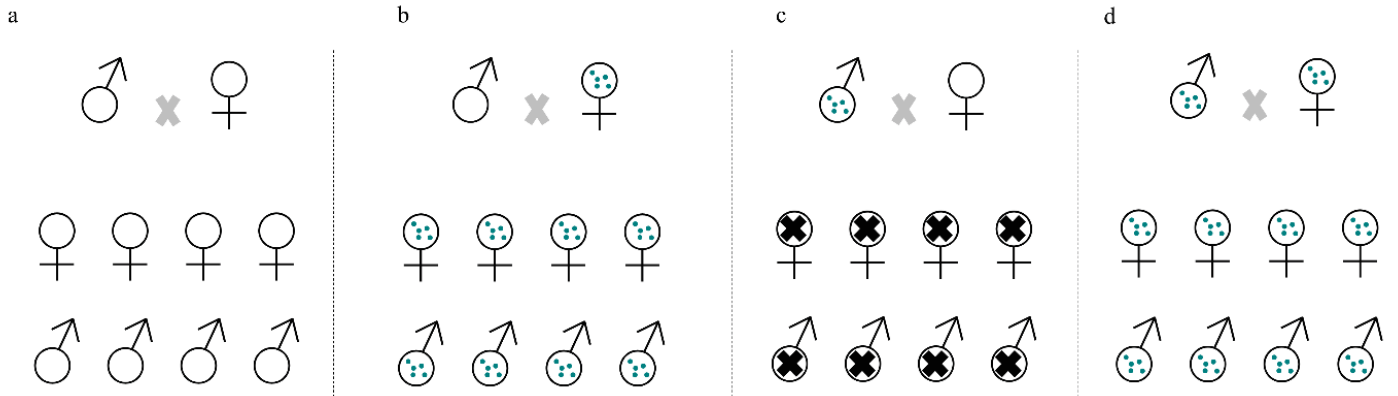


Figure 4 : Incompatibilité cytoplasmique.

Les points bleus représentent les *Wolbachia*.

a. Descendance attendue dans le cas où mâles et femelles ne sont pas infectés par *Wolbachia*.

b. Lorsque les femelles sont infectées par *Wolbachia* mais pas les mâles, la descendance sera viable et elle aussi infectée.

c. Lorsque les mâles sont infectés mais les femelles ne le sont pas, les croisements sont incompatibles et la descendance sera non viable.

d. Lorsque mâles et femelles sont infectés par *Wolbachia* la descendance sera viable.

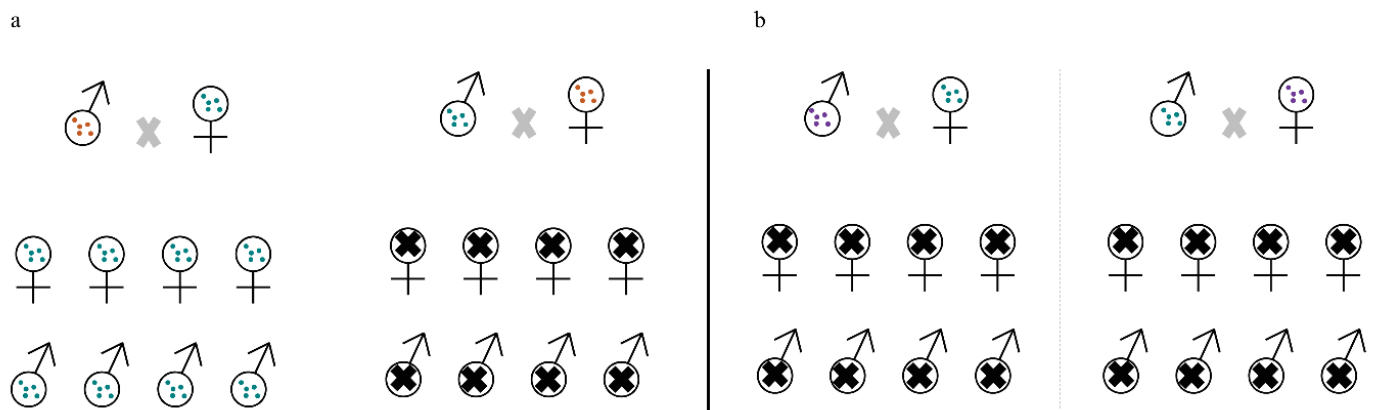


Figure 5 : Incompatibilité cytoplasmique entre individus infectés par des *Wolbachia* différentes et incompatibles.

Les points de couleur représentent différentes souches de *Wolbachia*.

a. IC unidirectionnelle : un sens de croisement est compatible alors que l'autre sens de croisement est incompatible.

b. IC bidirectionnelle : les deux sens de croisement sont incompatibles.

iv. Incompatibilité Cytoplasmique

Une autre stratégie de manipulation de la reproduction (et non des moindres !) sélectionnée chez les parasites de la reproduction s'appelle l'incompatibilité cytoplasmique (IC). A l'heure actuelle, l'IC a été uniquement associée à *Cardinium* et *Wolbachia*, en revanche, la diversité d'hôte chez laquelle ces symbiotes induisent l'IC est très vaste au sein des pancrustacés terrestres, allant des isopodes aux insectes (Coléoptère, Diptère, Hémiptère, Hyménoptère, Lépidoptère, Orthoptère...) en passant par les chélicérates comme les acariens et les pseudoscorpions (Duron *et al.* 2008; voir Werren *et al.* 2008 pour revue). Contrairement aux trois manipulations précédentes, l'IC n'entraîne pas de biais du sex-ratio car il n'y a ni mortalité ni transformation des mâles. L'IC a lieu lorsque des mâles infectés par un symbiote capable d'induire l'IC copulent avec des femelles non infectées et se traduit par la mort des embryons durant l'embryogenèse quel que soit leur sexe (Figure 4) (Callaini *et al.*, 1996). Les femelles infectées peuvent se reproduire et donner une descendance viable avec les mâles infectés et les mâles non infectés. Ainsi, les femelles infectées produiront en moyenne plus de descendants que les femelles non infectées, ce qui permet aux symbiotes de se propager dans les populations hôtes (Rousset & Raymond, 1991). Chez certaines espèces, l'IC peut avoir lieu entre individus infectés par des souches de *Wolbachia* différentes et incompatibles (Figure 5). Deux types de croisements réciproques sont alors observés : soit seul un sens de croisement est compatible et l'autre sens de croisement est incompatible et l'on parlera d'IC unidirectionnelle, soit les deux sens de croisement sont incompatibles et l'on parlera d'IC bidirectionnelle (Figure 5). C'est par exemple le cas chez *Aedes albopictus* où les individus infectés par la souche *wAlbA* sont bidirectionnellement incompatibles avec les individus infectés par la souche *wAlbB* (Dobson *et al.*, 2001), ou encore chez le cloporte *Porcellio dilatatus* où des individus d'une sous espèce infectés par *wDil* sont bidirectionnellement incompatibles avec ceux d'une sous espèce infectés par *wPet* (Sicard *et al.*, 2014a).

Cette stratégie de manipulation sera décrite de façon approfondie dans la suite du manuscrit, celle-ci étant au cœur des investigations qui y sont menées.

c. Les conflits génétiques entre hôtes et manipulateurs

Les gènes des parasites de la reproduction du fait de leur transmission cytoplasmique uniparentale vont favoriser le sexe transmetteur alors que les gènes nucléaires de l'hôte transmis par les deux parents vont favoriser un sex-ratio équilibré (Werren & Beukeboom, 1998). Ainsi, la sélection naturelle s'exerçant sur hôtes et symbiotes ne va pas dans la même direction et les gènes cytoplasmiques et nucléaires s'engagent dans un conflit génétique. Dans le cas du MK de la PT et de la F, les parasites du sexe vont entraîner un biais de sex-ratio en faveur des femelles. Comme le sexe le plus rare dans une population présente toujours le succès reproducteur le plus élevé, produire des femelles plutôt que des mâles dans une population biaisée vers les femelles est coûteux pour l'hôte en terme de transmission de

ses gènes nucléaires (Hamilton, 1967; Turelli, 1994). Fisher (1930) prédit que la sélection du sex-ratio par les gènes nucléaires doit favoriser un sex-ratio de 1:1 dans une population panmictique. Par conséquent, n'importe quel gène nucléaire qui pourrait éliminer l'infection symbiotique ou inhiber sa manipulation favorisant ainsi la production de mâles serait positivement sélectionné dans les populations hôtes. L'intensité du conflit entre hôtes et symbiotes dépend de la fréquence d'infection. Ainsi, plus l'infection est fréquente, plus le sex-ratio sera biaisé en faveur des femelles, plus le coût engendré par la distorsion du sexe ratio sera important et donc plus la sélection pour des gènes de résistance contrant la distorsion chez l'hôte sera importante. Parmi les stratégies de manipulation de la reproduction entraînant une distorsion de sex-ratio, la plus coûteuse est le MK (voir Charlat *et al.* 2003 pour revue). En effet, en plus de la distorsion de sex-ratio, les femelles infectées produisent moitié moins de descendants du fait de la mort de leurs fils que les femelles non-infectées contrairement à la F et la PT où le nombre de descendants reste le même. Par conséquent, pour des niveaux d'infection similaire, on peut prédire que la sélection pour des gènes de résistance, contrariant la manipulation, sera toujours plus forte chez les hôtes où les symbiotes induisent le MK que chez les hôtes où les symbiotes induisent la parthénogenèse thélytoque ou la féminisation (voir Charlat *et al.* 2003 pour revue).

Ces gènes de résistance sont appelés des restaurateurs nucléaires de sex-ratio et ils sont sélectionnés chez des hôtes où l'on observe les trois types de manipulation MK, F et PT (voir Cordaux *et al.* 2011 pour revue). Ce type de gène existe, comme prédit, dans le cas de résistance au MK par exemple chez *Hypolimnas bolina* où un gène de résistance autosomal dans certaines populations de guêpes parasitoïdes a été sélectionné car il inhibe tout effet de MK lié à la *Wolbachia* (Hornett *et al.*, 2006).

Un exemple de résistance à la féminisation par *Wolbachia* est la sélection du gène de résistance R chez le cloporte *A. vulgare* qui empêche la transmission du symbiote à la descendance et donc la transformation au cours du développement des fils en filles (Rigaud & Juchault, 1992). Chez la guêpe parasitoïde *Trichogramma kaykai* où *Wolbachia* induit la parthénogenèse thélytoque, la sélection positive dans certaines populations hôtes d'un chromosome surnuméraire nommé PSR (paternal sex-ratio chromosome) a été observée (Stouthamer *et al.*, 2001; van Vugt *et al.*, 2003). Ce chromosome lorsqu'il est présent dans les populations de *T. kaykai* code pour un déterminisme sexuel mâle et maintient la fréquence de *Wolbachia* assez basse dans les populations. Ainsi, la population hôte n'est jamais totalement parthénogénétique. A cette fin, le chromosome PSR converti les œufs diploïdes fécondés en œufs haploïdes qui donneront des mâles non transmetteur de la bactérie. Le mécanisme consiste à détruire le jeu de chromosome paternel en entraînant sa surcondensation après la fécondation les empêchant ainsi de prendre part à la première division embryonnaire (van Vugt *et al.*, 2003).

Les conflits génétiques associés à l'IC seront quant à eux développés plus loin dans le manuscrit.

d. Les conséquences évolutives du parasitisme sexuel

Du fait de l'interaction très intime entre un manipulateur et son hôte, le parasitisme du sexe va avoir des conséquences évolutives fortes sur ses populations hôtes.

i. Modification du déterminisme du sexe

De par leur impact sur la reproduction, les symbiotes parasites du sexe peuvent avoir des conséquences sur le déterminisme génétique du sexe. C'est le cas par exemple chez le cloporte *A. vulgare* où la détermination du sexe peut être codée par des gènes nucléaires, les femelles étant hétérogamétiques (ZW) et les mâles homogamétiques (ZZ). Cependant, en présence de *Wolbachia* qui induit la féminisation chez cette espèce, le sex-ratio est biaisé en faveur des femelles car les mâles ZZ de la descendance sont transformés en femelles qui vont à leur tour produire une descendance biaisée vers les femelles (Rigaud, 1997). Ainsi, les mâles féminisés ZZ produisent des femelles ne portant pas le chromosome W qui ne pourront donc pas le transmettre à leur descendance. Par conséquent, la fréquence du chromosome W diminue à chaque génération pour disparaître totalement dans certaines populations (voir Rigaud *et al.* 1997b pour revue). Dans de telles populations où le chromosome W a disparu, tous les individus sont des mâles génétiques ZZ ; ceux qui hériteront de *Wolbachia* se développeront en femelles alors que les autres se développeront en mâles. Le déterminisme du sexe dans de telles populations est alors sous le total contrôle de *Wolbachia* et constitue un parfait exemple de déterminisme cytoplasmique du sexe.

Chez cette espèce, il existe des populations dont le sex-ratio est biaisé en faveur des femelles mais où *Wolbachia* n'est pas présente. Ces femelles sont en fait des mâles génétiques ZZ convertis en femelle par la présence de l'élément f qui est un morceau du génome d'environ 3-Mb de la bactérie féminisante *Wolbachia wVulC* qui s'est inséré dans le génome de l'hôte (Legrand & Juchault, 1984; Leclercq *et al.*, 2016). Cet élément f provient d'un transfert horizontal entre le génome du symbiote et de l'hôte. Ces transferts horizontaux de gènes entre bactéries et animaux peuvent être ainsi considérés comme une nouvelle origine potentielle à l'apparition de chromosome sexuel. Les symbiotes distorateurs du sex-ratio peuvent donc être de puissants moteurs de l'évolution du déterminisme du sexe et de sa transition d'un déterminisme à un autre chez les animaux.

ii. Modification des patrons de sélection sexuelle et de choix du partenaire

La théorie de la sélection sexuelle propose que les mâles soient en compétition pour l'accès aux femelles car ils augmentent leur succès reproducteur au travers d'acte de copulations multiples (Bateman, 1948). A l'inverse, le succès reproducteur des femelles dépendrait surtout de la qualité des mâles avec lesquels elles s'accouplent plutôt que de la quantité de partenaires. Dans le cas d'une

population où le sex-ratio est biaisé en faveur des femelles à cause de la présence d'un manipulateur de la reproduction, les mâles vont devenir rares. Quand les mâles sont plus rares, la compétition entre eux pour l'accès aux femelles va diminuer et de la compétition entre femelles pour l'accès aux mâles peut survenir. Ainsi, on peut voir une inversion des comportements sexuels entre mâles et femelles dans de telles populations. C'est par exemple le cas chez l'espèce de papillon *Acrea encefon* où *Wolbachia* induit le MK (Jiggins *et al.*, 2000). Chez cette espèce, le niveau d'infection des populations est très fort ce qui entraîne un sex-ratio très biaisé en faveur des femelles. Un très grand nombre de femelles vierges a été observé, ce qui suggère que les mâles constituent dans ce cas un facteur limitant à la reproduction. Il a été constaté que les femelles ont développé un comportement de reproduction qui n'est habituellement observé que chez les mâles : elles se rassemblent et copulent très rapidement avec un mâle (il n'y a plus de temps accordé au choix du partenaire).

Selon la théorie de la sélection sexuelle, le choix du partenaire permet d'améliorer sa valeur sélective et la qualité de sa descendance. Une infection parasitaire, qu'elle soit due à une bactérie ou à un agent viral, impose un coût à son hôte et peut résulter en la transmission du parasite à la descendance. On peut donc s'attendre à ce que des mécanismes d'évitement des individus infectés soient sélectionnés. Dans le cas de l'infection par des manipulateurs de la reproduction, il est prédit que le choix du partenaire se porterait sur des individus non infectés ou dans le cas particulier de l'IC sur un partenaire possédant le même statut d'infection (voir Beltran-Bech & Richard 2014 pour revue). En effet, il a été montré chez l'acarien *Tetranychus urticae* chez lequel *Wolbachia* induit l'IC que les femelles non infectées préfèrent s'accoupler avec des mâles eux aussi non infectés (Vala *et al.*, 2004). En revanche, lorsque le coût du choix du partenaire par évitement de l'infection n'est pas compensé par le bénéfice de choisir un partenaire non infecté, les mécanismes d'évitement de l'infection ne vont pas être sélectionnés. Ainsi, chez *Drosophila innubila* chez laquelle *Wolbachia* induit le MK, l'absence de choix des femelles par les mâles pourrait résulter de la fluctuation de la prévalence de l'infection dans les populations (Sullivan & Jaenike, 2006). En effet, dans une population où la prévalence de l'infection est élevée, il est avantageux de pouvoir discriminer les femelles infectées des femelles non infectées beaucoup plus rares. En revanche, si la prévalence de l'infection vient à chuter, le coût du choix des mâles surpasse les bénéfices. Ainsi, dans un environnement changeant lorsque la prévalence de l'infection fluctue fortement, les mécanismes certes adaptatifs mais coûteux de choix des mâles peuvent ne pas être sélectionnés.

iii. Evolution vers l'asexualité

Les symbiotes parasites de la reproduction capables d'induire la PT chez les espèces hôtes haplodiploïdes peuvent entraîner la disparition du sexe dans les populations hôtes. En effet, les femelles infectées, peuvent produire des femelles en l'absence de mâles. Chez certaines espèces hôtes comme les

guêpes parasitoïdes du genre *Trichogramma*, les mâles ont disparu de certaines populations et ne peuvent être obtenus artificiellement qu'après un traitement aux antibiotiques permettant d'éliminer l'infection. Par conséquent, certaines populations de cette espèce se maintiennent uniquement grâce à la PT induite par le symbiote et on observe la perte de la sexualité (Stouthamer *et al.*, 1990). Pour qu'un symbiote seul soit responsable d'une perte de la reproduction sexuelle de la population dans laquelle il induit la PT, il doit être transmis à la descendance quasiment parfaitement (~100% des œufs infectés) et doit être capable de transformer complètement les embryons non fécondés haploïdes en embryons diploïdes donnant des femelles (voir Hurst & Frost 2015 pour revue). Puisque les symbiotes montrent souvent des problèmes de transmission ou des capacités de manipulation imparfaites, la transition vers l'asexualité obligatoire est souvent due à une évolution de l'hôte associée à la propagation du symbiote. Une hypothèse a été proposée qui stipule que la propagation des symbiotes dans les populations génération après génération favoriserait la perte de la fonction sexuelle chez les femelles (Stouthamer *et al.*, 2010). Dans les populations hôtes où le sex-ratio est biaisé en faveur des femelles, produire des mâles est avantageux. Chez les haplodiploïdes puisque les mâles sont obtenus à partir d'œufs haploïdes (non fécondés), rester vierge est avantageux pour une femelle non infectée. Ainsi, la sélection peut favoriser la production de fils chez ces femelles via l'apparition de « gènes de virginité ». Dans de tels cas, les femelles non infectées produiraient quasiment uniquement des mâles. Ainsi, les femelles non infectées vont en moyenne produire moins de filles que les femelles infectées, ce qui renforcera la propagation du symbiote dans les populations jusqu'à sa fixation et la perte de toute reproduction sexuelle (Stouthamer *et al.*, 2010). De telles prédictions théoriques ont été empiriquement observées dans des populations naturelles de *Apoanagyrus diversicornis*, *Asobara japonica* et *Leptopilina clavipes* où les femelles dérivées de lignées complètement asexuées n'étaient plus capables d'utiliser le sperme des mâles qui les inséminaient (Pijls *et al.*, 1996; Pannebakker *et al.*, 2005; Kremer *et al.*, 2009).

iv. Extinction

Les parasites manipulateurs du sexe pourraient dans des cas extrêmes entraîner l'extinction de leurs populations hôtes en diminuant la productivité d'une population (le nombre de descendants produit à chaque génération). L'extinction pourrait par exemple survenir chez les espèces diploïdes où l'on observe la F ou le MK avec un taux d'infection très fort responsable d'un sex-ratio très fortement biaisé en faveur des femelles (Ferdy *et al.*, 2016). Dans de telles populations, les mâles deviendraient tellement rares que la reproduction ne serait plus suffisante pour maintenir la viabilité de la population. Une telle situation a été observée chez *A. encefalon* où un grand nombre de femelles n'est pas fécondé dans les populations où le symbiote est présent (Jiggins *et al.*, 2000). De plus, sur le long terme, les symbiotes pourraient aussi augmenter les risques d'extinction en réduisant la diversité génétique de la population. En effet, la taille efficace d'une population où le sex-ratio est fortement biaisé en faveur des femelles est génétiquement parlant proche de celle du sexe le plus rare, *i.e.* le sexe mâle (Wright 1932; voir

Charlat *et al.* 2003 pour revue). Ainsi, lorsque le sex-ratio est très fortement biaisé, peu de mâles sont présents dans les populations et la diversité génétique est fortement réduite. Enfin, dans les populations où les symbiotes induisent la PT, l'absence de sexualité pourrait conduire à l'accumulation de mutations délétères (cliquet de Muller) (Maynard, 1978).

III. *Wolbachia* reine de la manipulation

Parmi tous les symbiotes capables de manipuler la reproduction de leur hôte, les *Wolbachia* occupent une place particulière. En effet, ce sont les seuls symbiotes chez lesquels ont été observées toutes les stratégies de manipulation de la reproduction présentées ci-dessus. De plus, *Wolbachia* infecte la plus grande diversité d'espèces d'arthropodes car il a été estimé que plus de 40% des espèces d'arthropodes seraient potentiellement infectées (Zug & Hammerstein, 2012). Si ces chiffres représentent la réalité, l'infection des arthropodes par *Wolbachia* pourrait constituer une des plus grandes pandémies de la planète (voir LePage & Bordenstein 2013 et Sicard *et al.* 2014b pour revue). Enfin, ce qui rend *Wolbachia* encore plus intéressante c'est, qu'en plus d'être un parasite de la reproduction, ce symbiote est aussi impliqué dans des interactions mutualistes avec ses hôtes.

a. *Wolbachia* l'endosymbiote le plus répandu chez les arthropodes

Les *Wolbachia* sont des alpha-protéobactéries de l'ordre des *Rickettsiales*, un groupe très divers de bactéries intracellulaires qui comprend par exemple les bactéries du genre *Anaplasma*, *Ehrlichia* et *Rickettsia* connues pour être des parasites de vertébrés entraînant par exemple le typhus chez l'Homme (*Rickettsia prowazekii*) (voir Werren *et al.* 2008 pour revue; Valbuena & Walker 2009). Cependant, contrairement aux autres genres, qui infectent les vertébrés, les bactéries du genre *Wolbachia* infectent uniquement des ecdysozoaires (arthropodes et des nématodes).

L'espèce *Wolbachia pipientis* (Hertig, 1936) a été décrite pour la première fois chez le moustique *Culex pipiens* (Hertig & Wolbach, 1924). Elle a ensuite été décrite chez de nombreuses espèces appartenant à deux grands clades au sein des ecdysozoaires : les Nématodes filaires et les Arthropodes. Chez les Arthropodes, *Wolbachia* infecte principalement trois grands clades : les insectes, les isopodes terrestres et les arachnides. Parmi les Arachnides on retrouve *Wolbachia* dans les groupes des Acariens (Breeuwer & Jacobs, 1996), des Araignées (Rowley *et al.*, 2004) et des Pseudoscorpions (Zeh *et al.*, 2005). Enfin, c'est chez les insectes que l'infection par *Wolbachia* est la plus répandue car on la retrouve chez les Diptères, les Lépidoptères, les Hémiptères, les Hyménoptères, les Coléoptères, les Orthoptères, les Siphonaptère, les Isoptères, les Phthiraptère, les Névroptères, les Blattoptères, ... (voir Werren *et al.* 2008 pour revue). On estime à l'heure actuelle que plus de 65% des espèces d'insectes

sont infectés par *Wolbachia* (Hilgenboecker *et al.*, 2008). Par ailleurs, *Wolbachia* a aussi été détectée chez les collemboles, groupe frère des insectes (Vandekerckhove *et al.*, 1999; Ma *et al.*, 2017).

b. Diversification des *Wolbachia* : la conséquence de transferts horizontaux ?

Le maintien et la large prévalence des *Wolbachia* dans les populations d'arthropodes sont essentiellement liés à leur capacité à manipuler la reproduction de leurs hôtes. Cette capacité a été sélectionnée chez ces symbiotes du fait de leur transmission quasiment systématiquement verticale. Cette transmission verticale peut expliquer le maintien du symbiote dans une espèce donnée mais ne peut, en revanche, pas expliquer le large spectre d'hôtes des *Wolbachia*. La seule explication possible réside dans des transferts infectieux entre espèces différentes. Ces transferts horizontaux entre espèces ont plus probablement lieu, dans la nature, lorsque les hôtes présentent des liens écologiques étroits, comme par exemple entre un hôte et son parasite, une proie et son prédateur (Cordaux *et al.*, 2001), ou encore par le partage d'une même plante hôte par deux espèces de phytophages (Sintupachee *et al.*, 2006), le cannibalisme/nécrophagie (Le Clec'h *et al.*, 2013a) ou même par contact sanguin suite à une blessure (Rigaud & Juchault, 1995). De nombreux transferts expérimentaux ont pu être réalisés de manière artificielle via des injections (Grenier *et al.*, 1998; Rigaud *et al.*, 2001; Riegler *et al.*, 2004; Blagrove *et al.*, 2012, 2013; Le Clec'h *et al.*, 2013b) et dans certains cas, l'infection peut se maintenir de génération en génération (Rigaud *et al.*, 2001). Les expériences utilisant des voies écologiques sont plus rares et peu de cas de transferts horizontaux interspécifiques de ce type (*i.e.* sans injection du symbiote) ont été décrits. De plus, dans la plupart des cas, l'infection par *Wolbachia* disparaît durant les quelques générations suivant le transfert horizontal (Heath *et al.*, 1999; Huigens *et al.*, 2004). Même si ces transferts horizontaux n'ont pas ou peu été « mimés » expérimentalement, de nombreuses études phylogénétiques sur des espèces variées montrent l'absence de co-diversification entre hôtes et *Wolbachia* traduisant les réalités des transferts horizontaux à l'échelle évolutive (Vavre *et al.*, 1999; Timmermans *et al.*, 2004; Werren *et al.*, 2008; Raychoudhury *et al.*, 2009; Watanabe *et al.*, 2012; Ahmed *et al.*, 2016). Des exemples marquants illustrent l'acquisition par transferts horizontaux interspécifiques. Ainsi, *C. pipiens* qui est un diptère et *H. bolina* qui est un lépidoptère sont infectés par des *Wolbachia* très proches phylogénétiquement (Bleidorn & Gerth, 2018). A l'opposé, deux espèces hôtes très proches comme *D. simulans* et *D. melanogaster* sont naturellement infectées par des souches de *Wolbachia* appartenant à des super-groupes phylogénétiques différents (Paraskevopoulos *et al.*, 2006). De plus, une même espèce hôte peut être infectée par des souches de *Wolbachia* appartenant à des super-groupes différents comme c'est le cas chez le moustique tigre *A. albopictus* (Sinkins *et al.*, 1995; Dobson *et al.*, 2001). De telles observations chez les arthropodes ne peuvent être expliquées que par l'occurrence de transferts horizontaux au cours de l'histoire évolutive de ces interactions. En revanche, chez les nématodes, la forte congruence entre les phylogénies des hôtes et celle des souches

Tableau 1 : Espèces hôtes et super-groupes des souches de *Wolbachia* qui les infectent.

Super-groupe de <i>Wolbachia</i>	Clade des espèces hôtes	Exemple d'espèces hôtes	Références
A	Arthropodes	<i>Drosophila melanogaster</i> <i>D. ananassae</i> <i>D. simulans</i> <i>Nasonia longicornis</i> <i>N. vitripennis</i> <i>Aedes albopictus</i> <i>Muscidifurax uniraptor</i>	(Lo <i>et al.</i> , 2002) (Baldo <i>et al.</i> , 2006) (Ros <i>et al.</i> , 2009)
B	Arthropodes	<i>D. simulans</i> <i>Armadilidium vulgare</i> <i>Culex pipiens</i> <i>Aedes albopictus</i> <i>Hypolimnas bolina</i> <i>Diaphorina citri</i> <i>Encarsia formosa</i> <i>Trichogramma brevicornis</i>	(Lo <i>et al.</i> , 2002) (Baldo <i>et al.</i> , 2006) (Ros <i>et al.</i> , 2009)
C	Nématodes filaires	<i>Onchocerca volvulus</i> <i>O. gutturosa</i> <i>O. ochengi</i> <i>Dirofilaria immitis</i> <i>D. repens</i>	(Lo <i>et al.</i> , 2002) (Taylor <i>et al.</i> , 2013) (Lefoulon <i>et al.</i> , 2016)
D	Nématodes filaires	<i>Brugia malayi</i> <i>B. pahangi</i> <i>Litomosoides sigmodontis</i>	(Lo <i>et al.</i> , 2002) (Taylor <i>et al.</i> , 2013) (Lefoulon <i>et al.</i> , 2016)
E	Collemboles	<i>Folsomia candida</i> <i>Mesaphorura italica</i>	(Ma <i>et al.</i> , 2017)
F	Nématodes filaires Arthropodes	<i>Syringophilopsis turdi</i> <i>Torotroglia rubeculi</i> <i>Nasutitermes nigriceps</i> <i>coptotermes acinaciformis</i> <i>Microcerotermes sp.</i> <i>Kalotermes flavicollis</i>	(Lo <i>et al.</i> , 2002) (Casiraghi <i>et al.</i> , 2005) (Baldo <i>et al.</i> , 2007) (Covacin & Barker, 2007) (Panaram & Marshall, 2007), (Lefoulon <i>et al.</i> , 2012)
(G)	Araignées Australiennes	<i>Pegoscopus gemellis</i> <i>Dysdera erythrina</i> <i>Diaeacircumlita</i>	(Rowley <i>et al.</i> , 2004)
H	termites	genre <i>Zootermopsis</i>	(Bordenstein & Rosengaus, 2005)
I	puces	<i>Ctenocephalides felis</i> <i>Orchopeas leucopus</i>	(Ros <i>et al.</i> , 2009)
J	Nématodes filaires	<i>Dipetalonema gracile</i>	(Ros <i>et al.</i> , 2009)
K	Acarien	genre <i>Bryobia</i>	(Ros <i>et al.</i> , 2009)
L	Nématode parasite de plante	<i>Radopholus similis</i>	(Haegeman <i>et al.</i> , 2009) (Głowska <i>et al.</i> , 2015)
M	Puceron	<i>Brevicoryne brassicae</i> <i>Aphis fabae</i>	(Wang <i>et al.</i> , 2014) (Augustinos <i>et al.</i> , 2011)
N	Puceron	<i>Pucerons</i>	(Wang <i>et al.</i> , 2014) (Augustinos <i>et al.</i> , 2011)
O	Mouche blanche	<i>Bemisia tabaci</i>	(Nirgianaki <i>et al.</i> , 2003)
P	Acariens	<i>Torotroglia merulae</i> <i>T. lusciniiae</i> <i>Syringophilopsis turdi</i>	(Głowska <i>et al.</i> , 2015)
Q	Acariens	<i>Torotroglia cardueli</i>	(Głowska <i>et al.</i> , 2015)

de *Wolbachia* indique la présence de co-spéciation et traduit la rareté des transferts horizontaux entre hôtes dans ce clade (Casiraghi *et al.*, 2001).

Grâce entre autres à ces transferts horizontaux, les *Wolbachia* sont à l'heure actuelle diversifiées en 16 super-groupes allant de A à F et de H à Q (Baldo *et al.*, 2006; Paraskevopoulos *et al.*, 2006; Bordenstein *et al.*, 2009; Ros *et al.*, 2009; Glowska *et al.*, 2015). Le super-groupe G contient uniquement des *Wolbachia* infectant des araignées australiennes (Rowley *et al.*, 2004), mais certains auteurs pensent que les *Wolbachia* de ce super-groupe seraient en fait des recombinants entre des souches des groupes A et B plutôt qu'un groupe distinct. Les classifications de ces différentes souches de *Wolbachia* dans les 16 super-groupes (Tableau 1, Figure 6) ont été permises grâce au développement d'un outil appelé MLST pour « Multi Locus Sequence Typing » qui contient plusieurs gènes permettant de discriminer les souches de *Wolbachia* (Baldo *et al.*, 2006; Paraskevopoulos *et al.*, 2006). Si les approches MLST semblent bien refléter les relations phylogénétiques entre souches de *Wolbachia* ayant divergé il y a longtemps, elles manquent de puissance pour discriminer les *Wolbachia* ayant récemment divergé (Bleidorn & Gerth, 2018).

c. Les génomes des *Wolbachia* : des génomes très dynamiques

i. Impact du mode de vie intracellulaire sur la taille et l'évolution des génomes

Du fait de leur mode de vie intracellulaire, les *Wolbachia* sont restreintes au cytoplasme des cellules de leur hôte et par conséquent possèdent de faible taille efficace de population. La transmission verticale résulte à chaque génération en un goulot d'étranglement dans la population de *Wolbachia* infectant un hôte. En effet, parmi toutes les *Wolbachia* qui infectent un hôte seules celles présentes dans les ovocytes vont se transmettre. Ces caractéristiques associées à leur reproduction asexuée, font que la sélection purificatrice, qui normalement purge les mutations faiblement délétères, soit beaucoup moins efficace chez ces bactéries. Par conséquent, les génomes des *Wolbachia* sont sujets à ce que l'on appelle le cliquet de Muller qui se traduit par l'accumulation de mutations faiblement délétères dans les génomes pouvant entraîner la pseudogénéisation sans possibilité de purification efficace (Moran, 1996). La conséquence d'un tel phénomène va être la disparition de gènes non essentiels à la bactérie (e.g des gènes impliqués dans la vie extracellulaire ou en redondance métabolique avec la cellule de l'hôte) et *in fine* une diminution de la taille des génomes. En effet, comparés aux génomes d'alpha-protéobactéries à vie libre dont les tailles sont comprises entre 3,8Mb et 9,7Mb (Teyssier *et al.*, 2004), ceux des *Wolbachia* ont des tailles de génome comprises entre 0,96Mb pour le génome de *wOo* infectant le nématode *Onchocerca ochengi* (Darby *et al.*, 2012) et 2,1Mb pour le génome de *wCon* infectant le cloporte *Cylisticus convexus* (Badawi *et al.*, 2018).

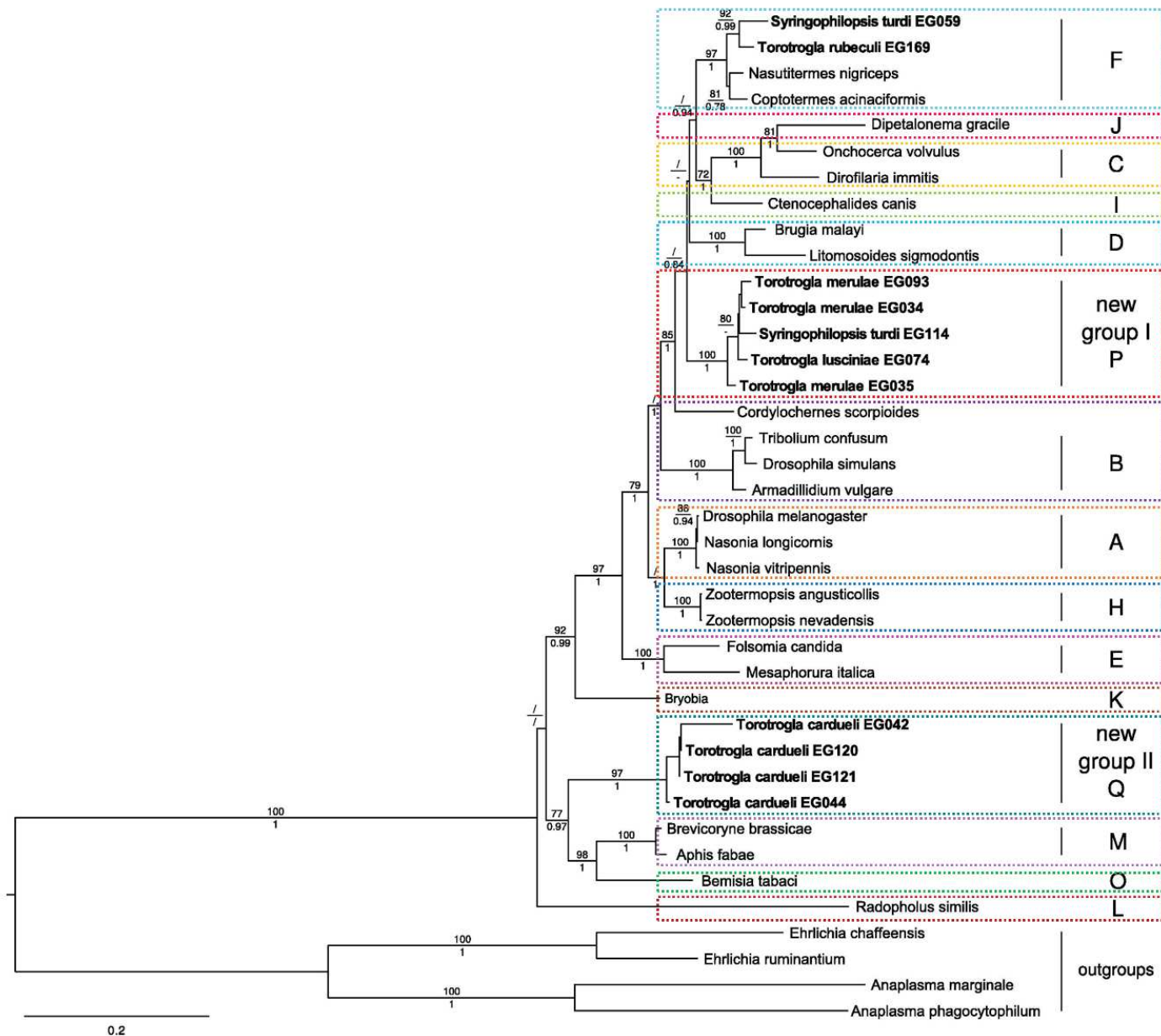


Figure 6 : Phylogénie des super-groupes de *Wolbachia* construite à partir d'une super-matrice composée des 5 gènes 16S, rRNA, *coxA*, *ftsZ*, *gltA*, *groEL* par la méthode de Maximum likelihood.

Les noms d'espèces représentent les hôtes infectés par les différentes souches de *Wolbachia*. Les numéros représentent les valeurs de bootstrap.
Figure adaptée de Glowska et al. (2015)

Une autre caractéristique des génomes de *Wolbachia* et des bactéries endocytobiotiques en général est un biais de composition nucléotidique vers les bases AT (McCutcheon & Moran, 2012). Ce biais de composition serait favorisé par la présence d'un biais de mutation de (G ou C) vers (A ou T) qui a été montré chez les bactéries (Hershberg & Petrov, 2010; Hildebrand *et al.*, 2010). Ce biais de composition serait aussi favorisé par la perte dans les génomes des endosymbiotes des gènes impliqués dans la réparation de l'ADN (Moran *et al.*, 2008). En effet, il a été montré que sans mécanisme de réparation de l'ADN, les altérations de celui-ci comme les déaminations des cytosines ou les oxydations des guanines entraînent souvent des changements de bases de (G ou C) vers (A ou T) (Lind & Andersson, 2008). Ainsi, les génomes des souches de *Wolbachia* *w*Pip infectant *C. pipiens*, *w*Mel infectant *D. melanogaster* et *w*Bm infectant *Brugia malayi* possèdent des taux de GC entre 34.2 et 35.2% (Klasson *et al.*, 2008), ce qui est significativement inférieur aux génomes de bactérie libre qui ont en moyenne un taux de GC de 49% (Rocha & Danchin, 2002).

Par ailleurs, la perte de ces mécanismes de réparation de l'ADN a aussi pour conséquence une augmentation du nombre de mutations qui ne vont pas être corrigées, ce qui va en plus de la relaxation de la sélection et du cliquet de Muller entraîner une évolution rapide des génomes (McCutcheon & Moran, 2012).

ii. Des génomes riches en éléments génétiques répétés mobiles et en protéines à domaine ankyrine

Même si les génomes des *Wolbachia* montrent des tailles réduites comparées aux bactéries libres proches phylogénétiquement, leurs génomes sont tout de même conséquents par rapport aux endosymbiotes mutualistes comme *Sulcia muelleri* ou la souche ASP du puceron *B. aphidicola* qui possèdent des génomes respectivement de 0,246 et 0,641 Mb (Moran *et al.*, 2008; McCutcheon & Moran, 2012). Cette différence de taille est majoritairement associée à la présence, dans leurs génomes, de nombreux éléments génétiques mobiles présents en plusieurs copies (éléments répétés) non présents dans les génomes des symbiotes primaires (*i.e.* symbiote essentiel à la survie de l'hôte) (Wu *et al.*, 2004; Klasson *et al.*, 2008, 2009; Salzberg *et al.*, 2009).

Trois grandes classes d'éléments répétés mobiles ont été identifiées dans les génomes des *Wolbachia* : les séquences d'insertions (IS), les rétrotransposons et les prophages (Wu *et al.*, 2004; Bordenstein & Reznikoff, 2005; Klasson *et al.*, 2008; Cerveau *et al.*, 2011). Les IS et les rétrotransposons sont rangés dans la catégorie des éléments transposables (Bordenstein & Reznikoff, 2005). Ce sont des séquences d'ADN capables de s'exciser d'une région du génome et de s'intégrer ailleurs dans le génome, voire dans un autre génome. Les IS se déplacent dans les génomes via l'utilisation de transposases et restent toujours sous forme d'ADN alors que les rétrotransposons utilisent des transcriptases reverses et alternent entre des phases ADN et ARN (Bordenstein & Reznikoff, 2005).

La quantité d'IS est variable selon les génomes de *Wolbachia* et les souches de *Wolbachia* mutualistes en possèdent moins que les souches *Wolbachia* manipulatrices de la reproduction (Cerveau *et al.*, 2011). En effet, on en décompte 52 dans le génome de *wBm* la *Wolbachia* mutualiste qui infecte le filaire *B. malayi* contre 105 dans le génome de *wMel*, 170 dans le génome de *wPip* et 171 dans celui de *wRi* (Cerveau *et al.*, 2011). Ce nombre d'IS est considérable, sachant que les génomes des procaryotes possèdent généralement très peu d'IS. En effet, l'étude de 262 génomes a révélé que le nombre médian d'IS dans les bactéries libres est de 12 copies pour un nombre allant de 0 à 342 copies (Touchon & Rocha, 2007).

Les séquences de prophages présentes dans les génomes des *Wolbachia* proviennent de l'insertion multiple du phage WO (Masui *et al.*, 2000, 2001). Ce phage a d'abord été découvert sous forme de particules virales dans les *Wolbachia* infectant le moustique *C. pipiens* (Wright *et al.*, 1978). L'insertion de son génome dans celui de *Wolbachia* n'a été mise en évidence que bien plus tard chez les souches de *Wolbachia wTai* infectant *Teleogryllus taiwanemma* (Masui *et al.*, 2000) et *wKue* infectant *Ephestia kuehniella* (Masui *et al.*, 2001). Ainsi, ce prophage est soit à l'état lysogénique, intégré dans le génome de *Wolbachia*, soit à l'état lytique et libre dans le cytoplasme de la bactérie (Masui *et al.*, 2001). Ce prophage a par la suite été détecté dans de très nombreux génomes de *Wolbachia* (Bordenstein & Wernegreen, 2004; Gavotte *et al.*, 2007) infectant des isopodes terrestres (Braquart-Varnier *et al.*, 2005), des insectes comme les drosophiles, les guêpes parasitoïdes ou encore les moustiques (Gavotte *et al.*, 2004; Klasson *et al.*, 2008), des acariens (Lu *et al.*, 2012). En fonction des souches de *Wolbachia*, les séquences de prophages ne sont pas les mêmes et le nombre de copies varie. Dans le génome de *wMel*, deux familles de prophages WO ont été détectées, le prophage WO-A et le prophage WO-B, ce dernier étant plus proche génétiquement des prophages détectés chez *wKue* que le prophage WO-A (Wu *et al.*, 2004). Chez *wPip*, la *Wolbachia* infectant *C. pipiens*, 5 copies des prophages WO ont été détectées toutes les 5 plus proches du prophage WO-B que du prophage WO-A de *wMel* (Klasson *et al.*, 2008). Ces 5 copies sont en revanche plus proches entre elles que du prophage WO-B de *wMel* suggérant des duplications postérieures à la séparation entre les souches *wPip* et *wMel* (Klasson *et al.*, 2008). De même chez *wRi*, 4 séquences de prophages ont été détectées WO-A, WO-B (en deux copies identiques) et WO-C (Klasson *et al.*, 2009).

En plus des éléments génétiques mobiles, les génomes de *Wolbachia* présentent des gènes codant pour des protéines à domaine Ankyrine (ANK) (Wu *et al.*, 2004; Foster *et al.*, 2005; Iturbe-Ormaetxe *et al.*, 2005; Walker *et al.*, 2007; Klasson *et al.*, 2008). Ces protéines sont communes chez les eucaryotes mais peu fréquentes chez les bactéries chez lesquelles elles jouent un rôle dans la médiation des interactions entre les protéines de l'hôte et celles de la bactérie (Wu *et al.*, 2004). Chez *wPip*, 56 gènes présentent un ou plusieurs motifs ANK contre 35 chez *wRi* et 23 chez *wMel*. Cela fait de *wPip* la bactérie qui contient le plus grand nombre de protéines à domaine ankyrine parmi toutes les bactéries

pour lesquelles l'information est disponible à l'heure actuelle (Wu *et al.*, 2004; Klasson *et al.*, 2008, 2009).

iii. Implications des éléments génétiques mobiles et des protéines à domaine ankyrine dans l'évolution des génomes

La présence de ces éléments génétiques mobiles et de ces protéines à domaine ankyrine dans les génomes des souches de *Wolbachia*, va avoir des conséquences sur la composition des génomes mais aussi potentiellement dans les phénotypes de manipulation qu'elles induisent chez leurs hôtes.

Ces éléments génétiques mobiles peuvent dans certains cas s'insérer à l'intérieur d'une séquence codante. Par exemple dans les génomes de *wPip* et *wRi*, les *Wolbachia* infectant *C. pipiens* et *D. simulans*, les séquences de 44 et 46 gènes respectivement présentent des insertions d'IS (Klasson *et al.*, 2008, 2009). De telles perturbations de séquences rendent les gènes non fonctionnels, ce qui en plus de l'accumulation de mutations faiblement délétères, est une source de pseudogénéisation suivie par la purge de gènes.

Ces éléments mobiles et séquences répétées sont, par ailleurs, très probablement à l'origine de très nombreux réarrangement intra et inter-chromosomiques observés chez les *Wolbachia*. En effet, l'introduction et la propagation d'éléments transposables et de prophages conduit à l'apparition de sites de recombinaisons homologues promoteurs de réarrangements chromosomiques (Klasson *et al.*, 2008). De tels réarrangements contribuent à la diversification rapide des *Wolbachia*. De plus, il a été montré que les phages sont capables de se déplacer entre génomes de *Wolbachia* via des transferts horizontaux et de transporter avec eux de nouveaux gènes permettant un apport de diversité génétique dans les génomes de *Wolbachia* (Bordenstein & Wernegreen, 2004; Gavotte *et al.*, 2004; Chafee *et al.*, 2010; Kent *et al.*, 2011b; a, Wang *et al.*, 2016a; b).

Le maintien d'une richesse importante en éléments mobiles et séquences répétées dans les génomes des *Wolbachia*, en particulier chez *wMel*, *wPip* et *wRi* pourrait s'expliquer par une sélection purificatrice relaxée comme observée dans les génomes d'autres endosymbiotes (Wu *et al.*, 2004). Une autre hypothèse fait intervenir un avantage à leur présence. Du fait de la multiplication asexuée des procaryotes qui tend à diminuer la diversité génétique, tout système générant de la diversité génétique serait positivement sélectionné. De par leur contribution à la plasticité des génomes, la présence de ces éléments génétiques mobiles pourrait donc conférer un avantage adaptatif aux *Wolbachia* (Aras *et al.*, 2003).

En plus d'être une source d'évolution des génomes, les éléments mobiles et les protéines à domaine ankyrine sont potentiellement impliqués dans les phénotypes de manipulation de la

reproduction induits par *Wolbachia*. La présence de très peu d'IS et l'absence totale de prophage dans le génome de la *Wolbachia* obligatoire et mutualiste *wBm* infectant *B. malayi* et leur surreprésentation dans les génomes des *Wolbachia* parasites du sexe, supportent fortement leurs implications dans l'induction des « stratégies » de manipulation de la reproduction (Wu *et al.*, 2004; Foster *et al.*, 2005; Klasson *et al.*, 2008, 2009). Du fait de leur fonction dans la médiation des interactions entre les protéines de l'hôte et de la bactérie notamment au niveau de la régulation des cycles cellulaires de l'hôte, les gènes codant pour des protéines à motifs ankyrines constituent d'excellents candidats. Ainsi, le gène *Spaid* de la bactérie *Spiroplasma poulsonii* récemment montré comme impliqué dans l'induction du male-killing chez *D. willistoni* contient des domaines ankyrines (Harumoto et Lemaitre 2018). Par ailleurs, de récentes études ont mis en évidence l'implication de gènes associés à des régions du prophage WO dans l'induction de l'incompatibilité cytoplasmique chez la Drosophile (Beckmann *et al.*, 2017; LePage *et al.*, 2017; Shropshire *et al.*, 2018). Mais nous reverrons ces derniers en détail car ils sont au cœur même de cette thèse.

d. *Wolbachia* peut même devenir mutualiste ou pas...

Bien que la très grande majorité des effets des *Wolbachia* chez ses hôtes soit associée à des phénotypes de type parasitisme sexuel, il existe des cas où les *Wolbachia* confèrent un avantage sélectif à son hôte. Ainsi, chez la mouche *D. melanogaster*, *wMel* confère une résistance lors d'infections par des virus ARN (Teixeira *et al.*, 2008) alors que chez le cloporte *A. vulgare* elle confère une résistance lors d'infections par des bactéries pathogènes (Braquart-Varnier *et al.*, 2015). Certaines études montrent que *Wolbachia* peut également être associée à une augmentation de la fécondité des femelles comme c'est le cas chez *Aedes albopictus* (Dobson *et al.*, 2002b) et *D. melanogaster* (Fry *et al.*, 2004). Dans les exemples précédents, des traits bénéfiques peuvent compenser totalement ou partiellement les coûts engendrés par l'infection. Cependant, les *Wolbachia* ne sont pas pour autant clairement mutualistes et encore moins obligatoires. En revanche, chez la punaise de lit *Cimex lectularius*, la souche *wCle* a été montrée comme essentielle à la croissance et à la reproduction de son hôte car elle est la source majeure de vitamine B absente du régime alimentaire de la punaise (Hosokawa *et al.*, 2010). Chez cette espèce, l'interdépendance semble avoir évolué secondairement, dans le contexte d'une interaction mutualiste. Il faut bien garder à l'esprit que le caractère obligatoire d'une association n'en fait pas une symbiose mutualisme. En fait, il est possible que les *Wolbachia* se soient « évolutivement parlant » rendues obligatoires. L'interaction entre la guêpe parasitoïde *Asobara tabida* et ses *Wolbachia* est un bon exemple de cette dépendance induite par la bactérie. Cette espèce de parasitoïde est infectée par trois souches de *Wolbachia* *wAtab1* et *wAtab2* qui induisent de l'IC (Dedeine *et al.*, 2004) et *wAtab3* qui est obligatoire pour que l'ovogenèse de l'hôte se déroule normalement (Dedeine *et al.*, 2001). D'ailleurs, l'espèce *Asobara rufescens* infectée uniquement par *wAtab1* ne nécessite pas la présence de *Wolbachia*

pour se reproduire (Dedeine *et al.*, 2005). Ces résultats suggèrent une dépendance spécifique de *A. tabida* à *wAtab3* lors de l'ovogenèse (Dedeine *et al.*, 2004, 2005). Le traitement aux antibiotiques de *A. tabida* entraîne une mort précoce par apoptose des cellules nourricières qui alimentent l'ovocyte (Pannebakker *et al.*, 2007). L'apoptose est un moyen de défense des eucaryotes contre les pathogènes intracellulaires car la mort cellulaire permet leur élimination (Everett & McFadden, 1999; Behar *et al.*, 2011). Par conséquent, de nombreux pathogènes comme par exemple *Listeria* ont développé des stratégies de contournement en manipulant le mécanisme d'apoptose (Balcewicz-Sablinska *et al.*, 1998; Gao & Abu Kwaik, 2000). Dans le cas de *A. tabida*, *Wolbachia* aurait dérégulé le programme d'apoptose de certaines cellules hôtes. *Wolbachia* limiterait le niveau d'apoptose dans les cellules qui l'hébergent, ce qui la protège de ce moyen de défense. En réponse, l'hôte aurait augmenté l'expression basale des facteurs apoptotiques permettant ainsi la survie des cellules nourricières infectées par *Wolbachia*. Cependant, lors d'un traitement par des antibiotiques qui mène à la destruction des *Wolbachia*, cet « équilibre » est rompu et l'apoptose surexprimée par l'hôte se retourne contre lui. Ainsi, *Wolbachia* est maintenant essentielle à *A. tabida* sans qu'aucune interaction à bénéfice réciproque n'ait été impliquée dans l'évolution de ce système (Pannebakker *et al.*, 2007). Il semble que cette « prise d'otage évolutive » soit en cours chez les nématodes filaires où le retrait artificiel des *Wolbachia* entraîne des problèmes de mue chez les larves et de reproduction chez les adultes (Taylor *et al.*, 2005; Landmann *et al.*, 2011).

IV. *Wolbachia* et l'incompatibilité cytoplasmique

L'IC est la stratégie de manipulation de la reproduction la plus communément induite par *Wolbachia* chez les arthropodes.

a. Direction et diversité de l'IC

i. Direction de l'IC

Comme décrit précédemment, l'IC qui permet aux *Wolbachia* d'envahir une population a lieu lorsqu'un mâle infecté copule avec une femelle non infectée et se traduit par la mort de la descendance lors de l'embryogenèse (Figure 4). Cependant, l'IC peut également avoir lieu entre individus mâles et femelles infectés par des souches de *Wolbachia* différentes et incompatibles. Deux types de croisements réciproques sont alors observés : soit seul un sens de croisement est compatible et l'autre sens de croisement est incompatible et l'on parlera d'IC unidirectionnelle, soit les deux sens de croisement sont incompatibles et l'on parlera d'IC bidirectionnelle (Figure 5) (Laven, 1967b; Breeuwer & Werren, 1990; O'Neill & Karr, 1990). La direction de l'IC (compatible/incompatible) peut varier selon i) la souche de *Wolbachia* impliquée comme chez *A. albopictus*, *C. pipiens*, *D. simulans* ou encore *L. heterotoma*,

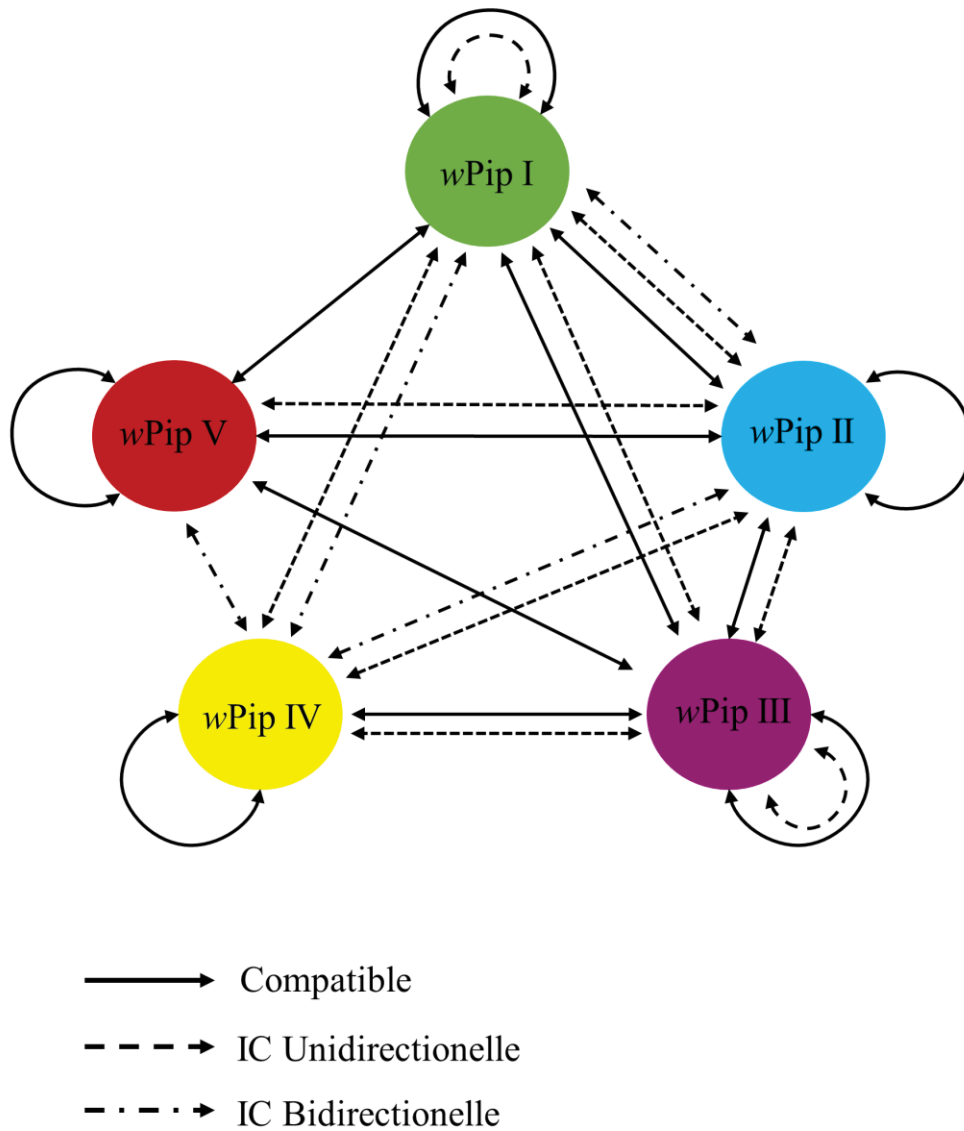


Figure 7 : Représentation schématique des relations d'incompatibilité cytoplasmique entre des lignées de *C. pipiens* infectés par des souches de *wPips* de groupes différents.

Figure adaptée d'Atyame et al. (2014)

(O'Neill & Karr, 1990; Dobson *et al.*, 2001; Vavre *et al.*, 2001; Atyame *et al.*, 2014) ou ii) par la présence d'infections multiples par différentes souches de *Wolbachia* comme chez *A. albopictus*, *L. heterotoma*, *D. simulans*, ou encore *Nasonia spp.* où les mâles multi-infectés sont unidirectionnellement incompatibles avec les femelles mono-infectées (Clancy & Hoffmann, 1996; Bordenstein & Werren, 1998; Dobson *et al.*, 2001; Vavre *et al.*, 2001).

ii. Diversité des phénotypes d'IC : *C. pipiens* mène la danse

Lorsque plusieurs souches de *Wolbachia* infectent une même espèce, les relations d'incompatibilités entre individus porteurs de ces différentes souches se complexifient comme c'est le cas chez *Drosophila spp.* ou *Nasonia spp.* (Poinsot *et al.*, 1998; Merçot & Charlat, 2004; Bordenstein & Werren, 2007; Atyame *et al.*, 2014).

L'espèce qui présente de très loin la plus grande diversité de relation d'incompatibilité décrite à ce jour est le moustique *C. pipiens*. Grâce à de très nombreux croisements réalisés depuis les années 1950 entre des *C. pipiens* provenant de l'ensemble de la planète, une incroyable matrice de phénotype d'IC a été révélée (Figure 7) (Laven, 1967b, 1956; Barr, 1980; Irving-Bell, 1983; Raymond *et al.*, 1986; Magnin *et al.*, 1987; O'Neill & Paterson, 1992; Guillemaud *et al.*, 1997; Rasgon & Scott, 2003; Duron *et al.*, 2006b, 2011, 2012, Atyame *et al.*, 2011b, 2014, 2015b). Dans le but d'étudier cette diversité, un protocole de caractérisation systématisée des phénotypes d'IC a été mis au point dans notre équipe. Ce système appelé « 4ref-cytotypes » repose sur des croisements entre lignées de *C. pipiens* provenant de la nature (dont on veut connaître le phénotype d'IC) avec 4 souches de référence de laboratoire choisies de façon arbitraire (Atyame *et al.*, 2014). Dans ce système, une lignée de moustiques infectée par une souche de *wPip* est caractérisée par un cytotype, c'est-à-dire ses relations d'incompatibilité (uni ou bi directionnelle) avec les quatre souches du 4ref-cytotypes avec lesquelles elle est croisée. Les cytotypes résultent de la combinaison d'un profil *mod* (défini par 4 valeurs correspondant aux 4 croisements avec les femelles de références) et d'un profil *resc* (défini par 4 valeurs correspondant aux 4 croisements avec les mâles de références). Ainsi, Atyame *et al.* 2014 ont mis en évidence l'existence de 14 cytotypes différents en croisant 51 lignées de *C. pipiens* récoltées sur l'ensemble du globe avec les 4 lignées de référence du 4ref-cytotypes. Au total, 8 profils *mod* (relation d'incompatibilité des mâles de ces 51 lignées vis-à-vis des femelles des lignées du 4ref-cytotypes) et 4 profils *resc* (relation d'incompatibilité des femelles des 51 lignées vis-à-vis des mâles des 4 lignées du 4ref-cytotypes) ont été établis grâce à ces différents croisements.

Jusqu'en 2004, la grande diversité de relation d'incompatibilité observée chez *C. pipiens* constituait un paradoxe pour les chercheurs car aucun polymorphisme n'était détecté chez *wPip* dans les gènes habituellement utilisés pour discriminer les souches de *Wolbachia*. En effet, même pour le gène *wsp* considéré comme étant le plus polymorphe au sein des MLST classiquement utilisés, aucune

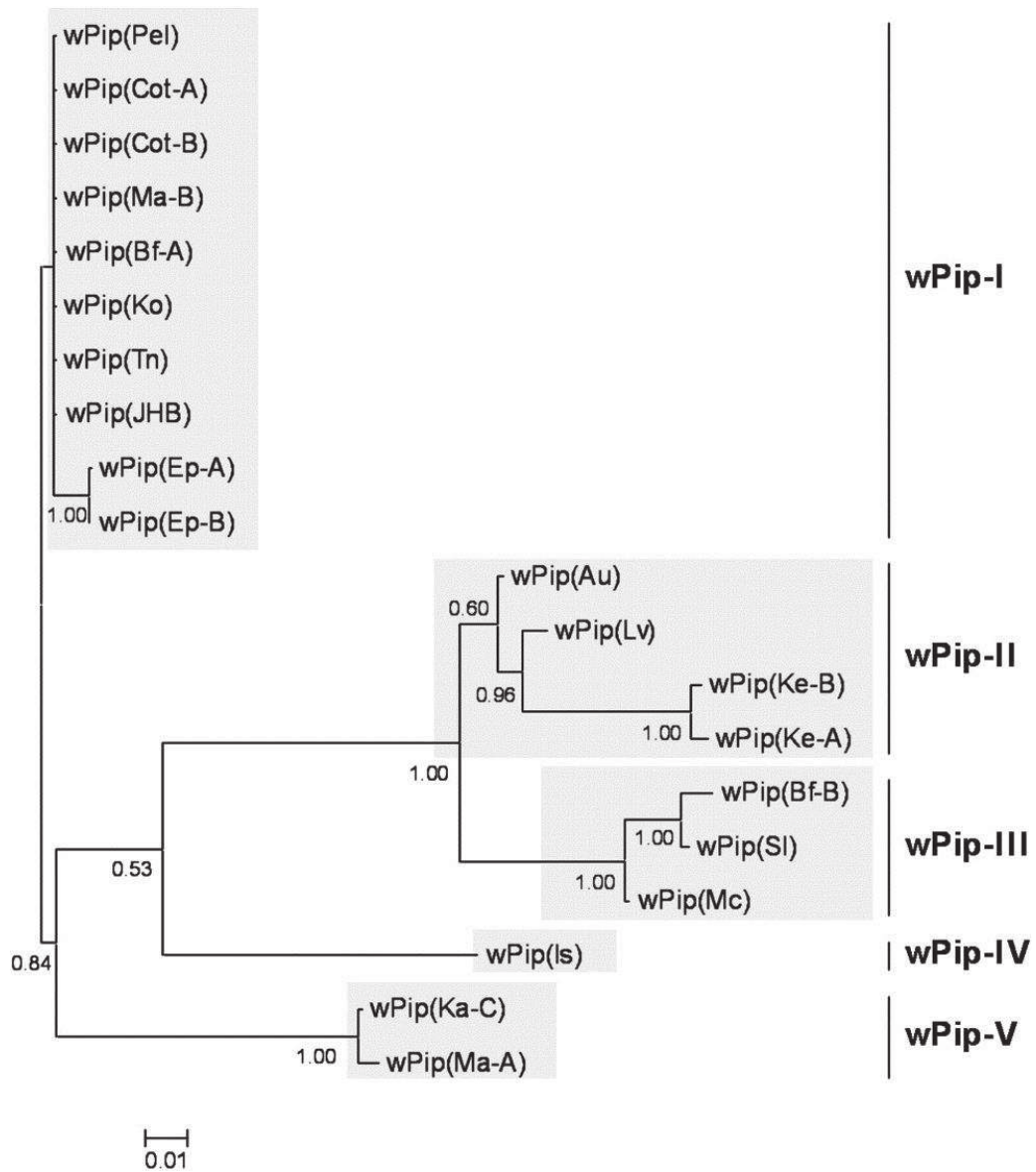


Figure 8 : Arbre phylogénétique des souches *wPips* obtenu à partir d'une analyse bayésienne de la concaténation des gènes *MutL*, *ank2*, *pk1*, *pk2*, *GP12*, *GP15*, et *RepA*.

D'après Atyame et al. (2011)

diversité n'était observée chez les *wPips* (Guillemaud *et al.*, 1997; Ruang-Areerate *et al.*, 2003). C'est la découverte de la présence de prophage dans le génome des *Wolbachia* (Masui *et al.*, 2000) qui a permis à Sanogo et Dobson en 2004 de montrer pour la première fois de la diversité génétique chez les *wPips*. Par la suite, le séquençage du génome de *wMel*, la souche de *Wolbachia* infectant *D. melanogaster* (Wu *et al.*, 2004) a révélé l'existence de gènes à domaine ankyrine, de transposons et de phages dont l'étude chez des souches de *wPips* a permis de révéler une diversité génétique jusqu'alors insoupçonnée (Duron *et al.*, 2005, 2006a, 2007a; Iturbe-Ormaetxe *et al.*, 2005; Walker *et al.*, 2007). Finalement, le séquençage du génome de *wPip* de référence (*wPipPel*), une souche de *Wolbachia* provenant d'une lignée de moustiques originaire du Sri Lanka (Klasson *et al.*, 2008) a permis d'accéder à encore plus de marqueurs génétiques. Grâce à ces marqueurs « hypervariables » plus de 100 souches différentes de *wPip* ont été caractérisées dans les populations naturelles sur l'ensemble du globe (Duron *et al.*, 2006a, 2011, Atyame *et al.*, 2011a; b, 2014, 2015b). Grâce à la mise au point d'un MLST spécifique des *wPip* basé sur les gènes *MutL*, *ank2*, *pk1*, *pk2*, *GP12*, *GP15*, et *RepA* (Atyame *et al.*, 2011a) il a été mis en évidence que les différentes *wPips* se rassemblent dans un clade monophylétique divisé en 5 groupes phylogénétiques allant de *wPipI* à *wPipV* (Atyame *et al.* 2011a; Dumas *et al.* 2013, Figure 8). Les groupes de *wPips* sont structurés géographiquement sur la planète (Dumas *et al.* 2013, Figure 9). Le groupe *wPipI*, le plus répandu, est présent majoritairement en Afrique (Sub-Saharienne et du Nord), en Amérique du Sud et en Asie du Sud-Est. Le deuxième plus répandu est le groupe *wPipIII* observé en Amérique du Nord et en Europe. Les trois autres groupes *wPip II*, *IV* et *V* sont moins communs. Les groupes *wPip II* et *V* sont restreints respectivement à l'Europe de l'Ouest et l'Asie sauf quelques exceptions. Le groupe *wPipIV* est plus dispatché car on le retrouve de manière sporadique en Europe, en Afrique du Nord et en Asie (Altinli *et al.*, 2018). Généralement, un seul groupe est retrouvé dans une région géographique donnée sauf en Europe où les 5 groupes sont présents (Dumas *et al.*, 2013). Cette coexistence des 5 groupes en Europe a conduit les auteurs à considérer l'Europe comme le berceau de l'infection par *Wolbachia* où les *wPip* se seraient diversifiées en 5 groupes puis étendues sur l'ensemble de la Planète notamment grâce aux migrations humaines. L'analyse de plus de 900 croisements entre lignées de *C. pipiens* provenant de l'ensemble de la planète, a mis en évidence une corrélation entre la proximité génétique des *wPips* et leur probabilité d'induire de l'IC entre leurs hôtes respectifs (Atyame *et al.*, 2014). En effet, les lignées de *C. pipiens* infectées par des *wPips* appartenant au même groupe sont très majoritairement compatibles entre elles (Figure 8). En revanche, lorsque des moustiques sont infectés par des souches de *wPips* de groupes différents, l'IC est beaucoup plus fréquemment observée (Figure 8). **Cependant, il n'est pas possible de prédire l'issue d'un croisement entre deux lignées de moustiques infectées en se basant simplement sur la caractérisation des groupes de *wPips* qu'elles hébergent. En effet, au sein d'un groupe de *wPip* des souches montrant des différences génétiques mais aussi de cytotypes existent. Pour prévoir l'issue d'un croisement, il faudrait connaître les bases génétiques de l'IC.**

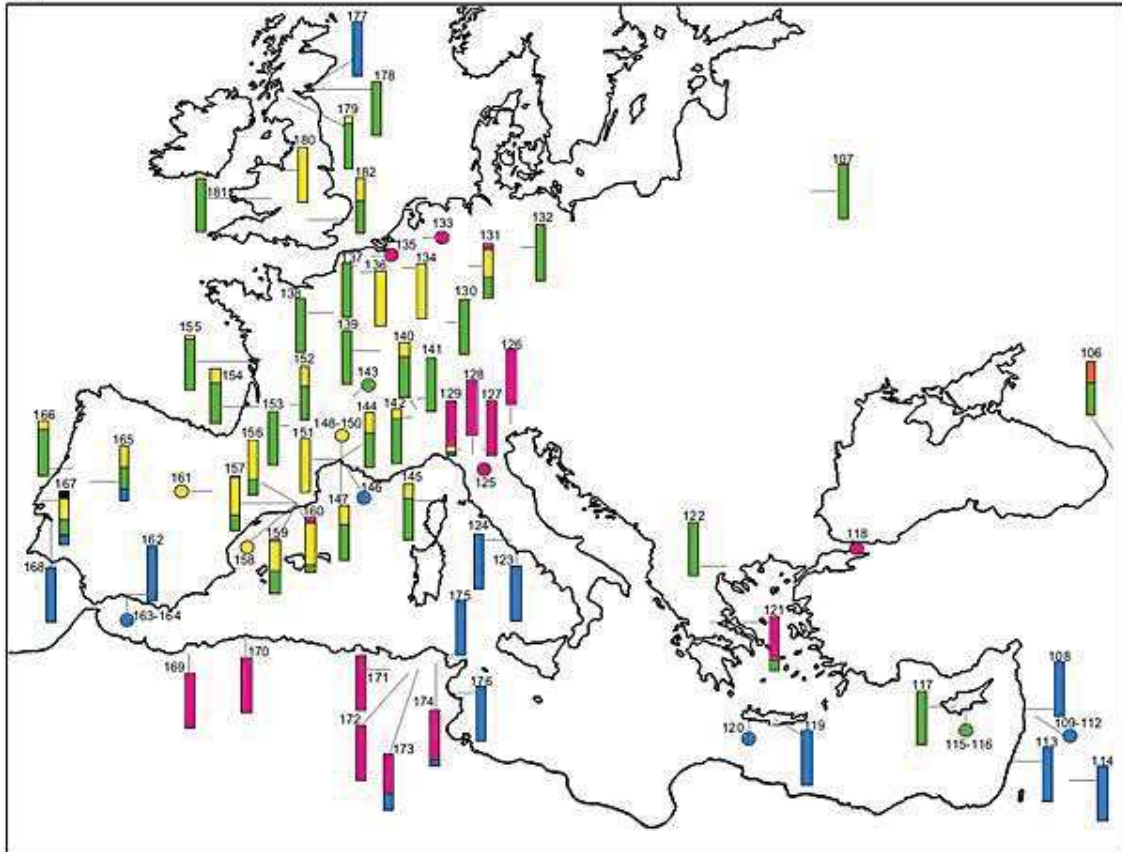
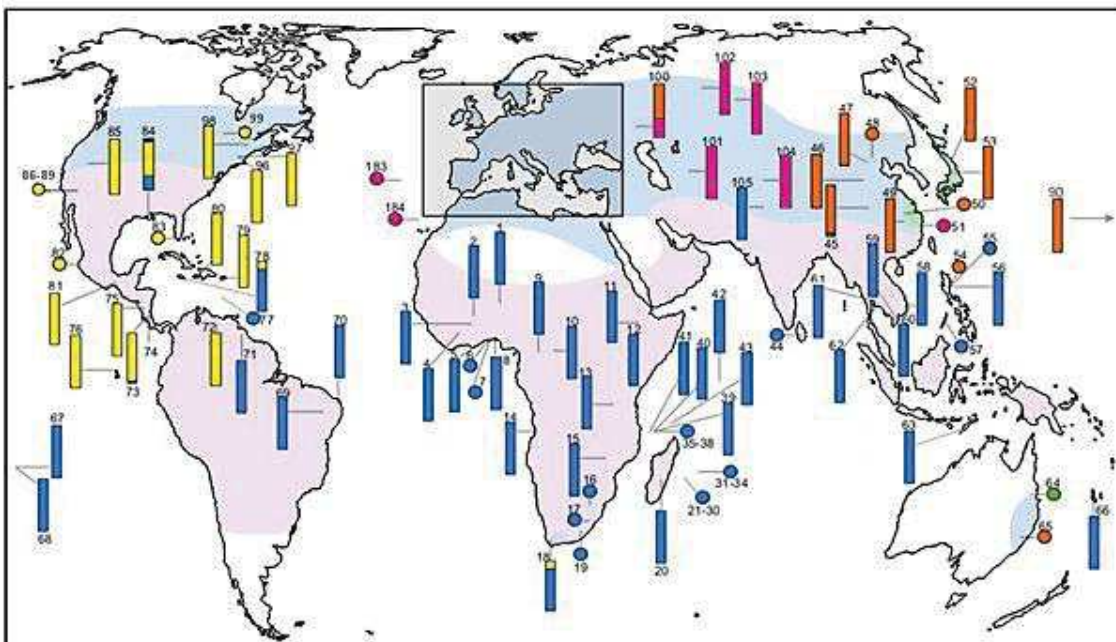
A**B**

Figure 9 : Distribution géographique des groupes de wPip dans le monde (a) et en Europe (b). Les barres et les points représentent respectivement les populations naturelles et les souches de laboratoire. Bleu : groupe wPipI ; vert : groupe wPipII ; jaune : groupe wPipIII ; rose ; groupe wPipIV et orange groupe wPipV.
Figure adaptée de Dumas et al. (2013)

b. Pénétrance de l'IC

La pénétrance de l'IC (*i.e.* proportion d'embryons morts du fait de l'IC) dépend de plusieurs facteurs mêlant géotypes des *Wolbachia* et de leurs hôtes. Ainsi, chez *D. simulans* la pénétrance de l'IC va être variable en fonction de la souche de *Wolbachia* allant d'une pénétrance très forte de *w*Ri (quasiment 100%) à une pénétrance plus faible de *w*No pouvant descendre jusqu'à 40% de mortalité (Merçot & Charlat, 2004). La densité de *Wolbachia* est un facteur majeur de la pénétrance de l'IC chez certaines espèces comme *D. simulans*, *D. melanogaster*, *Ephestia kuehniella*, *Laodelphax striatellus* et *Sogatella furcifera* (Boyle *et al.*, 1993; Clancy & Hoffmann, 1998; Noda *et al.*, 2001; Ikeda *et al.*, 2003). Enfin, chez *L. heterotoma*, les mâles multi-infectés induisent une IC plus forte que les mâles mono-infectés suggérant que l'infection multiple peut chez certaines espèces influencer la pénétrance de l'IC (Vavre *et al.*, 2001). Du côté hôte, l'âge a été montré comme pouvant influencer la pénétrance de l'IC. En effet, il a été montré chez les drosophiles que la proportion de spermatoocytes infectés par *Wolbachia* diminue avec l'âge et se traduit par une réduction de la pénétrance de l'IC avec le vieillissement des mâles (Jamnongluk *et al.*, 2000; Noda *et al.*, 2001; Kittayapong *et al.*, 2002; Reynolds & Hoffmann, 2002; Reynolds *et al.*, 2003). Chez *Nasonia spp.*, il a été montré que le passage de la souche *w*Av infectant normalement *N. vitripennis* chez *N. giraulti* induit une pénétrance d'IC plus forte suggérant une implication du fond génétique hôte sur la pénétrance de l'IC (Bordenstein & Werren, 1998; Poinot *et al.*, 1998; McGraw *et al.*, 2001). Enfin, il a été montré que des facteurs environnementaux tels que la température (Hoffmann *et al.*, 1986; Clancy & Hoffmann, 1998; Bordenstein & Bordenstein, 2011) et la densité d'élevage de la population hôte (Sinkins *et al.*, 1995) pouvait également avoir un impact sur la pénétrance de l'IC. Par exemple, le fait d'élever des drosophiles à des températures élevées entraîne une diminution de la pénétrance de l'IC probablement liée à une diminution de la densité de *Wolbachia* (Clancy & Hoffmann, 1998).

Contrairement aux espèces citées ci-dessus, chez le moustique *C. pipiens*, les différentes souches de *Wolbachia*, la densité, les variations génétiques hôtes et l'âge des mâles n'ont montré aucune influence sur la pénétrance de l'IC (Duron *et al.*, 2007b; Atyame *et al.*, 2011b). En effet, chez *C. pipiens*, quand un croisement est incompatible, il l'est à 99,99% quelles que soient les souches de *w*Pips impliquées où la provenance du moustique hôte (Guillemaud *et al.*, 1997; Duron *et al.*, 2006b, 2007b, 2011, Atyame *et al.*, 2011b, 2014). Cette pénétrance « totale » de l'IC semble être une caractéristique ancestrale que possèdent toutes les *w*Pips.

c. Dynamique de l'infection sous l'influence de l'IC

La capacité des *Wolbachia* à envahir les populations hôtes via l'IC a été explorée et démontrée théoriquement (Caspari & Watson, 1959; Fine, 1978; Hoffmann *et al.*, 1990; Turelli, 1994; Keeling *et al.*, 2003), expérimentalement au laboratoire chez les moustiques *A. albopictus* (Kambhampati *et al.*, 1992; Dobson *et al.*, 2002a) et *Aedes aegypti* (Xi *et al.*, 2005) et observée en milieu naturel chez *D. simulans* en Californie (Turelli & Hoffmann, 1991) et chez *Laodelphax striatellus* au Japon (Hoshizaki & Shimada, 1995). La capacité de *Wolbachia* à envahir les populations repose sur la conséquence directe de l'IC, à savoir que les femelles infectées pourront se reproduire et donner une descendance viable avec des mâles infectés et non infectés (100% des descendants seront viables) alors que les femelles non infectées ne pourront se reproduire et donner une descendance viable qu'avec les mâles non infectés (50% des descendants seront viables). Il est cependant important de noter que la production de fils infectés constitue un coût pour les femelles infectées car les fils infectés qu'elles produisent donneront une descendance non viable avec les femelles non-infectées. L'IC est donc avantageux pour la propagation de *Wolbachia*, mais induit un coût pour les femelles infectées. Malgré ce coût, en moyenne, les femelles infectées produisent plus de descendants que les femelles non infectées. Ce différentiel de production de descendants entre femelles infectées et non infectées constitue le carburant de la propagation de *Wolbachia* dans les populations. Cependant, cette propagation va dépendre de plusieurs caractéristiques intrinsèques à l'interaction entre la bactérie et son hôte. On peut citer le taux de transmission de *Wolbachia* dans les ovocytes et la pénétrance de l'IC mais aussi des traits parasitaires comme le coût à l'infection qui peuvent être contrebalancés par des traits bénéfiques comme, par exemple, la « protection » conférée contre des virus (Teixeira *et al.*, 2008). Ainsi, dans le cas « idéal » où la transmission de *Wolbachia* est de 100% (*i.e.* tous les ovocytes sont infectés), où la pénétrance de l'IC est totale et où l'infection ne confère aucun coût, *Wolbachia* pourra se propager dans la population jusqu'à fixation. Cependant, dans les populations naturelles, la transmission de *Wolbachia* et la pénétrance de l'IC sont rarement de 100% et il peut exister un coût sur la fécondité des femelles infectées comme cela a été montré chez *D. simulans* (Hoffmann *et al.*, 1990). Ainsi, la dynamique d'invasion de *Wolbachia* dans les populations naturelles n'est pas toujours « idéale » et de nombreux modèles ont été proposés afin de déterminer les conditions dans lesquelles l'invasion par *Wolbachia* reste possible (Caspari & Watson, 1959; Fine, 1978; Hoffmann *et al.*, 1990; Turelli, 1994; Engelstädter & Telschow, 2009).

Si le coût de l'infection est supérieur à l'avantage sélectif conféré par l'IC (*i.e.* le différentiel de descendants entre femelles infectées et non infectées) alors les femelles infectées seront désavantagées par rapport aux femelles non infectées et l'infection va disparaître. Si en revanche, le coût à l'infection est inférieur à l'avantage conféré par l'IC ou que la bactérie confère un avantage à son hôte alors l'invasion est possible et va dépendre de la fréquence initiale de l'infection dans les populations. Ce seuil de fréquence initiale va globalement dépendre du rapport entre le coût à l'infection et le bénéfice

conféré par l'IC. Plus le coût à l'infection est faible par rapport à l'avantage conféré par l'IC, moins le nombre initial de *Wolbachia* nécessaire à l'invasion sera élevé lors de la primo infection d'une population. Si la fréquence d'infection initiale est en dessous de ce seuil, l'infection va disparaître. En revanche, si la fréquence initiale d'infection est au-dessus de ce seuil alors l'infection va augmenter dans les populations jusqu'à théoriquement atteindre la fixation. La fixation n'est pas systématiquement observée dans les populations et ceci peut être dû à une i) transmission verticale imparfaite, ii) une pénétrance de l'IC non totale, iii) la migration, iv) à l'assortiment différentiel entre individus non infectés qui entraîne la production de manière récurrente d'individus non infectés dans les populations et qui empêchent la fixation (Hoffmann *et al.*, 1990; Turelli, 1994).

Chez certaines espèces hôtes, les dynamiques d'infection par *Wolbachia* peuvent devenir plus compliquées du fait de la présence de plusieurs souches différentes et incompatibles de *Wolbachia* dans une même population. Lorsque les mâles et les femelles sont infectés par des souches différentes et incompatibles, l'IC peut soit être unidirectionnelle (un sens de croisement est compatible alors que l'autre est incompatible) soit bidirectionnelle (les deux sens de croisement sont incompatibles). Dans une population panmictique, deux souches de *Wolbachia* incompatibles (unidirectionnellement ou bidirectionnellement) ne peuvent théoriquement pas coexister (Rousset *et al.*, 1991; Engelstädter & Telschow, 2009). En effet, dans le cas d'IC bidirectionnelle, c'est la souche de *Wolbachia* la plus fréquente dans la population qui va gagner la compétition et se maintenir dans la population. Dans le cas de l'IC unidirectionnelle, c'est la souche de *Wolbachia* qui est incompatible qui va théoriquement gagner la compétition et envahir la population (uniquement si le seuil de fréquence présenté précédemment est dépassé). En revanche, lorsque les populations hôtes sont structurées (*i.e.* non panmictique), alors, sous certaines conditions, des *Wolbachia* incompatibles peuvent coexister (Engelstädter & Telschow, 2009). C'est par exemple possible quand il y a homogamie, c'est-à-dire un assortiment préférentiel entre individus infectés par la même souche de *Wolbachia* qui va permettre l'évitement de l'IC et le maintien du polymorphisme d'infection (Rousset *et al.*, 1991; Jaenike *et al.*, 2006). Les populations hôtes peuvent aussi être structurées spatialement en sous-populations où des événements de migration peuvent avoir lieu. Dans une telle structure en métapopulations, le maintien du polymorphisme d'infection de souches de *Wolbachia* différentes et incompatibles va dépendre du taux de migration entre sous-populations (Telschow *et al.*, 2005, 2007; Flor *et al.*, 2007). Il existe un taux de migration critique en dessous duquel les souches de *Wolbachia* se maintiennent dans leurs sous-populations et au-dessus duquel une des deux souches va envahir. Ce taux de migration critique va dépendre de plusieurs paramètres comme le type d'IC (bidirectionnel ou unidirectionnel), le taux de transmission, la pénétrance de l'IC, le coût de l'infection, l'existence d'adaptation locale...

L'étude des dynamiques d'infections des populations par *Wolbachia* est particulièrement importante du fait de son utilisation via sa capacité à induire l'IC dans la lutte antivectorielle (voir Flores & O'Neill 2018 pour revue).

d. Conséquences évolutives de l'IC

L'incompatibilité cytoplasmique va avoir des conséquences sur l'évolution des populations hôtes. Une des conséquences qui peut être conceptualisée est une influence sur la spéciation (Coyne 1992; Hurst & Schilthuizen 1998; voir Charlat *et al.* 2003 pour revue). En effet, si on se place dans le cas de deux populations d'une même espèce, l'une infectée et l'autre non, l'incompatibilité entre les mâles de la population infectée et les femelles de la population non infectée peut réduire les flux de gènes entre ces populations. Ceci peut aussi se produire de la même manière lorsque les deux populations sont infectées par des *Wolbachia* différentes et incompatibles (uni ou bidirectionnellement). La réduction des flux de gènes entre hôtes qui peut s'en suivre conduirait à la divergence entre les populations devenues incompatibles et *in fine* pourrait conduire à de l'isolement reproductif et de la spéciation. Cependant, même si théoriquement l'IC constitue un moteur potentiel à la spéciation, aucune évidence empirique ne permet de le confirmer.

Une autre conséquence théorique potentielle de l'induction de l'IC par *Wolbachia* dans les populations hôtes est l'extinction (voir Charlat *et al.* 2003 pour revue). Premièrement, lors de l'invasion de *Wolbachia* dans une population hôte, le remplacement des individus non infectés par les individus infectés induit une diminution de la diversité génétique dans les populations au niveau nucléaire mais surtout au niveau de la diversité mitochondriale. En effet, le variant mitochondrial présent dans le même cytoplasme que *Wolbachia* va au cours de l'invasion remplacer progressivement ceux déjà présents dans la population (Turelli *et al.*, 1992; Rousset & Solignac, 1995). Deuxièmement, lors de la phase d'invasion par *Wolbachia* dans une population non infectée, on observe une diminution de la taille des populations due à la diminution du nombre moyen de descendants des femelles non infectées. C'est en partie pour cette raison que, depuis 1967, le couple *Wolbachia*/IC est fortement envisagé dans un but de lutte antivectorielle (Laven, 1967a). La diminution de la taille des populations augmente la probabilité de fixation de mutations faiblement délétères par dérive génétique (Ohta, 1987; Akashi *et al.*, 2012). L'accumulation de ces mutations faiblement délétères associée à une diminution de la diversité génétique peut théoriquement conduire les populations à l'extinction (Dobson *et al.* 2002a; voir Charlat *et al.* 2003 pour revue).

e. Conceptualisation des mécanismes l'IC

A l'heure actuelle, l'IC est généralement conceptualisée comme un modèle *mod/resc* ou toxine/antidote dans lequel l'implication de *Wolbachia* serait différente en fonction de si elle se trouve chez le mâle ou chez la femelle (Werren, 1997). Les *Wolbachia* présentes chez les mâles introduiraient dans les spermatozoïdes une toxine (fonction *mod*) lors de la spermatogénèse qui, après la fécondation, si l'œuf ne contient pas de *Wolbachia*, entrainerait la mort des embryons et le croisement serait qualifié d'incompatible. En revanche, si des *Wolbachia* sont présentes dans l'ovocyte, alors il y a mise en place

d'un « sauvetage » (la fonction *resc*) de l'effet de la toxine grâce à un antidote qui permet de restaurer la viabilité des embryons. Ce modèle théorique n'implique aucune hypothèse quant aux mécanismes cellulaires sous-jacents de la fonction *mod* et de la fonction *resc*.

Ainsi, plusieurs modèles plus précis sur le plan fonctionnel ont été proposés (voir Poinso *et al.* 2003 pour revue)

- Le premier modèle appelé « clé-serrure » a été formulé par Hurst en 1991 (Hurst, 1991a) et propose que la fonction *mod* soit basée sur le dépôt dans les spermatozoïdes par les *Wolbachia* paternelles d'une « serrure » qui viendrait se fixer sur le matériel génétique contenu dans les spermatozoïdes. Cette serrure bloquerait l'ADN paternel l'empêchant une fois la fécondation réalisée de prendre part à l'embryogenèse si aucune *Wolbachia* n'est présente dans l'œuf. En revanche, si des *Wolbachia* sont présentes dans l'œuf, alors elles apporteraient une « clé » (la fonction *resc*) qui permettrait de déverrouiller le matériel génétique paternel et le déroulement d'une embryogenèse normale. Dans ce modèle, les gènes bactériens responsables de la fonction *mod* et de la fonction *resc* seraient clairement différents et les produits de ces gènes seraient capables d'interagir directement l'un avec l'autre après la fécondation.
- Le deuxième modèle proposé par Kose & Karr 1995 appelé « titration-restitution » propose quant à lui que les *Wolbachia* paternelles enlèveraient lors de la spermatogenèse des protéines associées aux chromosomes contenus dans les spermatozoïdes (fonction *mod*) mais que de telles protéines seraient aussi enlevées dans les ovocytes par les *Wolbachia* maternelles. Après la fécondation, si des *Wolbachia* sont présentes dans l'œuf, elles restitueraient sur les chromosomes paternels et maternels ces protéines, ce qui permettrait à l'embryogenèse de se dérouler normalement. Dans ce modèle, les gènes responsables de la fonction *mod* et de la fonction *resc* pourraient être les mêmes. Ainsi, le passage de l'enlèvement (*mod*) à la restitution (*resc*) des protéines après la fécondation pourrait être régulé par des facteurs provenant de l'hôte. Ces gènes pourraient également être différents si, par exemple, les gènes *mod* codent des protéines responsables de l'enlèvement (*mod*) alors que les gènes *resc* codent des protéines capables d'inhiber l'enlèvement (*resc*) permettant la restitution.
- Enfin le dernier modèle qui porte le nom de « slow-motion » fut proposé par (Callaini *et al.*, 1997). Ce modèle propose que les *Wolbachia* produisent des facteurs qui se fixeraient sur le matériel génétique contenu dans les spermatozoïdes et que ces facteurs seraient responsables d'un ralentissement des mouvements des chromosomes paternels lors de la première division embryonnaire, ce qui induirait une asynchronie entre les jeux de chromosomes paternels et maternels. La fonction *resc* rétablirait la synchronie entre les deux jeux de chromosomes en fonctionnant de la même manière que la fonction *mod* (ralentissement) mais sur la chromatine maternelle. Dans ce modèle, le *mod* et le *resc* impliqueraient exactement les mêmes mécanismes et seraient pas conséquent très probablement codés par les « mêmes » gènes.

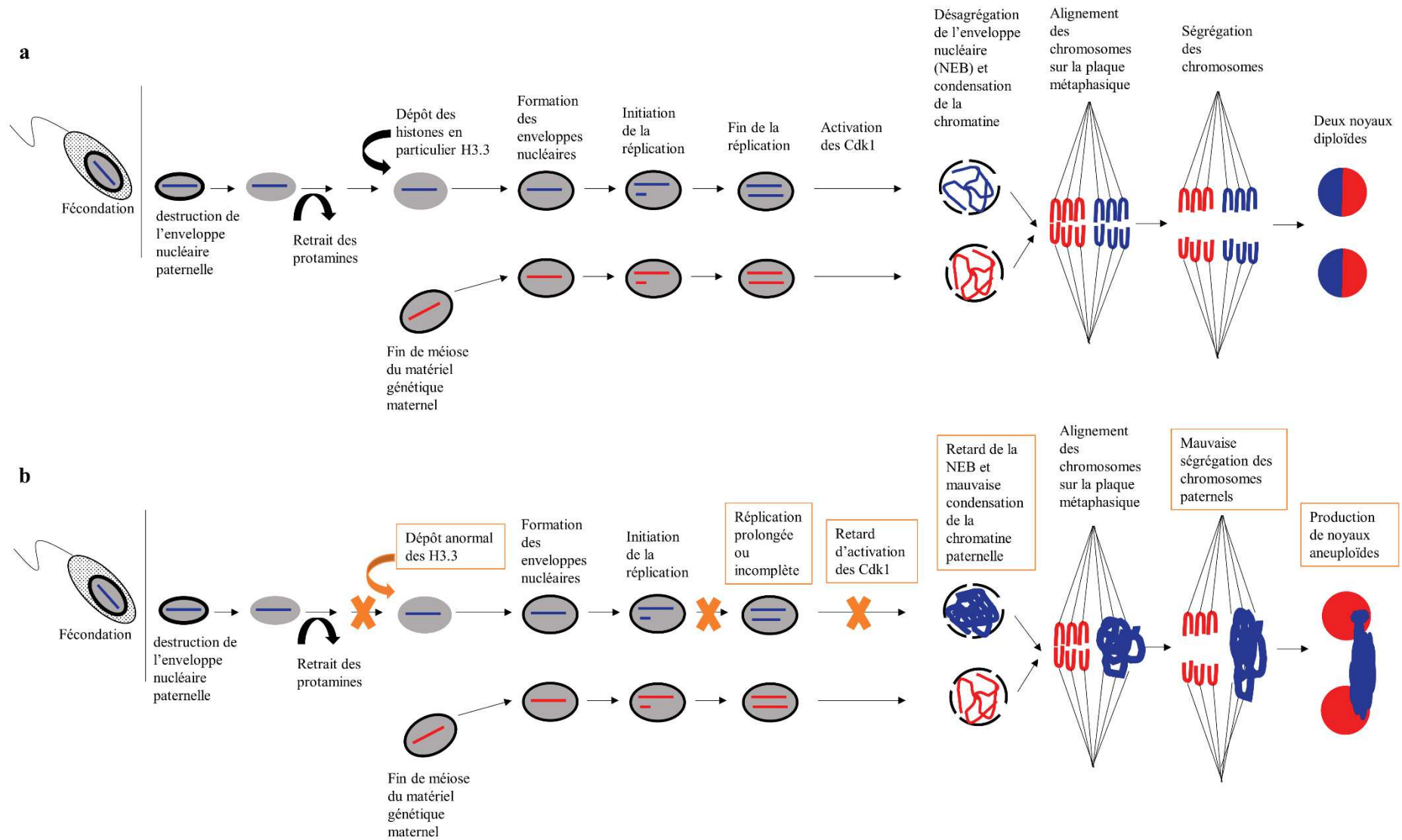


Figure 10 : Représentation schématique des événements clés depuis la fécondation à la fin de la première division embryonnaire chez un embryon issu d'un croisement compatible (a) et un embryon issu de croisement incompatible (b). Les défauts associés à l'IC sont entourés en orange. Figure adaptée de Landmann et al. (2009).

f. Les mécanismes cytologiques sous-jacents à l'IC

La mort de la descendance dans les croisements incompatibles a lieu lors de l'embryogenèse. Les *Wolbachia* induisent l'IC en perturbant le développement embryonnaire. Des études cytologiques chez *D. melanogaster*, *D. simulans* et *N. vitripennis* ont permis de lever le voile sur une partie des mécanismes cytologiques de l'incompatibilité cytoplasmique. Ces études se sont concentrées sur les mécanismes cytologiques responsables de la mort des embryons (fonction *mod*) et pour le moment aucune étude des mécanismes de restauration de la fertilité (fonction *resc*) n'a été réalisée. Chez ces trois espèces, le phénomène cellulaire, qui entraîne un arrêt de l'embryogenèse et donc la mort des embryons, est un problème de condensation de la chromatine paternelle résultant en sa mauvaise ségrégation lors de l'anaphase de la première division embryonnaire (Ryan & Saul, 1968; Breeuwer & Werren, 1990; Reed & Werren, 1995; Callaini *et al.*, 1996; Lassy & Karr, 1996; Tram & Sullivan, 2002; Landmann *et al.*, 2009). Une partie des mécanismes cellulaires qui conduisent à cette mauvaise ségrégation de la chromatine paternelle ont été révélés et sont récapitulés dans la figure 10 (Tram & Sullivan, 2002; Landmann *et al.*, 2009). Lors de la spermatogenèse, du fait de la petite taille des spermatozoïdes, l'ADN contenu dans le noyau est extrêmement condensé grâce au remplacement des histones par des protamines. Après la fécondation, l'enveloppe nucléaire contenant l'ADN paternel est éliminée et les protamines sont remplacées par les histones H3.3, H4, H2A et H2B fournies par l'ovocyte (Landmann *et al.*, 2009). Une fois la chromatine paternelle remodelée avec ces nouvelles histones, une enveloppe nucléaire se forme autour de l'ADN paternel et une autre autour de l'ADN maternel ayant complété sa deuxième division méiotique formant ainsi deux pronuclei. La phase S pendant laquelle la réplication de l'ADN a lieu est ensuite initiée. A la fin de la phase S, la protéine Cdk1 (Cyclin-dependent kinase 1) est activée, ce qui permet la désagrégation de l'enveloppe nucléaire (nuclear envelope breakdown, NEB), la condensation de la chromatine et l'assemblage des fuseaux mitotiques avant l'entrée en mitose. Une des particularités de la première division mitotique de l'embryogenèse chez beaucoup d'insectes est la présence de deux jeux de fuseaux mitotiques (fuseaux gonométriques), qui sont régulés de manière indépendante à travers les étapes de la mitose dont l'un est associé aux chromosomes paternels et l'autre aux chromosomes maternels (Ryan & Saul, 1968; Callaini *et al.*, 1997). Une fois les fuseaux mitotiques assemblés, les chromosomes paternels et maternels sont placés dans des régions différentes de la plaque métaphasique puis les chromatides se séparent et deux noyaux diploïdes contenant chacun du matériel génétique paternel et maternel sont formés. Ces noyaux vont subir d'autres divisions mitotiques et donneront à l'issue de l'embryogenèse une larve viable.

Dans les embryons issus de croisements incompatibles, la désagrégation de l'enveloppe entourant l'ADN paternel et l'enlèvement des protamines se passent normalement. En revanche, il a été observé chez *D. simulans* un retard dans l'incorporation des histones H3.3 lors du remodelage de la chromatine (Landmann *et al.*, 2009). Ainsi, lors de la formation de l'enveloppe nucléaire, la chromatine paternelle n'est pas parfaitement remodelée. Il a été mis en évidence, toujours chez *D. simulans*, en

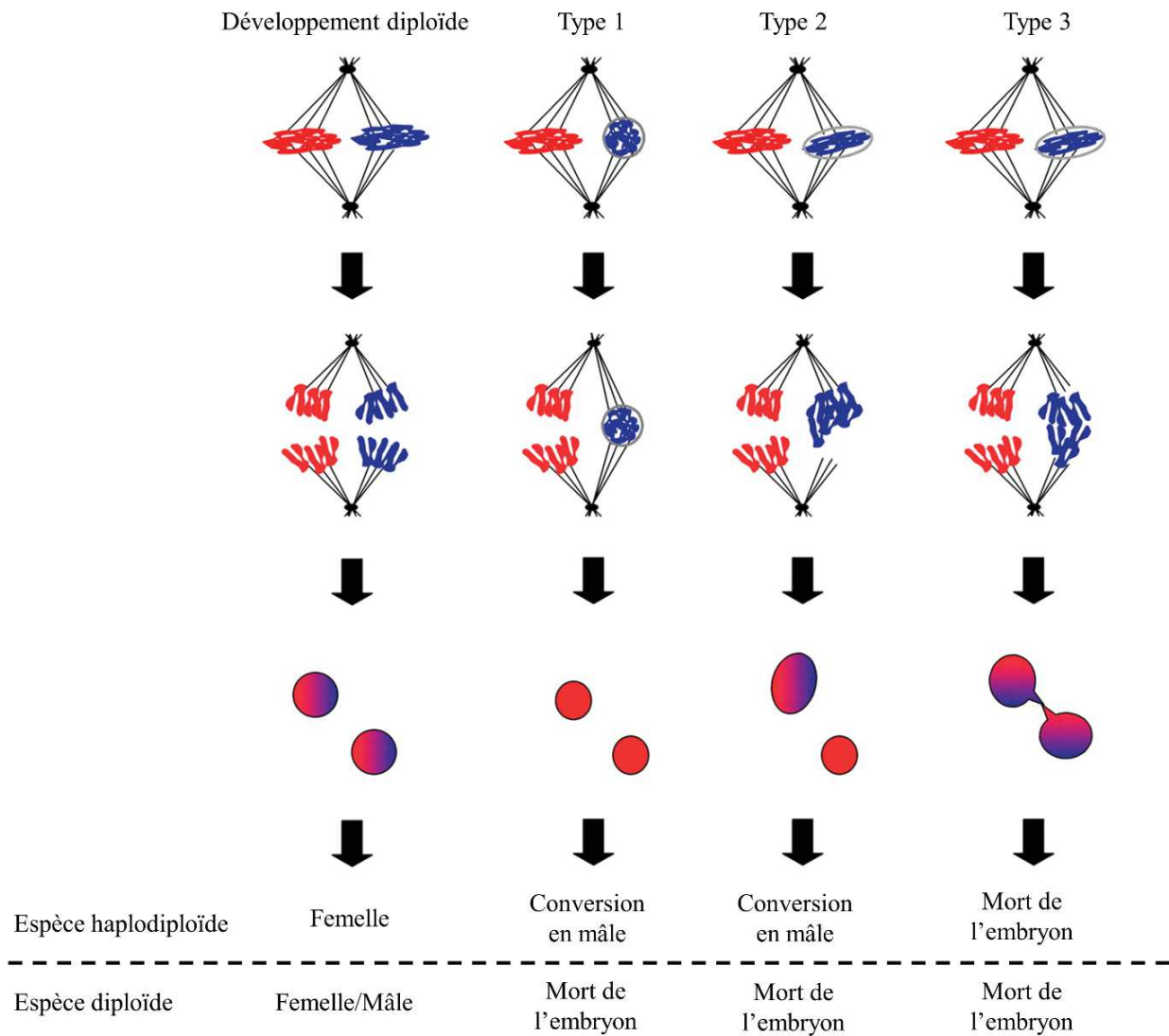


Figure 11 : Schématisation de l'issue de la première division embryonnaire en fonction du degré d'exclusion de la chromatine paternelle.

Lors d'un développement embryonnaire normal (à gauche) les chromatines maternelles (rouge) et paternelles (en bleu) s'alignent sur différentes régions de la plaque métaphasique puis les deux jeux de chromosomes ségrègent équitablement pour produire deux noyaux diploïdes. Dans les embryons issus de croisements incompatibles la chromatine paternelle est modifiée ce qui entraîne sa mauvaise ségrégation lors de l'anaphase de la première division embryonnaire. Trois cas peuvent être décrits en fonction du degré d'exclusion de la chromatine paternelle. Soit la chromatine paternelle est fortement modifiée et donc totalement exclue lors de la première division embryonnaire (type 1) ce qui permettra à la chromatine maternelle de ségréger et de produire deux noyaux haploïdes. Soit la modification de la chromatine paternelle est plus faible et elle ségrége anormalement. L'ADN paternel ségrége soit entièrement d'un côté du fuseau mitotique (type 2) ce qui produira un noyau haploïde et un noyau aneuploïde ; soit l'ADN paternel ségrége de manière inégale dans les deux noyaux (type 3) qui sont alors tous les deux aneuploïdes (on observe souvent des ponts chromatiniens entre ces noyaux). Chez des espèces haplodiploïdes comme *N. vitripennis* et *Leptopilina heterotoma* les embryons de type 1 et 2 (embryon haploïde) donneront des embryons mâles alors qu'ils seront non viables chez les espèces diploïdes comme *D. melanogaster* ou *C. pipiens*. Les embryons de type 3 seront toujours non viables chez ces deux types d'espèces.

Figure adaptée de Tram et al. (2006)

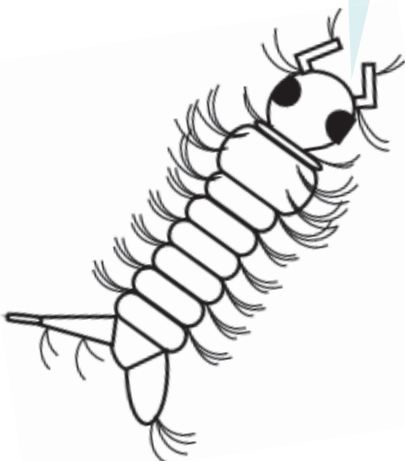
suivant la présence de PCNA (Proliferating Cell Nuclear Antigen) qui est un composant de la fourche de réplication de l'ADN, que malgré l'état structurel imparfait de la chromatine paternelle, la réplication de l'ADN dans les deux pronuclei est initiée en même temps (Landmann *et al.*, 2009). A la fin de la phase S, la réplication est terminée et PCNA n'est plus détecté sur la chromatine maternelle. En revanche, PCNA est détecté sur la chromatine paternelle jusqu'à la métaphase de la première division embryonnaire, ce qui indique que la réplication de la chromatine paternelle n'est pas entièrement complétée à la fin de la phase S (Landmann *et al.*, 2009). Ce retard dans la réplication de la chromatine paternelle pourrait être responsable de désynchronisation dans l'activation des Cdk1 associées respectivement aux chromosomes paternels et maternels observés chez *Nasonia spp.* (Tram & Sullivan, 2002; Landmann *et al.*, 2009). Ainsi, les Cdk1 associées aux chromosomes paternels vont être activées plus tardivement que celles associées à la chromatine maternelle, ce qui aura pour conséquence une condensation imparfaite des chromosomes paternels lors de l'alignement sur la plaque métaphasique et leur mauvaise ségrégation lors de l'anaphase (Tram & Sullivan, 2002; Landmann *et al.*, 2009). Cette mauvaise ségrégation aura pour conséquence la production de noyaux aneuploïdes, c'est-à-dire ne contenant pas tout le matériel génétique paternel et maternel attendu, entraînant un arrêt de l'embryogenèse et la mort des embryons.

Il a été mis en évidence chez *Drosophila* et *Nasonia spp.* que chez certains embryons, la chromatine paternelle peut i) former des ponts chromatiniens entre les chromatines maternelles en ségrégation, ii) ségréger uniquement dans un des deux noyaux issus de la division, entraînant la production d'un noyau haploïde constitué uniquement de matériel génétique maternel et d'un noyau aneuploïde ou iii) ne pas ségréger du tout ce qui entrainera la formation de deux noyaux haploïdes à l'issue de la première division embryonnaire (Figure 11) (Ryan & Saul, 1968; Callaini *et al.*, 1997; Tram *et al.*, 2006). Ce ou ces noyaux haploïdes vont continuer à se diviser permettant à l'embryogenèse de se poursuivre jusqu'à des stades tardifs de développement (Ryan & Saul, 1968; Callaini *et al.*, 1997; Duron & Weill, 2006; Tram *et al.*, 2006). Il peut donc y avoir deux issues à la première division embryonnaire : soit la chromatine paternelle est totalement exclue d'un ou des deux noyaux et un embryon haploïde sera obtenu, soit la chromatine paternelle est partiellement exclue et présente partiellement dans les deux noyaux ce qui entrainera la production d'un embryon aneuploïde. A l'issue de la première division embryonnaire, l'exclusion totale ou partielle de la chromatine paternelle résultant respectivement en la production d'embryons haploïdes ou aneuploïdes pourrait être due à une intensité plus ou moins forte des défauts cellulaires de l'IC (*i.e.* remodelage, réplication, condensation). Cette variabilité d'intensité entrainerait une activation différentielle du checkpoint associé aux fuseaux mitotiques régissant la ségrégation des chromosomes paternels (Breeuwer, 1997; Vavre *et al.*, 2001, 2002; Bordenstein *et al.*, 2003; Tram *et al.*, 2003, 2006). Ainsi, des défauts cellulaires sévères entraineraient l'activation du checkpoint du fuseau mitotique associé aux chromosomes paternels empêchant ainsi la ségrégation des chromosomes paternels et menant à la production de noyaux haploïdes. A l'inverse, des défauts moins

Box 1 : Le complexe d'espèces *Culex pipiens*

C. pipiens est un complexe d'espèces contenant deux sous espèces majeures *C. pipiens pipiens* Linnaeus 1758, *C. pipiens quinquefasciatus* Say 1823, et deux autres sous espèces dont le statut est moins clair *C. pipiens pallens* Coquillett 1889 et *C. australicus* Dobrotworsky & Drummond 1953. *C. pipiens* et *C. quinquefasciatus* sont morphologiquement discriminés sur la base de leur phallosomes et leurs hybrides présentent une forme de phallosomes intermédiaires à celles des deux parents (Barr, 1957), formes très similaire à celles de *C. p. pallens* et *C. australicus*. Ceci a conduit des auteurs à proposer que ces deux sous espèces seraient en fait des hybrides (Vinogradova E. B., 2000). Par conséquent seules les sous espèces *C. p. pipiens* et *C. p. quinquefasciatus* seront abordées par la suite. *C. p. pipiens* est présent dans les régions tempérées et possède une distribution holarctique alors que *C. p. quinquefasciatus* est observé dans les régions tropicales et subtropicales ainsi que dans les régions tempérées pendant les mois d'été. *C. p. quinquefasciatus* n'est pas capable de réaliser la diapause ce qui explique leur disparition des régions tempérées en hiver (Miller *et al.*, 1996).

La sous espèces *C. p. pipiens* possède deux formes, la forme *pipiens* et la forme *molestus* Forskal 1775 qui présentent des différences physiologiques et comportementales. La forme *pipiens* anautogène (les femelles ont besoin d'un repas sanguin pour produire leur première ponte), eurygame (l'accouplement nécessite de grands espaces pour la parade nuptiale), hétérodynome (les femelles diapauses en hiver) et ornithophile (les femelles se nourrissent préférentiellement sur les oiseaux). La forme *molestus* quant à elle est autogène (les femelles n'ont pas besoin d'un repas sanguin pour produire leur première ponte), sténogame (les accouplements peuvent se faire dans des espaces plus réduits), homodynamique (il n'y a pas de diapause en hiver) et mammophile (les femelles se nourrissent préférentiellement sur des mammifères). Dans la suite du manuscrit l'utilisation du terme *C. pipiens* fera référence aux deux sous-espèces *C. p. pipiens* et *C. p. quinquefasciatus* ainsi qu'aux deux formes *pipiens* et *molestus* de la sous espèces *C. p. pipiens*



sévères seraient insuffisants pour activer le checkpoint, ce qui entrainerait une mauvaise ségrégation des chromosomes du fait de leur mauvaise condensation et la production de noyaux aneuploïdes. Ces embryons haploïdes ne sont jamais viables chez des espèces diploïdes comme *D. simulans* et *C. pipiens* (Callaini *et al.*, 1997; Duron & Weill, 2006). Cependant, chez les guêpes parasitoïdes haplodiploïdes *N. vitripennis* et *Leptopilina heterotoma* les embryons haploïdes seront viables et donneront des mâles (Ryan & Saul, 1968; Vavre *et al.*, 2001; Bordenstein *et al.*, 2003; Tram *et al.*, 2006).

Problématiques de ma thèse portant sur le(s) mécanisme(s) cellulaire(s) de l'IC chez *Culex pipiens* :

Bien que *Wolbachia* et l'IC aient été découvertes chez *C. pipiens*, l'IC et sa diversité chez cette espèce ont uniquement été définies par l'analyse du taux d'éclosion des pontes issus d'une grande diversité de croisements. Les mécanismes cytologiques à l'origine de la mort des embryons et de cette diversité de phénotypes d'IC n'ont pas encore été étudiés. Dans ma thèse, nous avons investigué, pour la première fois, les mécanismes cellulaires responsables de la mort des embryons issus de croisements incompatibles chez *C. pipiens* afin de répondre aux questions suivantes :

- Quel est ou quels sont ces mécanismes de l'IC i) entre individus infectés et ii) entre mâles infectés et femelles non infectées ?
- Sont-ils différents de ceux observés chez *Nasonia spp.* et *Drosophila spp* ?
- La diversité des phénotypes d'IC observée chez cette espèce repose-t-elle sur une diversité de mécanismes cytologiques ?
- Est-ce le même mécanisme qui conduit au développement haploïde chez *C. pipiens* que celui décrit chez *Nasonia spp.* et *Drosophila spp.* ?

g. Les bases génétiques et moléculaires sous-jacentes à l'IC

Jusqu'en 2013, aucun candidat robuste pour les bases génétiques de l'IC n'avait été proposé. Plusieurs auteurs avaient proposé que les protéines à domaine ankyrine ou les gènes de phages pourraient être de bons candidats car i) ils sont quasiment absents des souches de *Wolbachia* mutualistes et très présents chez les souches manipulatrices de la reproduction, ii) les protéines à domaines ankyrines sont impliquées dans des interactions protéines-protéines en particulier au niveau de la régulation du cycle cellulaire, iii) les phages sont impliqués chez d'autres bactéries dans les interactions mutualistes ou parasitaires entre celles-ci et leur hôtes (Iturbe-Ormaetxe *et al.*, 2005; Sinkins *et al.*, 2005; Duron *et al.*, 2006a, 2007a; Oliver *et al.*, 2009). Puisque les *Wolbachia* sont expulsées des spermatozoïdes avec l'excès de cytoplasme lors de la spermatogenèse, les facteurs responsables de l'IC sont supposés être contenus dans les spermatozoïdes en l'absence de *Wolbachia*. La première vraie piste qui a ouvert un champ de découverte vers les mécanismes moléculaires de l'IC a été publiée en 2013 par Beckmann & Fallon. Dans cet article, les auteurs ont mis en évidence une protéine produite par *Wolbachia* en grande quantité dans le sperme de *C. pipiens* contenu dans les spermathèques de femelles non infectées préalablement croisées avec des mâles infectés. Cette protéine est codée dans le génome de *wPipPel* par le gène WPIP0282 (raccourci en 282) qui forme un opéron avec le gène WPIP0283 (raccourci en 283) (Beckmann & Fallon, 2013). Des homologues de cet opéron ont été trouvés dans les génomes des souches de *Wolbachia* *wMel*, *wRi*, *wSim*, *wVitB*, *wHa*, *wNo* et *wAlbB*, toutes capables d'induire l'IC et aucun n'a été trouvé dans les génomes des souches de *Wolbachia* mutualistes *wBm* et *wOo* infectant des nématodes filaires (Beckmann & Fallon, 2013). **Cet opéron et ses homologues sont donc apparemment associés aux souches de *Wolbachia* inductrices de l'IC, ce qui en fait des candidats robustes comme bases génétiques de l'IC. Dans les génomes de *wRi*, *wMel* et *wPip*, ces opérons sont localisés dans une zone du prophage WO et entourés par des éléments génétiques mobiles** (Beckmann & Fallon, 2013).

L'opéron 282/283 possède dans le génome de *wPipPel*, un paralogue formé par les gènes WPIP0294 (294) et WPIP0295 (295). Les analyses in-silico des séquences d'acides aminés de 282 et 294 n'ont pas permis d'obtenir d'indications précises sur les fonctions des protéines qu'ils codent. En revanche, l'analyse du repliement protéique (structure tertiaire) de ces séquences a révélé des similarités avec la structure des protéines Cullin qui font partie d'un complexe protéique impliqué dans la ligation des ubiquitines (SCF ubiquitin ligase complex). Les ubiquitines sont des petites protéines attachées à d'autres protéines et qui jouent un rôle dans la stabilité, la fonction et/ou la localisation des protéines sur lesquelles elles s'attachent (Glickman & Ciechanover, 2002; Schnell & Hicke, 2003). Un homologue particulièrement intéressant de ces protéines Cullin, l'Apc2, fait partie du complexe « anaphase-promoting » qui intervient dans la ségrégation des chromatines durant l'anaphase par des processus de protéolyses via ubiquitination (Deshaies, 1999; Beckmann & Fallon, 2013).

L'analyse des séquences d'acides aminés de 283 et 295 a révélé que 283 possède un domaine très similaire à un domaine eucaryote Ulp1C48 SUMO et 295 un domaine similaire à un domaine DUF 1703 (Beckmann & Fallon, 2013). Ainsi les protéines codées par 283 et 295 sont très probablement respectivement une déubiquitineuse (DUB) et une nucléase. Une DUB est une enzyme capable de cliver les protéines ubiquitines attachées à d'autres protéines (Reyes-Turcu *et al.*, 2009) alors qu'une nucléase est une enzyme capable de couper des acides nucléiques au niveau des liaisons phosphodiester (Nishino & Morikawa, 2002). L'opéron 294/295 est aussi un bon candidat car le domaine DUF 1703 que le gène 294 contient est présent chez un élément génétique égoïste appelé « l'élément Medea » qui induit chez *Tribolium castaneum* des phénotypes d'IC et de morts embryonnaires très similaires à ceux de *Wolbachia* (Lorenzen *et al.*, 2008).

Les fonctions protéiques prédites des gènes 282 et 283 en font de très bons candidats pour les bases de l'IC car les phénomènes d'ubiquitination et déubiquitination sont observés i) au niveau de la régulation des cycles cellulaires, ii) de l'architecture des histones, iii) de la réplication de l'ADN (PCNA), autant de mécanismes cellulaires impactés par l'IC (Shaheen *et al.*, 2010; Lee *et al.*, 2011; McAleenan *et al.*, 2012; Beckmann & Fallon, 2013; Cubeñas-Potts & Matunis, 2013). **La présence de la protéine 282 dans le sperme de *C. pipiens* ainsi que les fonctions protéiques théoriques des gènes 282 et 283 ont conduit Beckmann & Fallon (2013) à soutenir l'implication de ces gènes dans l'induction de l'IC.** Ils ont alors proposé que l'un des deux gènes de l'opéron 282/283 soit impliqué dans l'induction de l'IC (fonction *mod*) et l'autre dans l'empêchement de l'IC (fonction *resc*). Cette hypothèse est d'autant plus soutenue que la protéine 282 serait impliquée dans l'ubiquitination de protéines alors même que la protéine 283 serait hypothétiquement impliquée dans le mécanisme antagoniste qu'est la déubiquitination. Un tel système ressemble beaucoup au système toxine/antidote ou de « gènes d'addiction » souvent observé dans des éléments génétiques mobiles égoïstes (Rankin *et al.*, 2011; Yamaguchi *et al.*, 2011; Beckmann & Fallon, 2013).

Si Beckmann et Fallon avaient bien accumulé des faits pointant la potentielle implication des opérons 282/283 et 294/295 dans les mécanismes de l'IC chez *C. pipiens*, leur implication dans la diversité des phénotypes d'IC semblait peu probable du fait de la présence d'une seule copie de chacun de ces opérons dans le génome de référence *wPipPel* (Klasson *et al.*, 2008). Or, des études de modélisation à partir des croisements réalisés dans notre laboratoire avaient permis d'affirmer que la diversité de phénotypes d'IC chez *C. pipiens*, ne pouvait exister que grâce à la présence dans un même génome de *wPip* de plusieurs « gènes » ou allèles codant le *mod* et de plusieurs gènes ou allèles codant le *resc*. En effet, si on considère un cas simple avec deux souches A et B toutes les deux incompatibles avec la souche C, dans ce cas-là, un seul gène partagé par A et B permet d'expliquer ce phénotype. En revanche, si on rajoute une souche D compatible avec A mais pas avec B, ces phénotypes d'IC ne peuvent pas être expliqués par la présence d'un seul gène (deux gènes minimum ou deux allèles de ce gène sont alors nécessaires). Ainsi, la diversité des patrons d'IC chez *C. pipiens* se devait d'être codée

i) soit par des gènes *mod* avec des fonctions différentes, ii) soit des copies différentes d'un même gène *mod*. Le même raisonnement pouvant s'appliquer pour le resc (Atyame *et al.*, 2011b, 2014; Nor *et al.*, 2013). Ce qui implique que les *wPip* infectant les *C. pipiens* et qui induisent phénotypes d'IC différents doivent posséder des copies polymorphes de ces gènes différentes (Atyame *et al.*, 2014).

Problématiques de ma thèse portant sur les bases génétiques de l'induction et de la diversité de l'IC chez *C. pipiens* :

Lors de ma thèse, nous avons étudié à l'aide d'approches génomiques et génétiques les opérons 282/283 et 294/295 de souches de *Wolbachia*, issues de populations naturelles à travers le monde, induisant des phénotypes d'IC différents afin d'explorer :

- **Le rôle potentiel de ces gènes dans le mécanisme de l'IC chez *C. pipiens*.**
- **Leur implication dans l'unique diversité des phénotypes d'IC chez cette espèce.**

Partie I :

**Les bases génétiques de l'IC et de sa
diversité chez *Culex pipiens***

I. Introduction

L'IC induite par *Wolbachia* chez ses hôtes arthropodes se produit soit lorsque les mâles infectés copulent avec des femelles non infectées soit lorsque des individus infectés par des souches de *Wolbachia* incompatibles se croisent entre eux. Dans le deuxième cas, on peut voir apparaître des relations d'incompatibilité complexes entre les individus d'une même espèce. L'exemple emblématique de ce type d'IC est le moustique *C. pipiens* où la plus grande diversité de phénotypes d'IC a été décrite à ce jour. Le phénotype d'IC également appelé cytotype d'une lignée de *C. pipiens* infectée par une souche de *Wolbachia* donnée se caractérise par deux profils : le profil *mod* associé à la capacité des mâles de cette lignée à être incompatibles ou non envers les femelles d'autres lignées et le profil *resc* associé à la capacité des femelles de cette même lignée à restaurer la fertilité ou non lorsqu'elles sont croisées avec des mâles d'autres lignées. Grâce à de nombreuses études de croisements réalisées depuis les années 1950 entre *C. pipiens* provenant de l'ensemble de la planète, il a été décrit plus de 40 cytotypes différents (Laven, 1967b; Duron *et al.*, 2006b; Atyame *et al.*, 2014). Des études de modélisation ont mis en évidence qu'une telle diversité de cytotypes ne pouvait être expliquée que par la présence de plusieurs facteurs *mod* et plusieurs facteurs *resc* au sein d'une même souche de *wPip* (Nor *et al.*, 2012, 2013). Au sein d'un même génome, ces facteurs *mod* et *resc* peuvent être en théorie soit i) plusieurs gènes différents impliqués dans la fonction *mod* et plusieurs gènes différents impliqués dans la fonction *resc* soit ii) plusieurs copies polymorphes d'un même gène *mod* et plusieurs copies polymorphes d'un même gène *resc* (Nor *et al.*, 2013).

Afin de déterminer les gènes de *Wolbachia* impliqués dans la fonction *mod* et dans la fonction *resc*, de nombreuses études de génomique comparative entre génomes de *Wolbachia* inductrices d'IC et de *Wolbachia* mutualistes ont été réalisées. Ces études ont permis la mise en évidence i) de nombreux gènes codant pour des protéines à domaine ankyrine et ii) de gènes de phages présents dans les génomes des *Wolbachia* parasites de la reproduction mais absents ou en nombre très faible dans ceux des *Wolbachia* mutualistes (Wu *et al.*, 2004; Foster *et al.*, 2005; Klasson *et al.*, 2008, 2009). Ces deux types de gènes ont constitué les premiers candidats potentiels aux bases génétiques de l'IC. Cependant, aucune corrélation claire n'a été observée entre ces gènes et les phénotypes d'IC, en particulier ceux observés chez *C. pipiens* ne permettant pas de confirmer leur implication dans l'IC (Duron *et al.*, 2006b, 2007a). Une réelle avancée dans la compréhension des bases génétiques de l'IC est liée à la découverte d'une protéine de *Wolbachia* dans le sperme de *C. pipiens* (Beckmann & Fallon, 2013). Cette protéine est codée par le gène WPIP0282 du génome de *wPipPel* qui est associé en opéron avec le gène WPIP0283 (Klasson *et al.*, 2008; Beckmann & Fallon, 2013). Dans le génome de *wPipPel*, cet opéron est présent en une copie et possède un paralogue nommé 294/295 (Beckmann & Fallon, 2013).

Au cours de ma thèse, des preuves fonctionnelles, acquises chez la drosophile, de l'implication de l'opéron 282/283 dans l'IC ont été apportées par Beckmann *et al.*, 2017, LePage *et al.*, 2017, Shropshire

et al., 2018 et Shropshire *et al.* (2018). En revanche, aucune expérience de génétique fonctionnelle n'a pour le moment mis en évidence le rôle de l'opéron 294/295 dans l'induction de l'IC. Afin de simplifier les nomenclatures, les gènes WPIP0282 et WPIP0283 ont été nommés *cidA* et *cidB* pour « CI-inducing deubiquitinase » du fait de la fonction déubiquitinase (DUB) de *CidB* et WPIP0294 et WPIP0295 sont devenus *cinA* et *cinB* pour « CI-inducing nucléase » du fait de la fonction nucléase de *cinB* (Beckmann *et al.*, 2017). L'expression par transgénèse de l'opéron *cidA/cidB* provenant soit du génome de *wPip* (*cidA^{wPip}/cidB^{wPip}*) soit du génome de *wMel* (*cidA^{wMel}/cidB^{wMel}* aussi appelé *cifA/cifB*) dans des mâles *D. melanogaster* non infectés induit de l'IC lorsqu'ils sont croisés avec des femelles non infectées (Beckmann *et al.*, 2017; LePage *et al.*, 2017). Lorsque le domaine enzymatique DUB de *cidB^{wPip}* est rendu non fonctionnel par mutation, les mâles *D. melanogaster* transfectés n'induisent plus d'IC lorsqu'ils sont croisés avec des femelles non infectées (Beckmann *et al.*, 2017). Ces résultats soutiennent très fortement l'implication de *cidB* dans la fonction *mod*. En revanche, l'implication de *cidA* dans la fonction *mod* ou *resc* n'est pas encore claire. En effet, l'expression conjointe de *cidA^{wMel}* et de *cidB^{wMel}* dans les mâles transgéniques est nécessaire à l'induction de l'IC ce qui supporte le rôle de *cidA* dans la fonction *mod* chez la drosophile (LePage *et al.*, 2017). Cependant, l'expression par transgénèse de *cidA^{wMel}* chez des femelles non infectées croisées avec des mâles infectés permet de restaurer la fertilité du croisement, ce qui suggère fortement l'implication de *cidA^{wMel}* dans la fonction *resc* (Shropshire *et al.*, 2018). Enfin, la capacité de *cidA^{wPip}* à contrer la toxicité de *cidB^{wPip}* chez la levure supporte aussi l'implication de *cidA* dans la fonction *resc* (Beckmann *et al.*, 2017). Ainsi, à l'heure actuelle, dans la littérature, deux modèles s'opposent quant au rôle du gène *cidA* dans le fonctionnement de l'IC. Beckmann *et al.* (2013 & 2017) proposent un modèle toxine-antidote simple où *cidB* serait impliqué dans la fonction *mod* et *cidA* dans la fonction *resc* alors que LePage *et al.* (2017) ainsi que Shropshire *et al.* (2018) proposent un modèle dans lequel *cidB* est impliqué dans la fonction *mod* mais où *cidA* est impliqué à la fois dans la fonction *mod* quand exprimé chez les mâles et dans la fonction *resc* quand exprimé chez les femelles. La fonction de *cidA* est dans ce cas qualifiée de « deux en un » (Shropshire *et al.*, 2018).

Chez *C. pipiens*, les opérons *cidA/cidB* et *cinA/cinB* sont présents en une seule copie dans l'assemblage du génome de *wPipPel* qui constitue à l'heure actuelle le seul génome de référence des *Wolbachia* infectant *C. pipiens* (Klasson *et al.*, 2008). Même si l'expression par transgénèse de l'opéron *cidA^{wPip}/cidB^{wPip}* chez des mâles *D. melanogaster* permet d'induire l'IC lorsque ces mâles sont croisés avec des femelles non infectées (Beckmann *et al.*, 2017), le fait qu'il soit présent en une seule copie dans le génome de *wPipPel* alors que plusieurs facteurs *mod* et *resc* sont requis pour expliquer la diversité des phénotypes d'IC observée chez *C. pipiens* (Atyame *et al.*, 2011b, 2014; Nor *et al.*, 2013) rend, *a priori*, son implication dans cette diversité peu probable. Cette même observation peut être appliquée à l'opéron *cinA^{wPip}/cinB^{wPip}*.

Pour résoudre ce paradoxe, nous avons dans cette première partie, étudié la structure de l'opéron *cidA^{wPip}/cidB^{wPip}* ainsi que de son paralogue *cinA^{wPip}/cinB^{wPip}* dans les génomes de différentes souches de wPips infectant plusieurs de nos lignées de *C. pipiens* de laboratoire afin de statuer sur sa potentielle implication dans la diversité des phénotypes d'IC ainsi que sur le rôle respectif de *cidA* et de *cidB* dans l'IC chez *C. pipiens*.

Chapitre I :

La diversité des phénotypes d'IC observée chez le moustique *Culex pipiens* est gouvernée par un opéron amplifié et polymorphe présent dans le génome de *Wolbachia*

Article n°1 publié dans Nature Communications

Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia

Manon Bonneau¹, Celestine Atyame^{1,2}, Marwa Beji³, Fabienne Justy¹, Martin Cohen-Gonsaud⁴, Mathieu Sicard^{1‡} and Mylène Weill^{1‡*} (‡ These authors jointly supervised this work)

1- Institut des Sciences de l'Evolution de Montpellier (ISEM), UMR CNRS-IRD-EPHE-Université de Montpellier, Place Eugène Bataillon, 34095 Montpellier, CEDEX 5, France.

2- Processus Infectieux en Milieu Insulaire Tropical (PIMIT), UMR CNRS-INSERM-IRD-Université de La Réunion, Sainte-Clotilde, Ile de La Réunion, France.

3- Institut Pasteur Tunis, Laboratory of Epidemiology and Veterinary Microbiology, University of Tunis El Manar, Tunisia

4- Centre de Biochimie Structurale (CBS), UMR CNRS-INSERM-Université de Montpellier, 29 rue de Navacelles 34090 MONTPELLIER - France

*corresponding author: mylene.weill@umontpellier.fr

Abstract

Culex pipiens mosquitoes are infected with *Wolbachia* (*w*Pip) that cause an important diversity of cytoplasmic incompatibilities (CI). Functional transgenic studies have implicated the *cidA-cidB* operon from *w*Pip and its homolog in *w*Mel in CI between infected *Drosophila* males and uninfected females. However, the genetic basis of the CI diversity induced by different *Wolbachia* strains was unknown. We show here that the remarkable diversity of CI in the *C. pipiens* complex is due to the presence, in all tested *w*Pip genomes, of several copies of the *cidA-cidB* operon, which undergoes diversification through recombination events. In 183 isofemale lines of *C. pipiens* collected worldwide, specific variations of the *cidA-cidB* gene repertoires are found to match crossing types. The diversification of *cidA-cidB* is consistent with the hypothesis of a toxin-antitoxin system in which the gene *cidB* co-diversifies with the gene *cidA*, particularly in putative domains of reciprocal interactions.

Introduction

The most common way by which *Wolbachia* bacteria spread within insect populations is CI which results in the early death of embryos¹ when males infected with a given *Wolbachia* strain mate with uninfected females or females infected with an incompatible strain. The most popular model to date for conceptualizing CI involves the secretion, by *Wolbachia*, of a “modification” factor (mod, or toxin) in the sperm that impairs early embryogenesis, and a *Wolbachia* “rescue” factor (resc, or antitoxin) produced in the oocyte which allow for a viable, diploid zygote to develop if the cross is compatible^{2,3,4}. Previous studies searching for *Wolbachia* effectors involved in CI by genomic comparison between mutualistic and manipulative *Wolbachia* revealed a higher number of genes with ankyrin repeats in the latter^{5,6}. Ankyrin repeat genes were thus considered as potential mod and resc candidates due to their role in protein-protein interactions and cellular cycle regulation⁷. However, to date no correlation between the distributions of any ankyrin repeat genes and CI patterns was demonstrated⁸. The first convincing candidate gene for involvement in CI encodes a protein called CidA

(WP0282), secreted by *Wolbachia* into *C. pipiens* sperm⁹. The *cidA* gene is part of an operon also containing a second gene, *cidB* (WP0283). CidA and CidB proteins directly interact with each other as demonstrated by pull-down experiments¹⁰. The *cidA* and *cidB* genes (named *cifA* and *cifB* in *wMel*¹¹) are associated with the prophage WO modules¹¹, and paralogs of these genes have been found in all published genomes from *Wolbachia* strains known to induce CI in insects^{9,11,12}. The co-expression of *cidA/cidB* (or *cifA/cifB*) transgenes in *D. melanogaster* reproduces the disturbance of the first embryonic division hallmark of CI, leading to embryo death^{10,11}. In this context, disruption of the deubiquitylating (DUB) domain of CidB restores normal embryogenesis, demonstrating the contribution of this domain to the “mod” factor function¹⁰. The role of CidA in the mod-resc CI system is much more debated. Data obtained on yeast suggest that CidA may rescue CidB¹⁰, but CifA has been described as a CifB elicitor in *D. melanogaster*¹¹.

cidA-cidB operon is involved in the CI induced by infected males in the progenies of uninfected females^{10,11}, but the genetic basis of CI diversity (i.e. compatibility, uni or bidirectional incompatibility) between hosts infected with different *Wolbachia* strains has yet to be investigated. We addressed this question in *C. pipiens*, a powerful model in which hundreds of crosses between lines sampled worldwide have revealed unprecedented CI diversity^{13,14,15}. This unique CI diversity is solely governed by *wPip* diversification in this species complex since no other manipulative endosymbionts are present and no host genetic background influence has been yet demonstrated^{13,15}. All *C. pipiens* individuals are infected with *Wolbachia*, which recently diverged into five distinct phylogenetic groups (*wPipI* to *wPipV*)¹⁶. The *wPip* strains of the same group generally ensure compatibility between their hosts, whereas intergroup crosses are more likely to be incompatible¹³. An analysis of multiple crosses concluded that each *Wolbachia* (*wPip*) genome must contain several mod and resc factors to account for the high diversity of CI in *C. pipiens*^{15,17}. These multiple mod and resc factors could theoretically be encoded by different copies (i.e. variants) of same mod/resc genes or different mod and resc genes within the same *wPip* genome^{15,17}. It therefore seems likely that (i) different copies (i.e. variants) of *cidA-cidB*, (ii) paralogs of these genes resulting from ancient duplication events^{10,11,17} or (iii) totally unrelated genes¹³, could be involved in CI diversity in *Culex pipiens*.

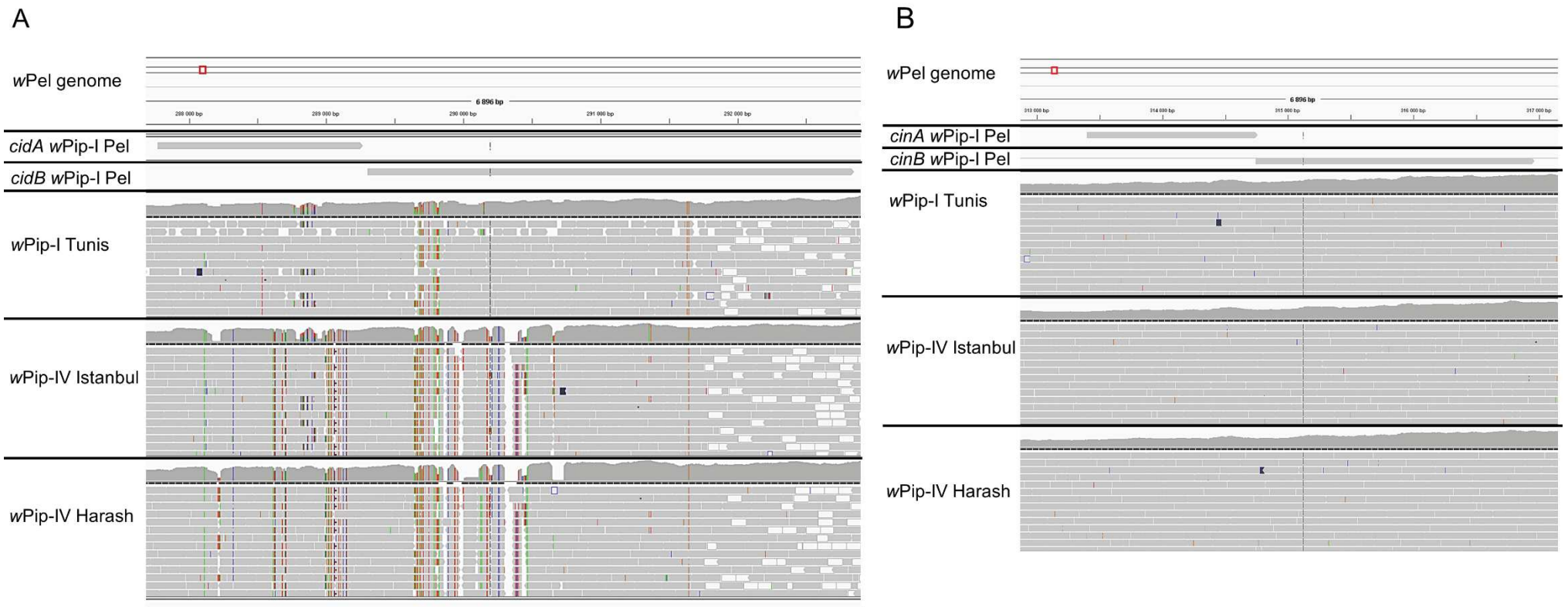


Figure 1 | The *cidA/cidB* operon displays polymorphism within and between different *wPip* strains, whereas no polymorphism is observed for the *cinA/cinB* operon. Mapping onto the reference genome *wPip_Pel* of the Illumina reads from the Tunis line infected with *wPipI*, and from the Harash and Istanbul lines infected with *wPipIV*. Colored residues are different from those in the reference *wPip_Pel* sequence. The pattern above the dotted lines represents the number of Illumina reads that have mapped for each position. The red box on the *wPip_Pel* reference genome allows the IGV user to locate the zone that is visualized with more details on the bottom panels.

A. Mapping of the *cidA/cidB* operon reads onto the *wPip_Pel* reference genome. Polymorphism between *wPip* groups was detected (differences between Harash and Istanbul infected with *wPipIV*, Tunis and *wPip_Pel* infected with *wPipI*). Polymorphism was also detected within *Wolbachia* groups (variations observed between Tunis and *wPip_Pel* or between Harash and Istanbul). The operon was also found to be polymorphic within the same isofemale line of *C. pipiens*.

B. Mapping of the *cinA/cinB* operon reads onto the *wPip_Pel* reference genome. No polymorphism was detected between reads from the three isofemale lines and the reference line.

We thus investigated the contribution of polymorphism of the *cidA-cidB* operon and its *cinA-cidB* paralog with predicted nuclease activity¹⁰ to CI diversity in *C. pipiens* with the major hypothesis that if these genes encode for CI diversity determinants, they should be different between *C. pipiens* lines with different crossing types. We show that all tested *wPip* genomes contain several variants of the *cidA-cidB* operon while it is not the case for *cinA-cinB* operon. The *cidA-cidB* repertoires of variants differ between *wPip* strains from different groups exhibiting different crossing types. In 180 isofemale lines of *C. pipiens* exclusively infected with *wPip* IV strains, specific variations of the *cidA-cidB* gene repertoires are found to match crossing type variations. These variations occurred only in specific domains for both proteins which consist exclusively of protein-protein interaction motifs and thus might be involved in the physical interaction between CidA and CidB.

Results

Illumina sequencing of several *wPip* strains

We first used Illumina technology to sequence three *wPip* genomes from strains inducing different crossing types¹³: two strains from group IV, originating from Algeria and Turkey, and one strain from group I, originating from Tunisia. The mapping of the reads onto the reference genome *wPip_Pel*¹⁸ revealed an abnormally high coverage of the *WOPip1* region, which includes the *cidA* and *cidB* genes. The coverage for this region was, on average, three times greater than that for 14 single-copy housekeeping genes^{19,20} (Supplementary Fig. 1). Moreover, the *cidA-cidB* operon was highly variable between and within the three isofemale lines sequenced (Fig. 1A). Contrary to the assembly of *wPip_Pel* genome which exhibits only one copy of *cidA-cidB* operon¹⁸, our findings suggest the occurrence of several copies of the *cidA-cidB* operon within each *wPip* genome from Algeria, Tunisia and Turkey. Consistent with this observation, the Sanger sequences of *cidA* and *cidB* obtained from single individuals displayed multiple overlaps (Fig. 2A). No such diversity was found in *wMel* and *wRi* from *Drosophila melanogaster* and *D. simulans* (Fig. 2B, 2C). The *cinA* and *cinB* paralogs, which are also present in all sequenced *wPip* genomes and have been identified as possibly involved in incompatibility between *C. pipiens* lines¹⁰, displayed no polymorphism within and between *wPip* strains (Fig. 1B), ruling out their role in CI diversity¹⁰. These findings suggest a previously unsuspected

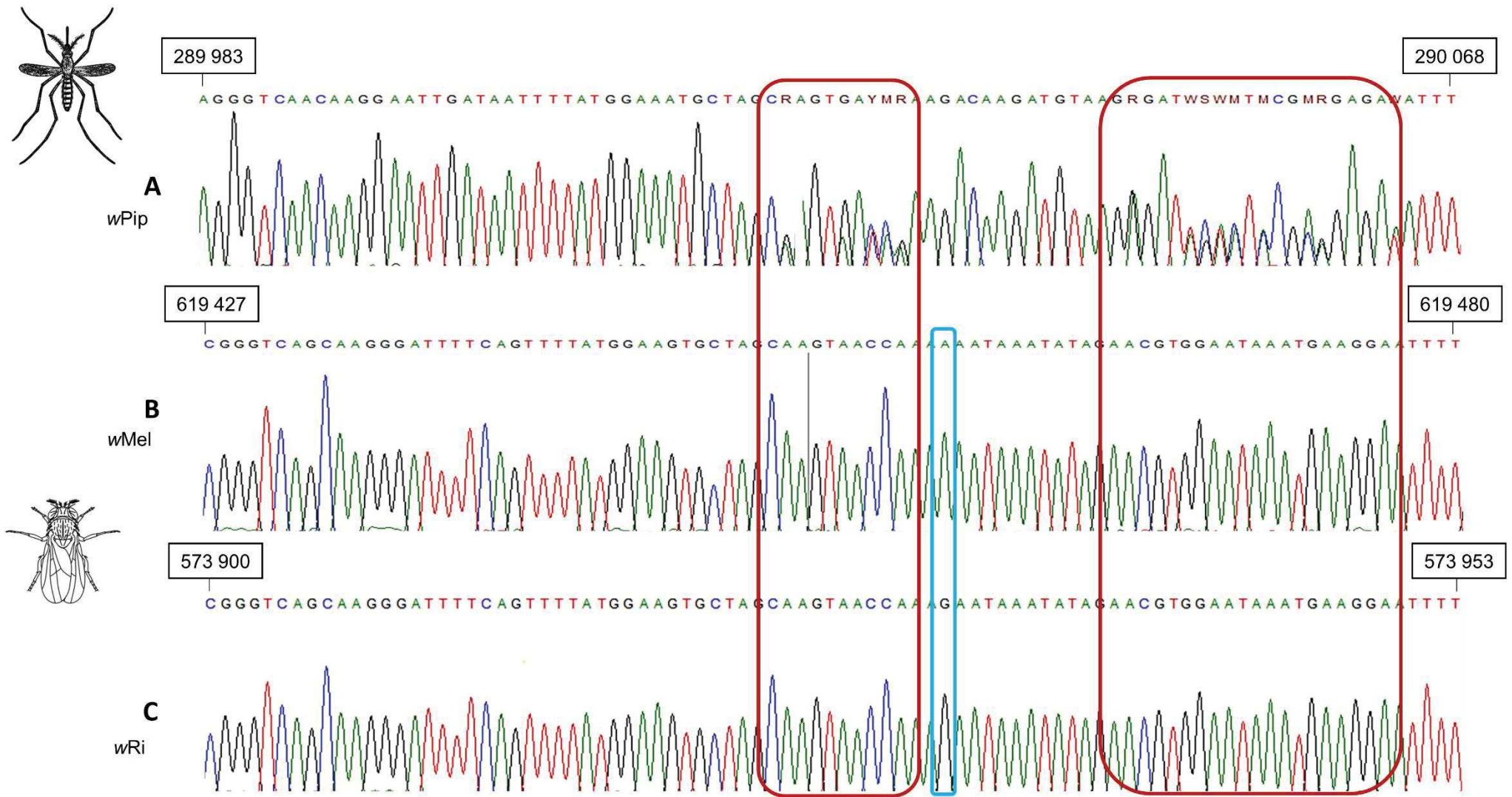


Figure 2 | Sanger sequencing electrophoregrams of homologous regions of (A) *cidB* in *wPip* (B) *cifB* in *wMel*, (C) and *cifB* in *wRi*.

The numbers in the black boxes indicate the positions of these regions in the *wMel*, *wRi* and *wPip_Pel* genomes. The blue rectangle highlights the only polymorphism between *wRi* and *wMel*. Unlike the *Wolbachia* strains of *Drosophila*, *wPip* strains consistently gave mixed signals (showed in red rectangles), suggesting the presence of at least two different sequences of this gene in DNA samples. A similar pattern was observed when *cidA* was compared with *cifA*.

diversity of *cidA-cidB* copies within *wPip* genomes, resulting from gene amplifications and divergence that may account for CI diversification in *C. pipiens*.

High *cidA-cidB* diversity between *wPip* groups

We explored the association between *cidA-cidB* genetic variants and CI diversity in *C. pipiens* further, by investigating the variability of these genes in four isofemale lines differing by their geographical origin, each infected with a *wPip* strain from a different group (I to IV) and exhibiting a different crossing type (Table 1). Each of these isofemale lines was founded with one initial egg-raft from a single female. Cloning and sequencing revealed a large number of *cidA-cidB* variants in each of these *wPip* strains. The variable domains that differ between variants were not contiguous but restricted to some specific zones of the proteins (Fig. 3A and 4A). Within a single line there were up to six different variants of *cidA* (Fig. 3B), and up to four different variants of *cidB* (Fig. 4B). Some specificity in the *cidA-cidB* repertoire was also detected, with none of the variants shared between strains from different *wPip* groups (Fig. 3B and 4B). Network analyses of *cidA* and *cidB* suggested that the differences between copies stemmed mostly from block rearrangements within and/or between *wPip* genomes (Fig. 3C and 4C). These rearrangements were never observed in the region corresponding to the CidB catalytic domain, but in other regions of the CidA and CidB proteins (Fig. 3, 4; Supplementary Fig. 2 and 3). The stability of the *cidA-cidB* repertoire over a decade (between 2006 and 2017) in addition with the stability in the *cidA* and *cidB* copy numbers relative to *wsp* gene (known to be present as a single copy in all *Wolbachia* genomes) inferred by q-PCR in the Istanbul *C. pipiens* line are convincing elements that the *cidA-cidB* genes are present in multiple copies in *wPip* genome (data not shown). The copy-number ratio between *cidA-cidB* genes and *wsp* was always between 4 and 6 (Table 2). We also found that all copies of these genes were transcribed. Indeed, in the *wPip* strain from Istanbul the transcripts of the six polymorphic copies (i.e. variants) of *cidA* and the four polymorphic copies of *cidB* were detected in RNA extract after RT-PCR, cloning and sequencing (data not shown).

Table 1 | Crossing types of the four mosquito lines each infected with *Wolbachia* *wPip* from different groups (I, II, III, IV).

♀ \ ♂	Tunis	Lavar	Maclo	Istanbul	Crossing type
Tunis (<i>wPip</i>I)	+	-	+	-	A
Lavar (<i>wPip</i>II)	+	+	+	-	B
Maclo (<i>wPip</i>III)	+	+	+	+	C
Istanbul (<i>wPip</i>IV)	-	-	-	+	D

+ means that the cross is compatible (many larvae hatch from the egg rafts) and – indicates an incompatible cross (no larvae hatch from the egg rafts). Each isofemale line has a different crossing type, as deduced from reciprocal crosses between the four lines.

Table 2 | Quantification of *cidA-cidB* operon in different *wPip* genomes.

Line \ Ratio	Tunis <i>wPip</i> -I	Lavar <i>wPip</i> -II	Maclo <i>wPip</i> -III	Istanbul <i>wPip</i> -IV
<i>cidA/wsp</i>	5,99 ± 0,77	5,40 ± 0,84	5,1 ± 0,41	5,62 ± 0,72
<i>cidB/wsp</i>	6,21 ± 1,40	5,06 ± 1,46	4,17 ± 0,64	6,13 ± 1,26
<i>cidA/cidB</i>	0,99 ± 0,21	1,10 ± 0,19	1,24 ± 0,21	0,95 ± 0,19

The number of copies of the *cidA* and *cidB* genes was estimated by real-time quantitative PCR, relative to the reference gene *wsp* known to be present as a single copy per genome. The number of copies was determined by calculating the ratio between *cidA* (or *cidB*) signals and *wsp* signals on four individuals per line. ± means standard deviation. The *cidA* to *cidB* ratio was always close to one. The numbers of *cidA* and *cidB* copies present per *Wolbachia* genome were therefore similar, as expected for an operon.

***cidA-cidB* repertoire variations are linked to crossing type**

We then focused on a simpler situation, by analyzing CI diversity within *wPip* group IV. We used six isofemale lines with two different crossing types. These lines had the same rescue properties but two different mods¹³. Males from “incompatible” lines sterilize females whenever they host a *wPip* group I, II or III strain, whereas males from “compatible” lines always produce progenies with these females (Fig. 5). We compared the *cidA-cidB* repertoires of “compatible” and “incompatible” lines. We found that one CidA variant [*cidA_IV*(α)] was present in all *wPip*IV lines, regardless of their crossing type or geographic origin (Fig. 5). Such a ubiquitous distribution is expected for the resc factor mediating compatibility between all isofemale lines hosting *wPip* strains from the same group. By contrast, no CidB variant was shared between compatible and incompatible lines (Fig. 5). Strikingly, *cidA* and *cidB* repertoire variations were clearly associated with crossing type variations, as *cidA_IV*(δ) and *cidB_IV*(a/2) variants were found uniquely in “incompatible” lines (Fig. 5). This may suggest that both *cidA_IV*(δ) and *cidB_IV*(a/2) are required for the incompatibility phenotype (Fig. 5).

We further investigated the distribution of these two specific *cidA/cidB* variants [*cidA_IV*(δ) and *cidB_IV*(a/2)] at a larger scale, by performing diagnostic RFLP tests on 180 isofemale lines from 15 populations [13 in Algeria and Tunisia; one each in China and Turkey], for which we determined crossing types. A very strong association was established between the presence of both *cidA_IV*(δ) and *cidB_IV*(a/2) in “incompatible” isofemale lines and their absence from “compatible” isofemale lines: all 17 “incompatible” lines had both these variants, whereas 147 of the 163 “compatible” lines had neither of them ($\chi^2= 78$, $df = 1$, $p < 2.2e-16$). Among the 16 “compatible” isofemale that appeared discordant: (i) eight were truly discordant as they display both *cidA_IV*(δ) and *cidB_IV*(a/2) while being compatible but (ii) eight were only discordant for *cidA* as they displayed *cidA_IV*(δ) but lacked *cidB_IV*(a/2) (as confirmed by cloning and sequencing ; Supplementary Data 1). To summarize the data: the absence of these variants is always associated with a “compatible” crossing type while its presence is mostly associated with “incompatible” crossing type with few discrepancies.

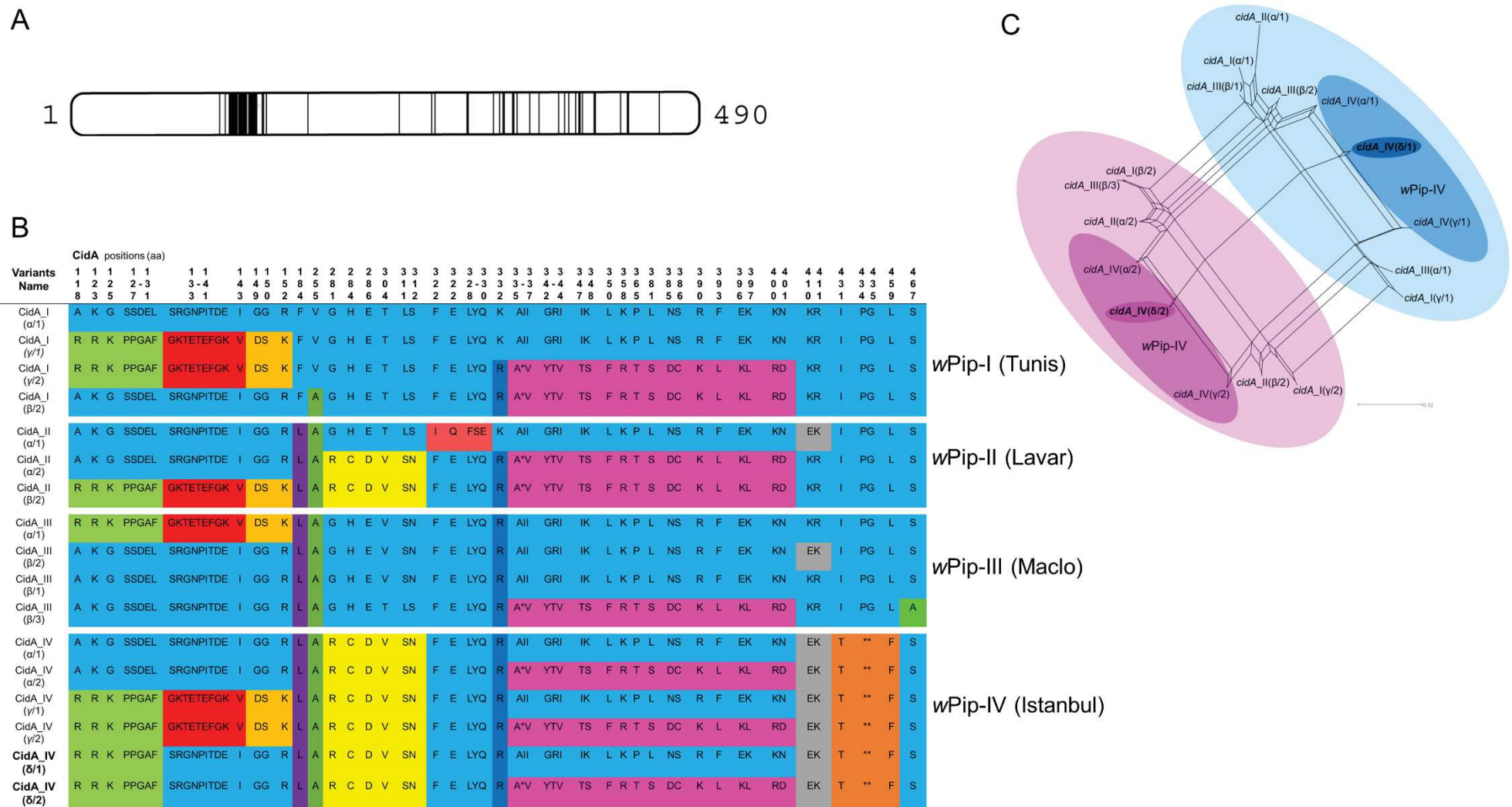


Figure 3 | The repertoire of CidA variants in the different wPip groups

A. Localization of polymorphic zones within CidA proteins. Schematic representation of the CidA protein with polymorphic zones between CidA variants of the four wPip groups highlighted in black.

B. The polymorphism between CidA variants is distributed as blocks of variable amino acids. Protein sequences alignment of the CidA variants found in the four *Wolbachia* strains Tunis, Lavar, Maclo and Istanbul. The first sequence is used as a reference to determine the polymorphic region. For more clarity, only polymorphic positions of the alignment are represented and amino acid positions are not contiguous. When more than two contiguous amino acid were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic wPip group (I to IV). However, no variant (i.e. complete CidA sequence) is common to wPip strains from different groups.

C. cidA variants result from block recombination. Each edge (or set of parallel edges) corresponds to a split in the dataset and has a length equal to the weight of the split. Incompatible splits produced by recombination are represented by boxes in the network. Most of the *cidA* variants are connected by multiple pathways resulting from block recombination between them. The largest circles highlight the two groups of *cidA* variants sharing the same amino-acid sequence from position 336 to 401 (panel B). Intermediate darker circles highlight *cidA* variants observed in the wPipIV group. The smallest bold circles highlight the two versions (1 and 2) of the *cidA*_IV(δ) variant matching the “incompatible” crossing type of mosquito lines infected with wPipIV (see the text).

Putative ankyrin interaction domains between CidA and CidB

The analysis of CidA and CidB polymorphic regions revealed that the sequence variations associated with crossing type variations are restricted to about 30 amino acids positions, for both proteins (Fig 6A). Extensive fold-recognition analysis of the CidA and CidB proteins were performed using the @tome2 in-house server²¹. Except for the deubiquitylating protease domain, no tridimensional model could be firmly obtained. Nevertheless, both CidA and CidB exhibit sequence identity homology (under 20%) with protein-protein interaction domains with repeated structure. Careful examination of hydrophobic patterns suggests internal helical structure repetitions as found in ankyrin- or HEAT-repeats²²). The size of ~33 amino acids is the universal expected size for tandem ankyrin repeat domains⁶. However, due to the lack of available homolog sequences in databases, it is difficult to determine the helical structure repetition boundaries in both N- and C- terminal part of the protein domains. As the protein variations matching with crossing type variations occur in putative ankyrin domains, we hypothesize that CidA and CidB co-diversify in these domains because they are involved in their reciprocal interactions (Fig. 6).

Discussion

In the *C. pipiens* mosquitoes, the genetic architecture of the *Wolbachia cidA-cidB* operon is characterized by variable multi-copies associated to the unique CI diversification observed in this species. These different copies of *cidA* and *cidB* likely result from genic amplification followed by diversification since we revealed a high diversity within each *wPip* strain. As these genes are associated with the prophage WO modules¹¹, the number of *cidA-cidB* copies assessed by q-PCR may not reflect the exact number of *cidA-cidB* copies in a single *wPip* genome since viral particle multiplication resulting from lytic phage activity can influence the accuracy of this quantification²³. However, the diversity of *cidA-cidB* sequences in a single mosquito, the geographical and temporal stability of sequences found in *wPip* IV group clearly point out that several copies of the operon do co-exist within the same *wPip* genome. As *cidA* and *cidB* are amplified in all *wPip* genomes from each *C. pipiens* individual we sequenced so far, the presence of only one copy of the *cidA/cidB* operon in *wPip_Pel* seems to result from an error during assembly likely caused by too many repeats. The reason(s) why

such amplification and diversification of *cidA/cidB* has evolved in *wPip* is still an open question. However, this diversity likely reflects high levels of intra- or intergenome recombination, which may be favored by multiple infections with different *wPip* and/or promoted by lytic phage activity in *C. pipiens*²⁴.

In a more global context, our results provide insights into the respective roles of CidA and CidB in the mod/resc model. Indeed, the association between *cidA_IV*(δ)-*cidB_IV*(a/2) presence/absence and incompatible/compatible crossing types, together with the results of functional studies^{10,11}, demonstrate that CidB is involved in the mod function responsible for the variation in crossing types. However, some discrepancies were found in few isofemale lines which harbor *cidA_IV*(δ) and/or *cidB_IV*(a/2) but are “compatible”. These discrepancies suggest that either (i) our strategy, involving larval field sampling, the establishment of isofemale lines and crossing experiments, generated a few errors or (ii) other genetic or epigenetic mechanisms may disrupt the relationship between genomic potential and the resulting crossing type.

The role of CidA is still unclear. The association of *cidA_IV*(δ) with variation in crossing types does not necessarily implicate CidA in the mod function. Indeed, as CidA and CidB interact with each other¹⁰, their co-diversification could result in specific CidA variations whenever involved in the mod (i.e. toxin) or in resc (i.e. antitoxin) function. Several lines of evidence suggest that CidA could be the antitoxin of CidB and that the *cidA-cidB* operon encodes a classic toxin-antitoxin system: (i) the simple two-gene structure of the operon, with the putative antitoxin gene located upstream from the putative toxin gene⁹, (ii) the overexpression of the antitoxin gene relative to the toxin gene¹¹, (iii) the ability of the two proteins to form a complex¹⁰ and (iv) the required co-expression of *cidA* and *cidB* for the production of live transgenic *D. melanogaster* and *Saccharomyces cerevisiae*¹⁰. Additional argument of CidA being an antitoxin is provided by the presence of *cidA_IV*(α), a ubiquitous CidA variant in *wPip* strains within the same group *wPipIV*, potentially accounting for their hosts’ compatibility. Moreover, if CidA is indeed the antitoxin, then a new toxin (i.e. CidB) produced by block recombination between variants would be positively selected only if a compatible CidA, capable of interaction with this new toxin, was already present in the same *wPip* genome to prevent its producer from being killed^{2,25}.

A	Line Name	Crossing type	Polymorphism location		<i>cidA_IV</i>			
			Upstream	α	β	γ	δ	
			Downstream	1/2	1/2	1/2	1/2	
	Istanbul	Incompatible		P	A	P	P	
	Hang Zhou	Incompatible		P	A	A/P	P	
	Ichkeul 09	Incompatible		P	A	A	P	
	Ichkeul 21	Incompatible		P	A	A	P	
	Harash	compatible		P	P	A	A	
	Ichkeul 13	compatible		P	P/A	A	A	

B	Line Name	Crossing type	<i>cidB_IV</i>					
			a1	a2	a3	b1	b2	b3
	Istanbul	Incompatible	P	P	A	P	P	A
	Hang Zhou	Incompatible	P	P	P	P	A	P
	Ichkeul 09	Incompatible	P	P	A	A	A	A
	Ichkeul 21	Incompatible	P	P	P	P	A	P
	Harash	compatible	P	A	P	P	A	P
	Ichkeul 13	compatible	A	A	A	P	A	P

Figure 5 | *cidA_IV*(δ) and *cidB_IV*(a/2) are present specifically in lines infected with *wPipIV* strains displaying the “incompatible” crossing type. The “incompatible” or “compatible” crossing type of the *C. pipiens* lines infected with *wPipIV* strains were determined by crossing males with females infected with *wPipI*, *wPipII* or *wPipIII*. A “compatible” type corresponds to a line in which all males are compatible when crossed with all tested females, and an “incompatible” type corresponds to a line in which males are incompatible when crossed with all tested females. The column of the two variants *cidA_IV*(δ) and *cidB_IV*(a/2) associated with the “incompatible” crossing type were shaded in green (presence) and grey (absence).

A. Distribution of *cidA* variants (P for present, A for absent) in six *wPip-IV* strains. Each column corresponds to the specific upstream polymorphic region named α , β , γ or δ . The downstream polymorphism is represented by either the sequence 1 or the sequence 2 (1/2). For a given upstream sequence P/A means that the strain exhibits the sequence 1 but not the sequence 2. On the opposite A/P means that the strain exhibits the sequence 2 but not the sequence 1.

B. Distribution of *cidB* variants (P for present, A for absent) in six *wPip-IV* strains.

This is coherent with our results since we found more *cidA* variants than *cidB* variants in the same isofemale line (except for Tunis which has as many *cidA* as *cidB*).

Most toxin-antitoxin systems require an interaction between both proteins²⁵. Pull-down experiments have demonstrated that CidA can interact with CidB¹⁰. Our analysis of CidA-CidB repertoires in *wPipIV* group showed that genetic variations linked to crossing type variations occurred only in specific domains of about ~ 30 amino acids in length, for both proteins (Fig. 6A). These domains consist exclusively of protein-protein interaction motifs (ankyrin or HEAT repeats) and thus might be involved in the reciprocal interaction between CidA and CidB. The variability of these domains that co-diversify may account for specific interaction between one variant of CidA and one variant of CidB. Such variability in interactions between CidA and CidB could drive the unrivalled diversity of crossing types in *C. pipiens*.

Methods

Isofemale lines construction

C. pipiens larvae and pupae were collected in the field and reared to adulthood in the laboratory. Females were then fed on blood to lay eggs that served to establish isofemale lines. Each egg raft (containing 100–300 eggs) was individually isolated for hatching, and the *Wolbachia* group present was determined by performing *pk1* PCR-RFLP tests²⁶ on two first-instar larvae (L1). Isofemale lines were created by rearing the offspring resulting from a single egg-raft (thus from a single female). We established 162 isofemale lines for this study. We also used 21 isofemale lines from laboratory stocks of various geographic origins (Supplementary Data 1 and Supplementary Table 1). Isofemale lines were reared in 65 dm³ screened cages kept in a single room at 22 to 25°C, under a 12 h light/12 h dark cycle. Larvae were fed with a mixture of shrimp powder and rabbit pellets, and adults were fed on honey solution.

Crossing type determination

Crossing types were characterized by crossing males (25–50 virgin males) from each of the studied isofemale lines with females (25–50 virgin females) from four reference laboratory isofemale lines hosting different *Wolbachia* strains [Tunis (*wPipI*), Lavar (*wPipII*), Maclo (*wPipIII*) and Istanbul

(wPipIV)] or with females from Tunis only (Supplementary Data 1 and Supplementary Table 1).

Illumina sequencing

The genomes of three *Wolbachia* strains from three *C. pipiens* isofemale lines were fully sequenced: Tunis from group wPipI, and Harash and Istanbul from group wPipIV. *Wolbachia* is an intracellular uncultivable bacterium. To obtain satisfactory amounts of *Wolbachia* DNA, we therefore developed a protocol based on that described by Ellegaard *et al.*²⁷. This protocol involved *Wolbachia* enrichment on egg rafts immediately after oviposition, to maximize the ratio of *Wolbachia* to *Culex* genomes. The freshly laid egg rafts were washed in bleach to crack their chorion. They were then rinsed in water and homogenized in phosphate-buffered saline (1 x PBS). *Wolbachia* cells were isolated from each egg raft separately and concentrated. Each egg raft was crushed with a sterile pestle and the resulting suspension was centrifuged at 400 x g for 5 min at 4°C to remove the cell debris. The supernatant was then centrifuged again at 6,000 x g for 5 min at 4°C, to obtain a pellet of *Wolbachia* cells. These cells were resuspended in phosphate-buffered saline (1 x PBS) and passed through a filter with 5 µm pores (GVS Filter Technology), and a filter with 2.7 µm pores (Whatman), to remove any remaining particles bigger than *Wolbachia*. The resulting filtrate was centrifuged at 6,900 x g for 15 min at 4°C to obtain a pellet of *Wolbachia* cells. Multiple-displacement amplification (MDA) was carried out directly on the *Wolbachia* pellet, with the Repli-g- mini kit (Qiagen). For each strain, five amplifications were performed independently, to randomize potential amplification errors, and the amplicons were then pooled for sequencing. Illumina sequencing was performed by the 3 kb Long Jumping Distance Mate-Pair method, on a HiSeq2000 instrument (Eurofins MWG Operon platform in Ebersberg, Germany), generating 2 x 100 bp sequences for each fragment.

Cloning and Sanger sequencing of the *cidA* and *cidB* genes

For all experiments, total DNA was extracted with the Dneasy Blood & Tissue Spin-Column kit (Qiagen; bench protocol: animal tissues). For amplification of the full-length *cidA* gene, we used two pairs of overlapping primers, 1/2 and 3/4 (Fig. 7, Supplementary Table 2). For amplification of the full-length *cidB* gene, we used four pairs of overlapping primers: 5/6, 7/8, 9/10 and 11/12 (Fig. 8,

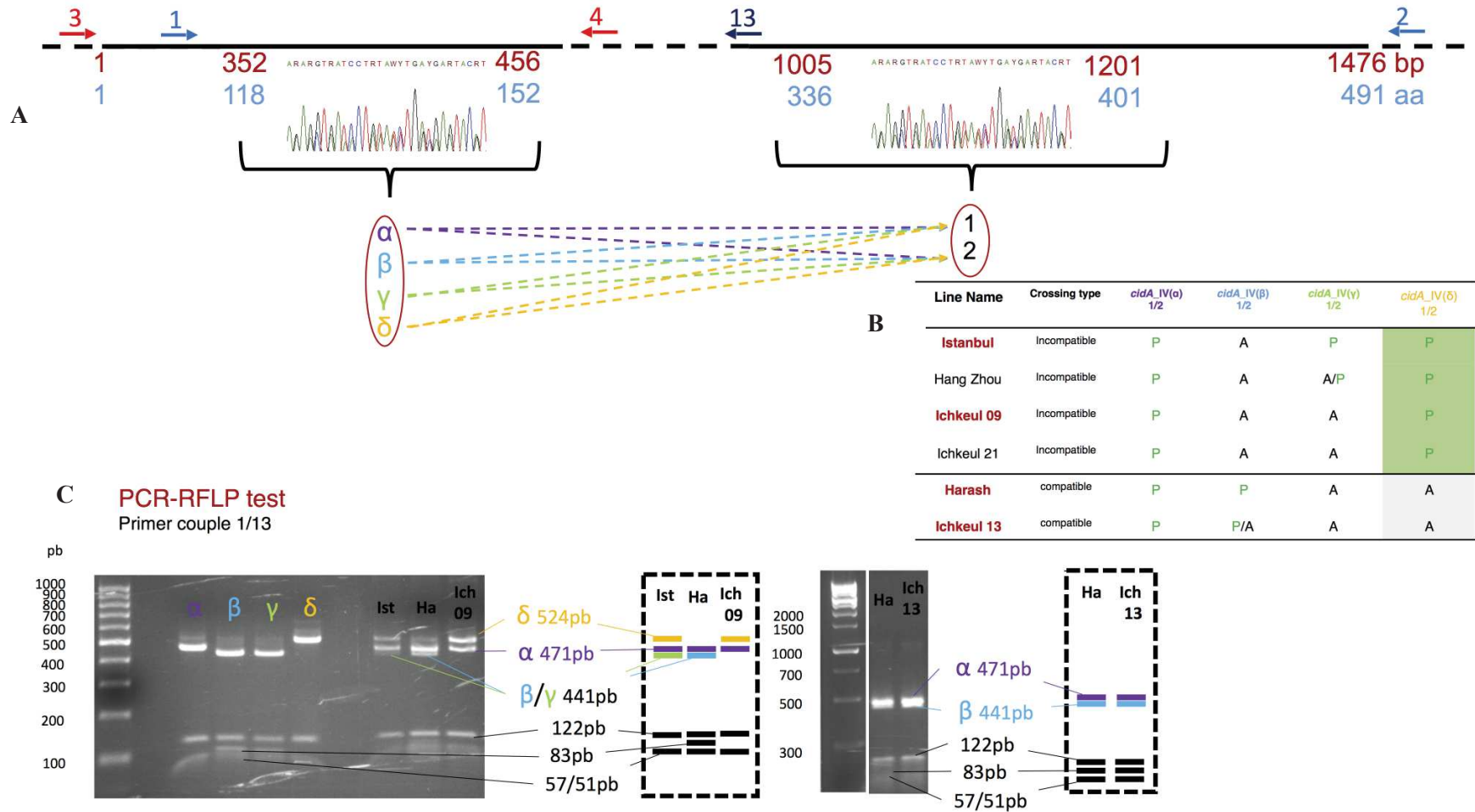


Fig. 7 *cidA_IV* gene polymorphic regions and PCR-RFLP tests for specific variants.

A Schematic representation of the architecture of *cidA_IV* polymorphism. The black line represents the *cidA_IV* sequence with a focus on the two polymorphic regions. Non-polymorphic regions were shortened and represented as a dashed line. Numbers in red under the line represent nucleotide positions, and the numbers in blue indicate amino-acid positions. The overlapping oligonucleotides used for PCR amplification are represented by arrows numbered as in Supplementary Table 2. Both polymorphic regions were studied, but only the upstream matched the compatibility profile; its four different sequences (α , β , γ , and δ) were followed by one of the two possible sequences (1 or 2) in the downstream polymorphic region. A different color code was used for α (purple), β (light blue), γ (light green), and δ (yellow) sequences in the upstream part of the *cidA* gene.

B The repertoire of *cidA_IV* variants is different in “compatible” and “incompatible” lines. *cidA_IV*(δ) is present only in lines with “incompatible” crossing type. The names of the *C. pipiens* lines used to set-up the PCR-RFLP (c) test are highlighted in red.

C PCR-RFLP tests for distinguishing between *cidA_IV* variants on the basis of the upstream polymorphic region. A 778 bp fragment was amplified with primers 1/13. Double digestion with ApoI and Hpy188I distinguished between *cidA_IV*(α) (six fragments: 471; 122; 57; 53; 51; and 24 bp), *cidA_IV*(β) and *cidA_IV*(γ) (six fragments: 441; 122; 83; 57; 51; and 24 bp), and *cidA_IV*(δ) (five fragments: 524; 122; 57; 51; and 24 bp). Electrophoresis gel of the left: PCR-RFLP on DNA from clones and the Istanbul, Harash and Ichkeul 09 lines. On the right of the gel, a schematic representation of the PCR-RFLP profiles of these lines. Electrophoresis gel of the right: PCR-RFLP on DNA from the Harash and Ichkeul 13 lines. On the right of the gel, a schematic representation of the PCR-RFLP profiles of these lines. Digestion bands that are specific of the variants are represented with the color code established in a. Bands that are not used to discriminate between the variants are represented in black on the bottom of the schematic gel.

Supplementary Table 2). We used *cidA* fragments amplified with primers 1/2 and *cidB* fragments amplified with primers 7/8 for cloning experiments, because these two regions contain all the polymorphism observed in *cidA* and *cidB* variants (Fig. 3 and Fig. 4). PCR products were extracted from the electrophoresis gel with the W-tracta disposable gel extraction tool (Star Lab) and purified with the QIAquick Gel Extraction Kit (Qiagen). The TOPO TA cloning Kit pCR 2.1-TOPO Vector (Invitrogen) was used for ligation. Before electroporation, the DNA was dialyzed with VSWP membrane filters with 0.025 μm pores (Merck Millipore) to remove excess salt from the ligation. Electroporation was performed with One Shot TOP10 Electrocomp *E.coli* (Invitrogen) and a Bio-Rad Micropulser with 2 mm-path electroporation cuvettes (Eurogentec) according to the manufacturer's instructions. For repertoire acquisition of *cidA* and *cidB*, at least 48 clones for each individual were amplified and sequenced with M13 primers. Once the PCR-RFLP tests (see below) had been developed, more sequences were obtained for the variants initially underrepresented. For all the *C. pipiens* isofemale lines analyzed by cloning-sequencing, direct Sanger sequencing results for *cidA* and *cidB* were compared with the repertoire obtained by cloning-sequencing, to ensure that no variants were missing. For Sanger sequencing, the PCR products were purified with the Agencourt Ampure PCR purification kit (Agencourt) and directly sequenced with an ABI Prism 3130 sequencer using the BigDye Terminator Kit (Applied Biosystems).

Sanger sequencing data analyses

Sequence variants of *cidA* and *cidB* were aligned, with Muscle implemented in Seaview 6.4.1 software²⁸. Alignments were then analyzed within a phylogenetic network framework, to account for potentially conflicting signals due to recombination. A phylogenetic network was constructed from uncorrected P distances by the neighbor-net method²⁹ implemented in SplitsTree4³⁰.

PCR-restricted fragment length polymorphism (RFLP) tests

We designed PCR-RFLP tests to detect the presence of the *cidA*_{IV}(δ) [*cidA*_{IV}(δ /1) *cidA*_{IV}(δ /2)] and *cidB*_{IV}(2) [*cidB*_{IV}(a/2) and *cidB*_{IV}(b/2)] variants in isofemale lines. For *cidA*_{IV} variants, only the first polymorphic region was correlated with crossing types. We therefore designed the PCR-RFLP test

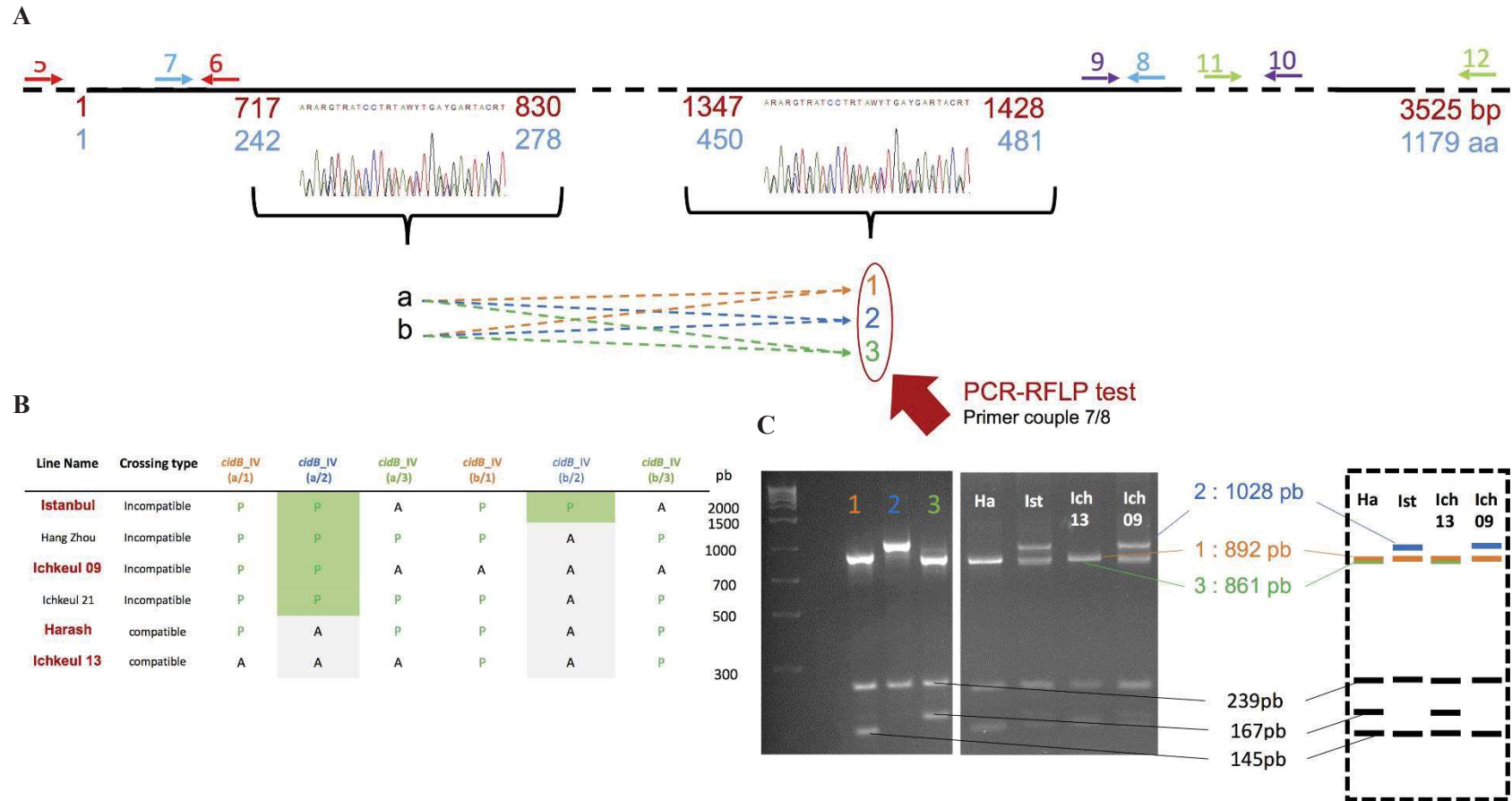


Figure 8 | *cidB_IV* gene polymorphic regions and RFLP tests for specific variants

A. Schematic representation of the polymorphism architecture of *cidB_IV*. The black line represents the *cidB_IV* sequence with a focus on the polymorphic region. Non polymorphic regions were shortened and represented as a dashed line. Numbers in red represent nucleotide positions and those in blue indicate amino-acid positions. The overlapping oligonucleotides used for amplification are represented by arrows numbered as in Supplementary Table 2. Both polymorphic regions are indicated but only the downstream (on the left, base pairs 1347 to 1428, amino acids 450 to 481) matches the compatibility profiles. The two sequences of the upstream polymorphic zone (a and b) are followed by one of the three possible sequences (1, 2 and 3) in the downstream polymorphic region. A different color code was used for 1 (orange), 2 (blue) and 3 (green) different sequences in the downstream part of the *cidB* gene.

B. The *cidB_IV* variant repertoire differs between “compatible” and “incompatible” lines. *cidB_IV*(a/2) is present only in lines with “incompatible” crossing type. The names of the *C. pipiens* lines used to set up the PCR-RFLP (panel C) test are highlighted in red.

C. PCR-RFLP tests for distinguishing between *cidB_IV* variants on the basis of the second polymorphic region. A 1267-1276 bp fragment was amplified with primers 7/8. Double digestion with *BanI* and *TaqI* distinguished between *cidB_IV*(1) (3 fragments: 892, 239, 145 bp), *cidB_IV*(2) (2 fragments: 1028, 239 bp) and *cidB*(3) (3 fragments: 861, 239, 167bp). Panel to the left of the electrophoresis gel: PCR-RFLP on DNA from clones; right panel: PCR-RFLP on DNA from lines. The panel on the right shows schematic representations of the PCR-RFLP profiles for these lines. Digestion bands which are specific of the variants are represented with the color code established in the panel A. Bands which are not used in to discriminate between the variants are represented in black on the bottom of the schematic gel.

to distinguish between the *cidA_IV*(α), *cidA_IV*(β), *cidA_IV*(γ) and *cidA_IV*(δ) sequences (Fig. 7). A 778 bp fragment was amplified with the 1/13 primer pair (Supplementary Table 2). Double digestion of the PCR products with *ApoI* and *Hpy188I* (New England Biolabs) identified three groups of sequencing: *cidA_IV*(α) (6 fragments: 471, 122, 57, 53, 51, 24 bp), *cidA_IV*(β) and *cidA_IV*(γ) (6 fragments: 441, 123, 83, 57, 51, 24 bp) and *cidA_IV*(δ) (5 fragments: 524, 122, 57, 51, 24 bp).

For *cidB_IV* variants, only the second polymorphic region was found to be correlated with crossing types, so the PCR-RFLP test was designed to discriminate between the *cidB_IV*(1), *cidB_IV*(2) and *cidB_IV*(3) sequences (Fig. 8). A 1267-1276 bp fragment was amplified with the 7/8 primer pair (Supplementary Table 2). Double digestion of the PCR products with *BanI* and *TaqAI* (New England Biolabs) distinguished three different groups of sequences: *cidB_IV*(1) (3 fragments: 892, 239, 145 bp), *cidB_IV*(2) (2 fragments: 1028, 239 bp) and *cidB_IV*(3) (3 fragments: 861, 239, 167 bp).

Gene copy number quantification

The copy numbers of the *cidA* and *cidB* genes were estimated per *Wolbachia* genome, for four laboratory lines, on male and female adults and a pool of larvae at the earliest stage, by real-time quantitative PCR with a Roche Light Cycler. Three PCRs were performed, in triplicate, on each sample, one for the *Wolbachia wsp* locus present as single copy per genome (primer pair 18/19), one for *cidA* (primer pair 14/15) and one for *cidB* (primer pair 16/17). These primer pairs specifically bind to a region conserved in all *wPip* groups for each gene (Supplementary Table 2). The ratio between *cidA* or *cidB* and *wsp* signals was used to estimate the relative number of copies of *cidA* or *cidB* per *Wolbachia* genome.

***cidA-cidB* expressed repertoire**

Fresh mosquito samples from adult females and males and first (L1) and last (L4) instar larvae were used for RNA extraction with Trizol (Life Technologies). The RNA obtained was treated with DNase with the TURBO DNA-free™ Kit (Life Technologies), in accordance with the manufacturer's instructions. We reverse-transcribed 2-5 μ g of each total RNA sample with the SuperScript III Reverse Transcriptase Kit and 30 ng of random oligomer primers ((RP)10; Invitrogen, Life Technologies).

The polymorphic zones of the *cidA* and *cidB* cDNAs were amplified, cloned and 48 clones were sequenced, as described above.

Statistical analyses

All statistical analyses were performed with R version 3.0.2 software. We compared the proportions of “compatible” and “incompatible” isofemales harboring *cidA*_{IV}(δ) and *cidB*_{IV}(a/2) in a two-sample test for equality of proportions with continuity correction based on Chi-squared.

Mapping of illumina reads on the reference genome

Illumina Long Jumping Distance sequencing data were trimmed with Trimmomatic version 0.36.

Paired reads and single reads for each line were compiled into Fastq files and mapped onto the reference genome *wPip_Pel*, with *bwa mem* and default parameters. Mappings were visualized with IGV version 2.3.68.

For each *Wolbachia* strain, the coverages of WP0282 (*cidA*), WP0283 (*cidB*), WP0294 (*cinA*), WP0295 (*cinB*), and MLST genes^{19,31} were calculated with *genomecov* from *bedtools*, with the *-d* option, to calculate “per-base” genome coverage. Several adjustments were made: (i) coverage was normalized by dividing by the mean coverage on the whole reference genome of the strain concerned, to eliminate bias due to the higher coverage for strains for which there were larger numbers of reads in the Illumina dataset; (ii) Polymorphic regions of *cidA* and *cidB* were removed, due to the considerable variability of coverage in these regions; (iii) The region of *cidB* from position 2461 to the end of the gene was removed because this region is identically duplicated in the *wPip_Pel* genome⁹, artificially increasing its coverage.

Data availability

The Illumina raw reads data of the three *wPip* genomes have been deposited in the Sequence Read Archive with Bioproject PRJNA393398 and BioSample SAMN07315580 (*wPipI* Tunis),

SAMN07327402 (wPipIV Harash) and SAMN07327406 (wPipIV Istanbul). Nucleotidic and proteic sequences of *cidA-cidB* variants have been deposited in GenBank and accession numbers are available in the Supplementary Table 3. The authors declare that all other data supporting the findings of this study are available within the article and its Supplementary Information files, or are available from the authors upon request.

Supplementary information

See annexes at the end of the PhD manuscript

Acknowledgments

We would like to thank Sylvain Charlat, Pierrick Labbé, Benjamin Lopin, Nicole Pasteur, and François Rousset for helpful comments on the manuscript. We also thank Khalid Belkhir for bioinformatic analyses performed on the Montpellier Bioinformatics Biodiversity cluster computing platform. Frédéric Delsuc for phylogenetic analyses, and Arnaud Berthomieu, Ali Bouattour, Emilie Dumas, Patrick Makoundou and Sandra Unal for technical support. This work was funded by the French ANR (project "CIAWOL" ANR-16-CE02-0006-01 and ANR-10-BINF-0003).

Author contributions

MB, MS, and MW conceptualized and designed the study; MB, CA, MaBe and FJ performed the experiments; MB, MC-G, MS and MW analyzed and interpreted the data; MB, MS and MW wrote the manuscript.

Competing interests

The authors declare no competing financial interests.

References

1. Yen, J. H. & Barr, A. R. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature* **232**, 657–658 (1971).
2. Poinsoot, D., Charlat, S. & Merçot, H. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: Confronting the models with the facts. *BioEssays* **25**, 259–265 (2003).
3. Hurst, L. D. The Evolution of Cytoplasmic Incompatibility or When Spite can be Successful. *J. theor. Biol.* **148**, 269–277 (1991).
4. Werren, J. H. Biology of *Wolbachia*. *Annu. Rev. Entomol.* **42**, 587–609 (1997).
5. Foster, J. *et al.* The *Wolbachia* genome of *Brugia malayi*: Endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* **3**, 0599–0614 (2005).
6. Siozios, S. *et al.* The Diversity and Evolution of *Wolbachia* Ankyrin Repeat Domain Genes. *PLoS One* **8**, e55390 (2013).
7. Sedgwick, S. G. & Smerdon, S. J. The ankyrin repeat: a diversity of interaction on a common structural framework. *Trends Biochem. Sci.* 311–316 (1999).
8. Duron, O. *et al.* Variability and expression of ankyrin domain genes in *Wolbachia* variants infecting the mosquito *Culex pipiens*. *J. Bacteriol.* **189**, 4442–4448 (2007).
9. Beckmann, J. F. & Fallon, A. M. Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: Implications for cytoplasmic incompatibility. *Insect Biochem. Mol. Biol.* **43**, 867–878 (2013).
10. Beckmann, J. F., Ronau, J. A., Hochstrasser, M. *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2**, 17007 (2017).
11. Lepage, D. P. *et al.* Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **543**, 243–247 (2017).
12. Sutton, E. R., Harris, S. R., Parkhill, J. & Sinkins, S. P. Comparative genome analysis of

- Wolbachia* strain wAu. *BMC Genomics* **15**, 928 (2014).
13. Atyame, C. M. *et al.* *Wolbachia* divergence and the evolution of cytoplasmic incompatibility in *Culex pipiens*. *PLoS One* **9**, 21–26 (2014).
 14. Laven, H., Wright, J. W. & Pal, R. in *Genetics of Insect Vectors of Disease* 251–275 (1967).
 15. Atyame, C. M. *et al.* Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito populations. *Mol. Ecol.* **20**, 286–298 (2011).
 16. Atyame, C. M., Delsuc, F., Pasteur, N., Weill, M. & Duron, O. Diversification of *Wolbachia* endosymbiont in the *Culex pipiens* mosquito. *Mol. Biol. Evol.* **28**, 2761–2772 (2011).
 17. Nor, I. *et al.* On the genetic architecture of cytoplasmic incompatibility: inference from phenotypic data. *Am. Nat.* **182**, E15–24 (2013).
 18. Klasson, L. *et al.* Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol. Biol. Evol.* **25**, 1877–1887 (2008).
 19. Baldo, L. *et al.* Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol.* **72**, 7098–7110 (2006).
 20. Lo, N. *et al.* Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. *Int. J. Syst. Evol. Microbiol.* **57**, 654–657 (2007).
 21. Pons, J. L. & Labesse, G. @TOME-2: A new pipeline for comparative modeling of protein-ligand complexes. *Nucleic Acids Res.* **37**, 485–491 (2009).
 22. Andrade, M. A., Perez-Iratxeta, C. & Ponting, C. P. Protein Repeats: Structures, Functions, and Evolution. *J. Struct. Biol.* **134**, 117–131 (2001).
 23. Bordenstein, S. R., Marshall, M. L., Fry, A. J., Kim, U. & Wernegreen, J. J. The Tripartite Associations between Bacteriophage , *Wolbachia* , and Arthropods. *PLoS Pathog.* **2**, e43 (2006).
 24. Wright, J. D., Sjostrand, F. S., Portaro, J. K. & Barr, A. R. The ultrastructure of the rickettsia-

- like microorganism *Wolbachia pipientis* and Associated virus-like bodies in the mosquito *Culex pipiens*. *J. Ultrastructure Res.* **63**, 79–85 (1978).
25. Aakre, C. D. *et al.* Evolving New Protein-Protein Interaction Specificity through Promiscuous Intermediates. *Cell* **163**, 594–606 (2015).
 26. Dumas, E. *et al.* Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol. Biol.* **13**, 181 (2013).
 27. Ellegaard, K. M., Klasson, L., Näslund, K., Bourtzis, K. & Andersson, S. G. E. Comparative Genomics of *Wolbachia* and the Bacterial Species Concept. *PLoS Genet.* **9**, e1003381 (2013).
 28. Gouy, M., Guindon, S. & Gascuel, O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol. Biol. Evol.* **27**, 221–224 (2010).
 29. Bryant, D. & Moulton, V. Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Mol. Biol. Evol.* **21**, 255–265 (2004).
 30. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).
 31. Paraskevopoulos, C., Bordenstein, S. R., Wernegreen, J. J., Werren, J. H. & Bourtzis, K. Toward a *Wolbachia* multilocus sequence typing system: Discrimination of *Wolbachia* strains present in *Drosophila* species. *Curr. Microbiol.* **53**, 388–395 (2006).

Chapitre II :

La fonction *mod* de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens* est gouvernée par le gène *cidB* et non par le gène *cidA*

Article n°2 soumis à Proceedings of the Royal Society of London B: Biological Science

**Cytoplasmic incompatibility *mod* diversity is governed by the *Wolbachia*
cidB gene but not *cidA* in *Culex pipiens* mosquitoes**

Manon Bonneau¹, Beniamino Caputo², Aude Ligier¹, Rudy Caparros³, Sandra Unal¹, Daniele Arnoldi⁴,
Mathieu Sicard¹ and Mylène Weill^{1*}

¹ ISEM, University of Montpellier, CNRS, IRD, Montpellier, France

² Dipartimento di Sanità Pubblica e Malattie Infettive, Laboratory affiliated to Istituto Pasteur Italia - Fondazione Cenci Bolognetti, Sapienza University of Rome, Rome, Italy

³ Functional and Evolutionary Entomology, Gembloux Agro-Bio Tech, University of Liège, Passage des Déportés, 2, 5030 Gembloux, Belgium

⁴ Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, 38010 San Michele, Italy

*corresponding author

Email: mylene.weill@umontpellier.fr

Abstract

In arthropods, endosymbiotic *Wolbachia* bacteria frequently induce cytoplasmic incompatibility (CI), resulting in the death of embryos generated by crosses between infected males and uninfected females or between individuals infected with incompatible *Wolbachia* strains. CI can be seen as a modification-rescue system in which paternal *Wolbachia* produce a *mod* factor, inducing embryonic defects unless the maternal *Wolbachia* produces a compatible *resc* factor. The *Wolbachia* *cidA* and *cidB* genes are associated with CI induction and diversity in *Culex pipiens*. However, although *cidB* is clearly required for *mod* function, the role of *cidA* remains unclear. According to the “two-by-one” model based on *Drosophila melanogaster* transgenesis experiments, *cidA* is involved in both *mod* function, when expressed in males, and in *resc* function, when expressed in females. In accordance with this model, correlative studies in natural *C. pipiens* populations have shown specific *cidA* variants to be associated with variations of *mod* phenotype, suggesting a role for *cidA* in *mod* function. However, this correlation may also result from evolutionary codiversification with *cidB*, as predicted in toxin-antidote systems. The presence of a *cidA* variant in all mosquitoes infected with compatible wPip highlights the role of *cidA* as a *resc* factor. By studying new natural *C. pipiens* populations, we confirm here the role of *cidB*, but rule out the involvement of *cidA* in *mod* phenotype variations. These results are consistent with a simple “one-by-one” model of CI in *C. pipiens* mosquitoes, with *cidA* encoding the *resc* function and *cidB* encoding the *mod* function.

Keywords : *Wolbachia*, Cytoplasmic Incompatibility, *Culex pipiens*, *cidA*, *cidB*

Introduction

Wolbachia are maternally inherited endosymbiotic bacteria commonly found in arthropods and filarial nematodes [1–3]. More than 50% of terrestrial arthropod species are thought to be infected [4–6]. The pervasiveness of this genus is mostly attributed to its ability to manipulate host reproduction, facilitating its spread within arthropod populations [1]. The most commonly described *Wolbachia*-induced phenotype in arthropods is cytoplasmic incompatibility (CI) [1,7]. CI is a form of conditional sterility resulting in embryonic lethality in diploid species [8] when infected males mate with uninfected females or with females infected with a different, incompatible *Wolbachia* strain [9–14]. When both males and females are infected, CI may be unidirectional (crossing is compatible in one direction but incompatible in the other) or bidirectional (crosses in both directions are incompatible) [9,15]. CI can be seen as a modification/rescue model in which the *Wolbachia* present in the male produces a *mod* factor during spermatogenesis, which induces CI through embryonic defects after fertilization unless the *Wolbachia* present in the eggs produce a compatible *resc* factor [16,17]. Recent work has implicated the syntenic *Wolbachia* genes *cidA* and *cidB*, located within the WO phage region, in CI induction [18,19]. Indeed, the expression of these two genes in transgenic *Drosophila melanogaster* reproduces CI phenotypes [18,19]. An analysis of the protein domain homology of the *cidA* and *cidB* genes revealed that *cidB* encodes a deubiquitylating enzyme (DUB) [18–20]. In *Drosophila melanogaster*, the expression of a *cidB* construct with an inactivated catalytic DUB domain in males abolished CI, confirming the role of the CidB protein as the *mod* factor [18]. However, the role of *cidA* in the CI mechanism is less clear. Both *cidA*^{wMel} and *cidB*^{wMel} are required to induce CI in transgenic *D. melanogaster*, and the expression of *cidA*^{wMel} in transgenic males infected with wMel amplifies the CI phenotype in crosses with uninfected *D. melanogaster* females, suggesting that *cidA*^{wMel} is also involved in *mod* function [19]. By contrast, the requirement for the co-expression of *cidA* with *cidB* for the production of live transgenic *D. melanogaster* and *Saccharomyces cerevisiae* suggests that CidA may serve as the antidote to CidB [18]. Furthermore, the expression of *cidA*^{wMel} in transgenic *D. melanogaster* females restores fertility when such females are crossed with infected males, strongly suggesting a role for *cidA*^{wMel} in the *resc* function [21]. Based on these findings, Shropshire et al. (2018)[21] proposed a

“two-by-one” model, in which that *cidA*^{wMel} acts as a *mod* factor when expressed in males and as a *resc* factor when expressed in females, in *D. melanogaster*.

In *Culex pipiens* mosquitoes, the genomes of all studied wPip *Wolbachia* strains have several polymorphic copies of these two genes [22]. Moreover, wPip strains inducing different CI phenotypes harbor different *cidA*^{wPip} and *cidB*^{wPip} variants in their genomes. These results provide strong evidence for a role of these variants in the tremendous diversity of CI phenotypes described in this species [15,22]. The putative roles of the *cidA*^{wPip} and *cidB*^{wPip} genes in CI diversity have been investigated by analyzing the diversity of the *cidA*^{wPip} and *cidB*^{wPip} repertoires in 180 isofemale lines from natural populations from China, North Africa and Turkey, all infected with *Wolbachia* from phylogenetic wPip group IV [22,23]. Mosquitoes infected with wPipIV present two different *mod* phenotypes: the “incompatible” *mod* phenotype, in which males cannot produce viable progenies with females infected with wPip from group I, II or III, and the “compatible” *mod* phenotype, in which males produce viable progenies with such females [22,23]. The association of a specific *cidB*^{wPip} variant [*cidB*_IV(a/2)] with the “incompatible” *mod* phenotype in these 180 isofemale lines, together with the results of functional studies in *D. melanogaster* [18,19], demonstrated the involvement of *cidB*^{wPip} in *mod* function and of its responsibility for the variation of the CI *mod* phenotype in *C. pipiens*. However, as in *Drosophila*, the role of *cidA*^{wPip} remains unclear in *C. pipiens*. An association was found between a specific *cidA*^{wPip} variant [*cidA*_IV(δ)] and the “incompatible” *mod* phenotype, suggesting a possible role for *cidA*^{wPip} in *mod* phenotype diversity [22], but the presence of a ubiquitous *cidA*^{wPip} variant [*cidA*_IV(α)] in all wPipIV *Wolbachia* strains, regardless of their geographic origins and potentially accounting for full compatibility, suggests a role for *cidA*^{wPip} in the *resc* function [15,22]. As the CidA and CidB proteins interact with each other in a specific manner [18], the diversity of CI phenotypes described in *C. pipiens* probably results from differential interactions between specific CidA and CidB variants. Indeed, we have already suggested that the zone of interaction between CidA and CidB proteins is probably located in the polymorphic regions shown to be correlated with CI phenotype variations [22]. It is possible that the association of the *cidA*_IV(δ) variant with the wPipIV “incompatible” *mod* phenotype also results from codiversification with the *cidB*_IV(a/2) variant for binding adjustment, as predicted for toxin-antidote

systems. A distribution of *cidA* and *cidB* variants may, therefore, be independent of any role of *cidA* in *mod* function diversity [22].

We investigated the role of *cidA* in the CI mechanism, by studying the *cidA*^{wPip}/*cidB*^{wPip} variant repertoires in wPipIV strains from newly sampled natural populations with putatively different evolutionary histories of codiversification between *cidA* and *cidB*. We studied *mod* variation in four North Italian populations presenting a mixture of “compatible” and “incompatible” *mod* phenotypes. The presence of the *cidB*_IV(a/2) variant in all “incompatible” isofemale lines confirmed the role of *cidB* in *mod* variation. By contrast, *cidA* IV(α), thought to be associated with intra-wPipIV compatibility, was present in all lines, regardless of their *mod* phenotype. However, no *cidA* variant was associated with either “incompatible” or “compatible” *mod* phenotypes, demonstrating the absence of *cidA*^{wPip} involvement in CI *mod* variation in *C. pipiens*.

Results

Only *cidB*_IV(a/2) and *cidB*_IV(b/2) variants are associated with the “incompatible” *mod* phenotype in North Italy

Isofemale lines were established from larvae collected at four sites in the province of Trento of North-East Italy (San Michele all’Adige, Roveré Della Luna, Mezzocorona and Zambana), because it was known that wPip strains from the wPipIV group occurred in this area [24]. Crosses between males from 67 isofemale lines and females from the wPipI Tunis laboratory line generated the two *mod* phenotypes described by Atyame et al. (2015)[23]: 62 isofemale lines contained males incompatible with Tunis females (“incompatible” isofemale lines) and five contained males compatible with Tunis females (“compatible” isofemale lines) (Supplementary DataSet 1). Thus, 92.5% of the isofemale lines in North Italy were incompatible.

For investigation of the diversity of *cidA*_IV and *cidB*_IV variant repertoires in North Italian wPipIV-infected *C. pipiens* and identification of the *cidA*_IV and *cidB*_IV variants associated with “compatible”

A

Line name	<i>mod</i> phenotype	<i>cidA_IV</i>							
		$\alpha 1$	$\alpha 2$	$\beta 1$	$\beta 2$	$\delta 1$	$\delta 2$	$\epsilon 1$	$\epsilon 2$
Luna 8	Compatible	P	P	A	A	A	A	P	P
Luna 27	Compatible	P	P	A	A	A	A	P	P
Luna 1	Incompatible	P	A	A	A	A	A	P	P
Luna 3	Incompatible	P	A	A	A	A	A	P	P
Michele 1	Incompatible	P	P	A	A	A	A	P	P
Mezzo 9	Incompatible	P	P	A	A	A	A	P	P
Michele 26	Compatible	P	P	A	A	A	A	P	P

B

Line name	<i>mod</i> phenotype	<i>cidB_IV</i>							
		a1	a2	a3	b1	b2	b3	c1	c3
Luna 8	Compatible	A	A	P	A	A	P	A	A
Luna 27	Compatible	A	A	P	A	A	P	A	A
Luna 1	Incompatible	A	P	P	A	P	A	A	A
Luna 3	Incompatible	A	P	P	A	P	P	A	A
Michele 1	Incompatible	P	P	P	P	P	P	A	A
Mezzo 9	Incompatible	A	P	P	A	P	P	A	A
Michele 26	Compatible	P	P	P	A	A	A	P	P

Figure 1. *cidA* and *cidB* variant repertoires for seven *wPipIV* strains from North Italy

A. *cidA_IV* variants previously detected in natural populations from North Africa (Bonneau et al., 2018) and new variants described in this study were either present (P) in or absent (A) from the isofemale lines originating from North Italy studied here.

B. *cidB_IV* variants previously detected in natural populations from North Africa (Bonneau et al., 2018) and new variants described in this study were either present (P) in or absent (A) from the isofemale lines from North Italy studied.

The CI phenotypes of the Italian *C. pipiens* isofemale lines infected with *wPipIV* strains were determined by crossing males with females from the *wPipI* Tunis laboratory line. A “compatible” *mod* phenotype corresponds to an isofemale line in which males are compatible with females from the Tunis line. An “incompatible” phenotype corresponds to a line in which males are incompatible with females from the Tunis line.

or “incompatible” *mod* phenotypes in these new populations, we analysed two “compatible” (Luna 8 and Luna 27) and four “incompatible” (Luna 1, Luna 3, Michele 1 and Mezzo 9) isofemale lines (Fig. 1). Cloning and Sanger sequencing revealed the absence of *cidA_IV*(δ) and the presence of *cidA_IV*(α) in the six isofemale lines analyzed (Fig. 1). A new variant called *cidA_IV*(ϵ), slightly different from the *cidA_IV*(β) previously described in North Africa due to the replacement of a valine with an isoleucine residue in position 143, was detected in the six isofemale lines studied (Fig. 1 Sup. Fig. 1). No *cidA_IV* variants were found associated with either “compatible” or “incompatible” *mod* phenotypes.

The sequencing of *cidB* gene repertoires revealed the presence of *cidB_IV*(a/2) and *cidB_IV*(b/2) in the four “incompatible” isofemale lines and their absence in the two “compatible” lines (Fig. 1). No other *cidB_IV* variants were found associated with *mod* phenotypes (Fig. 1).

For extension of the cloning and sequencing results to the population scale, we performed a PCR-RFLP screen of *cidA_IV* and *cidB_IV* variants on the 67 isofemale lines from the four sites (Supplementary Dataset 1). The *cidA_IV*(α) variant was detected in all 67 lines. By contrast, the *cidA_IV*(δ) variant was not found in any of the 67 lines, including the 62 “incompatible” isofemale lines. For *cidB_IV* variants, the PCR-RFLP test developed by Bonneau et al. (2018)[22] distinguishes between the downstream polymorphic sequences *cidB_IV*(1), *cidB_IV*(2), and *cidB_IV*(3), regardless of the upstream sequences present. The detection of *cidB_IV*(2) thus accounts for the presence of *cidB_IV*(a/2) and/or *cidB_IV*(b/2) variants (see Materials and method for further information about the screening methods used). We also developed an additional *cidB_IV*(2)-specific direct PCR test (also accounting for the *cidB_IV*(a/2) and/or *cidB_IV*(b/2) variants) to confirm the PCR-RFLP test results (Supplementary Dataset 1).

All 62 “incompatible” isofemale lines carried the *cidB_IV*(2) downstream region (dr), as confirmed by the two independent PCR-based methods (Supplementary Dataset 1). However, this variant was also detected in Michele26, one of the five “compatible” isofemale lines. Similar incongruences (*i.e.* “compatible” lines harboring the *cidB_IV*(2) variant) were observed in 4.9% of isofemale lines from natural populations from North Africa [22].

Table 1. Distribution of *cidA*_IV(α), *cidA*_IV(δ) and *cidB*_IV(2) in the 247 isofemale lines from North Africa, Turkey, China and North Italy

<i>cidA cidB</i> variants	<i>mod</i> phenotype	
	« Compatible » N=168	« Incompatible » N=79
	China = 0	China = 1
	North Africa = 163	North Africa = 15
	North Italy = 5	North Italy = 62
	Turkey = 0	Turkey = 1
<i>cidA</i> _IV(α)	China : 0 North Africa : 163 North Italy : 5 Turkey : 0 Total : 100%	China : 1 North Africa : 15 North Italy : 62 Turkey : 1 Total : 100%
<i>cidA</i> _IV(δ)	China : 0 North Africa : 16 North Italy : 0 Turkey : 0 Total : 9.5%	China : 1 North Africa : 15 North Italy : 0 Turkey : 1 Total : 21.5 %
<i>cidB</i> _IV(2)	China : 0 North Africa : 8 North Italy : 1 Turkey : 0 Total : 5.4%	China : 1 North Africa : 15 North Italy : 62 Turkey : 1 Total : 100%

Number of isofemale lines with “compatible” or “incompatible” *mod* phenotypes are indicated for each geographic area. The prevalence of each variant is given for “compatible” and “incompatible” isofemale lines in each area, together with the total percentage of “compatible” or “incompatible” isofemale lines carrying the variant concerned, regardless of geographic origin

The global distribution of the *cid* variants of interest in all the natural populations studied (Table 1) revealed the presence of *cidA_IV*(α) in all 247 *wPipIV*-infected isofemale lines, regardless of *mod* phenotype and geographic origins. *cidA_IV*(δ), which had been reported to be associated with the “incompatible” *mod* phenotype, was not detected in “incompatible” isofemale lines in North Italy. Overall, *cidA_IV*(δ) was detected in 21.5% of the “incompatible” isofemale lines. By contrast, *cidB_IV*(2) dr was detected in all the “incompatible” isofemale lines, regardless of geographic origin, and in 5.4% of the “compatible” isofemale lines.

How can Michele26 be compatible when it carries the *cidB_IV*(a/2) variant associated with incompatibility?

The Michele26 isofemale line harbors a unique cidB variant repertoire

The *cidA* and *cidB* variant repertoires of the *wPipIV* strain infecting the Michele26 isofemale line were cloned and Sanger sequenced, and compared with the repertoires of the four “incompatible” and the two “compatible” isofemale lines from North Italy. No features of the *cidA_IV* variant repertoire specific to this isofemale line were identified, as the sequences obtained were similar to those of both “compatible” and “incompatible” isofemale lines (Fig 1). The *cidB_IV* variant repertoire of Michele 26 included *cidB_IV*(a/2), but lacked *cidB_IV*(b/2), which was present in the incompatible lines. Furthermore, two variants unique to Michele26, *cidB_IV*(c/1) and *cidB_IV*(c/3), were identified (Fig. 1 and Sup. Fig. 2). The Michele26 isofemale line thus had a *cidB* variant repertoire different from the other “compatible” and “incompatible” repertoires reported for the same area.

Table 2. Total *cidA* and *cidB* variant expression levels, specific *cidB_IV(2)* expression levels and *Wolbachia* density in the testes of “compatible” Michele26 and “incompatible” Mezzo9 males

Line	Ratio	<i>Wolbachia</i> testes density	<i>cidA</i> expression	<i>cidB</i> expression	<i>cidB_IV(2)</i> expression
Mezzo9		18.10 ± 4.72 (a)	1.67 ± 0.56 (a)	0.72 ± 0.23 (a)	0.06 ± 0.01 (a)
Michele26		10.25 ± 3.42 (b)	1.33 ± 0.31 (a)	0.58 ± 0.18 (a)	0.04 ± 0.02 (b)

For each variable, letters indicate groups of statistical similarity. Mean values are expressed ± standard deviation. *Wolbachia* density in the testes was significantly lower in Michele26 than in Mezzo9 males. The levels of *cidA* and *cidB* expression did not differ significantly between males. *cidB_IV(2)* was significantly less strongly expressed in males from the Michele26 line than in males from the Mezzo9 line.

Michele26 males are as compatible as other males from “compatible” isofemale lines, even if they carry $cidB_{IV(a/2)}$

We compared the “compatible” *mod* phenotype of Michele26 with that of other “compatible” isofemale lines lacking the $cidB_{IV(a/2)}$ “incompatible” variant, we measured the hatching rate for crosses involving males from the Michele26 and Luna8 isofemale lines crossed with Tunis females. The mean hatching rates (30 egg-rafts per cross) were 0.82 ± 0.10 for Luna8 and 0.79 ± 0.18 for Michele26, and were not significantly different between these two lines ($W=463$, p -value=0.853, Supplementary Dataset 2).

$cidB_{IV(2)}$ is less strongly expressed in “compatible” Michele26 males than in “incompatible” Mezzo9 males

We then investigated possible differences in the expression of the *cidA* and *cidB* genes between “compatible” Michele26 males and “incompatible” Mezzo9 males. The overall levels of *cidA* and *cidB* expression were not significantly different (*cidA* Wilcoxon, $W=62$, p -value=0.182 and *cidB* Wilcoxon, $W=60$, p -value=0.243, Table 2). The expression of $cidB_{IV(2)}$, which was specific to variants associated with incompatibility, was studied by real-time quantitative PCR of a fragment of the $cidB_{IV(2)}$ dr sequence accounting for both the $cidB_{IV(a/2)}$ and $cidB_{IV(b/2)}$ variants (a/2 present in Mezzo9 and Michele26, and b/2 present only in Mezzo9, see Fig. 1). Expression of the $cidB_{IV(2)}$ dr fragment in males from the Michele26 isofemale line was twice as low as in males from the Mezzo9 isofemale line (Wilcoxon, $W=12$, p -value=0.003, Table 2, Supplementary Data Set 3).

Less Wolbachia in the testes of males from the “compatible” Michele26 isofemale line

As *mod* factors (CidB proteins) are probably restricted to the sperm, we determined *Wolbachia* density in the gonads of Michele26 and Mezzo9 males. The testes of males from the Michele26 isofemale line contained significantly less *Wolbachia* than those of males from the Mezzo9 isofemale line: $10.25 \pm$

3.42 and 18.10 ± 4.72 *Wolbachia* per host cell, respectively (Wilcoxon, $W=1$, p -value=0.008, Table 2, Supplementary DataSet 4).

Discussion

We investigated the role of the *cidA* gene in CI function, by sampling natural populations of *C. pipiens* infected with *wPipIV Wolbachia* in a new geographic region and monitoring the distribution of *cidA* and *cidB* variants relative to *mod* phenotype variations. We confirmed that mosquitoes in North Italy were infected with *wPipIV* strains, as previously reported by Dumas et al. (2013)[24]. Crosses between males from 67 isofemale lines and females from the Tunis *wPipI* laboratory line revealed the co-existence of the two *mod* phenotypes in these Italian populations: some males were compatible whereas others were not. However, unlike North African populations infected with *wPipIV* strains, in which 8.4% of the isofemale lines were incompatible [22,23], 92.5% of the isofemale lines in North Italian populations were incompatible.

In natural populations from North Africa, *cidA_IV*(δ) and *cidB_IV*(a/2) variants were associated with the “incompatible” *mod* phenotype (Table 1)[22]. In North Italy, only *cidB_IV*(2) dr, which accounts for both the *cidB_IV*(a/2) and *cidB_IV*(b/2) variants, was found (Table 1). Our findings demonstrate that only *cidB_IV*(2) dr was systematically found in “incompatible” repertoires, revealing the association of this downstream polymorphic region of *cidB_IV* variants with CI phenotype variations in *C. pipiens*. By contrast to the natural populations from North Africa, *cidA_IV*(δ) was not detected in any of the *wPipIV* strains infecting the 67 isofemale lines in North Italy, demonstrating that this variant was not essential for “incompatibility” (Fig 1, Table 1). We were unable to identify any *cidA* variant associated with either “incompatible” or “compatible” *mod* phenotypes, suggesting that *cidA* plays no role in *mod* phenotype diversity in *C. pipiens* and is probably not involved in *mod* function (Fig. 1). We can thus hypothesize that the association between *cidA_IV*(δ)/*cidB_IV*(a/2) and the “incompatible” *mod* phenotype in natural populations from North Africa resulted from the codiversification of these two variants. As previously suggested, *cidA* and *cidB* may encode a toxin-antidote (TA) system in which

CidA acts as the antidote to CidB [18]. Such a TA system may have driven the association of *cidA_IV*(δ) and *cidB_IV*(a/2) variants in North Africa if these variants interact particularly well together independently of any involvement of *cidA* in the *mod* phenotype. However, in the absence of *cidA_IV*(δ), another *cidA_IV* variant, such as *cidA_IV*(α), may interact with *cidB_IV*(a/2). Furthermore, as in the North African populations, the *cidA_IV*(α) variant was detected in all *wPipIV* strains infecting the 67 Italian isofemale lines. This observation strongly supports a role for *cidA* in the *resc* function in *C. pipiens*, as the presence of a ubiquitous *cidA_IV*(α) variant in each *wPipIV* tested could account for the total compatibility of mosquitoes infected with *wPipIV* strains [15]. This conclusion is further supported by the recent findings of Shropshire et al. (2018)[21], revealing the involvement of *cidA*^{wMel} in the *resc* function in transgenic *D. melanogaster*.

In North Africa, ~5% of isofemale lines were “incongruent”. In Italy, we also found one “incongruent” isofemale line (Michele26) phenotypically characterized as “compatible” but carrying the *cidB_IV*(a/2) variant associated with “incompatibility”. We investigated this discordant isofemale line further, to search for possible causes of this apparent dissociation of genotype and phenotype. We confirmed that, phenotypically, even if Michele 26 males harboured the *cidB_IV*(a/2) variant, they were as “compatible” as other “compatible” males. Genetically, we found that the Michele26 isofemale line had specific genetic features distinguishing it from both “compatible” and “incompatible” lines. Indeed, the *wPipIV* strain infecting Michele26 presented a specific *cidB* variant repertoire, including two variants [*cidB_IV*(c/1) and *cidB_IV*(c/3)] not detected in any other *wPipIV* strain. The presence of these additional *cidB_IV* variants may have resulted in the deregulation of titration or binding with other *Wolbachia* or host targets, preventing the CidB_IV(a/2) protein from inducing CI. The *wPipIV* strain infecting Michele26 also lacked the *cidB_IV*(b/2) variant presents in all “incompatible” isofemale lines from Italy. It was tempting to speculate that both *cidB_IV*(a/2) and *cidB_IV*(b/2) were required to induce CI. However, this hypothesis was ruled out by the lack of detection of *cidB_IV*(b/2) in “incompatible” isofemale lines from North African populations. Michele26 males express *cidB_IV*(2), so compatibility cannot be accounted for by a lack of expression. However, the levels of *cidB_IV*(2) dr expression in “compatible” Michele26 males are only half those in “incompatible” Mezzo9 males. This is undoubtedly

because Michele26 carries only the *cidB_IV(a/2)* variant, whereas Mezzo9 carries both *cidB_IV(a/2)* and *cidB_IV(b/2)*. As *mod* factors (*i.e.* CidB proteins) are probably released into the sperm in the testes during spermatogenesis, we measured the density of *Wolbachia* in the male gonads. We found that the density of *Wolbachia* in Michele26 was significantly lower than that in Mezzo9, which might also result in insufficient amounts of CidB_IV(2) dr protein being produced for the induction of CI.

In conclusion, our findings confirm that *cidB^{wPip}* variant repertoire is associated with the diversity of *mod* phenotypes observed in *C. pipiens*. Together with functional transgenic data [18], they clearly demonstrate the involvement of *cidB* in both *mod* function and *mod* phenotype diversity in *C. pipiens*. By contrast, we show that *cidA^{wPip}* is not involved in *mod* phenotype diversity, as lines with different *mod* phenotypes had the same *cidA^{wPip}* repertoire. In *C. pipiens* a “one-by-one” model, with CidB involved in *mod* function and acting as a toxin and CidA involved in *resc* function and acting as the antidote, is more consistent with our data than the “two-by-one” system proposed for *D. melanogaster* [21]. The *cidA* and *cidB* genes may not, therefore, behave in the same way in the *Wolbachia* bacteria infecting *D. melanogaster* and *C. pipiens*. Further investigations of this putative divergence in the molecular mechanisms of CI induced by *cidA^{wMel}* and *cidA^{wPip}* in these two species are required to shed light on this difference.

Materials and Methods

Mosquito collection and the construction of isofemale lines

C. pipiens larvae and pupae were collected from four natural breeding sites in North Italy in 2017 (Roveré della Luna, San Michele all’Adige, Zambana and Mezzocorona sites, Supplementary Dataset 1) and reared to adulthood in the laboratory. Females were then fed turkey blood (bcl Wholly Wild World) with a Hemotek membrane feeding system (Discovery Workshops, United Kingdom), to enable them to lay eggs, from which isofemale lines were established. Each egg-raft (containing 100–250 eggs) was individually isolated for hatching, and the *Wolbachia* group present was determined by performing the *pk1* PCR-RFLP test [24,25] on two first-instar larvae (L1) with a cetyl trimethylammonium bromide

(CTAB) protocol (Rogers and Bendich 1988). Isfemale lines were created by rearing the offspring resulting from a single egg-raft (thus from a single female). We established 67 isfemale lines for this study (Supplementary Dataset 1). Isfemale lines were reared in 65 dm³ screened cages in a single room maintained at 22–25°C, under a 12 h light/12 h dark cycle. Larvae were fed with a mixture of shrimp powder and rabbit pellets, and adults were fed on honey solution.

Determination of CI phenotypes

CI phenotypes were characterized by crossing males (25–50 virgin males) from each of the studied isfemale lines with females (25–50 virgin females) from the Tunis laboratory isfemale line infected with a *Wolbachia* strain from the *wPipI* group [26]. After five days in cages, the females were fed a blood meal and, five days later, egg-rafts were collected and deposited into 24-well plates. The CI status of each cross was determined by assessing egg-raft hatching status. All unhatched egg-rafts were checked for fertilization by observing embryonic development with a light microscope (Axiophot2, Zeiss), as described by Duron & Weill (2006)[27]. Isfemale lines in which the males were incompatible with females from the Tunis line were described as “incompatible” isfemale lines, whereas isfemale lines in which the males were compatible with females from the Tunis line were described as “compatible” isfemale lines.

Cloning and Sanger sequencing of *cidA* and *cidB* variants

The *cidA* and *cidB* genes of seven isfemale lines (Luna1, Luna3, Luna8, Luna27, Michele26, Michele1 and Mezzo9) were cloned and Sanger sequenced, as described by Bonneau et al. (2018)[22], using the same DNA samples used to determine *Wolbachia* phylogenetic group. Muscle implemented in Seaview 6.4.1 software [28] was used to align variant sequences.

Screen of *cidA* and *cidB* variants in natural populations from North Italy

Detection of cidA_IV(δ) and cidB_IV(2) variants

We investigated the presence of these variants in the 67 isofemale lines, using the same DNA samples used to determine *Wolbachia* phylogenetic group. We used the PCR-RFLP tests described by Bonneau et al. (2018)[22]. The *cidA_IV* and *cidB_IV* variants described by Bonneau et al. (2018)[22] have two polymorphic regions: an upstream and a downstream region. For instance, *cidA_IV*($\delta/1$) and *cidA_IV*($\delta/2$) have the same upstream sequence (δ) but two different downstream sequences (1/2), whereas *cidB_IV*(a/2) and *cidB_IV*(b/2) have two different upstream sequences (a/b) but the same downstream sequence (2) (Supplementary Figs. 1 and 2). Only the upstream polymorphic region of *cidA_IV* variants was associated with the CI phenotype. The *cidA_IV* PCR-RFLP test was, therefore, designed to distinguish between the various upstream polymorphic sequences — *cidA_IV*(α), *cidA_IV*(β) and *cidA_IV*(δ) — regardless of the downstream sequences present. The detection, with this test, of *cidA_IV*(δ) in an isofemale line accounts for the presence of *cidA_IV*($\delta/1$) and/or *cidA_IV*($\delta/2$). Only the downstream polymorphic region of *cidB_IV* variants was associated with the CI phenotype. The PCR-RFLP test was, therefore, designed to distinguish between the *cidB_IV*(1), *cidB_IV*(2), and *cidB_IV*(3) sequences, regardless of the upstream sequences present. Thus, the detection, in isofemale lines, of *cidB_IV*(2), also referred to as the *cidB_IV*(2) downstream region (dr), accounts for the presence of the *cidB_IV*(a/2) and/or *cidB_IV*(b/2) variants.

A specific PCR test was also designed to detect the presence of *cidB_IV*(2) dr (which also accounts for the *cidB_IV*(a/2) and/or *cidB_IV*(b/2) variants), to confirm the PCR-RFLP test results in isofemale lines. A 107 bp fragment was amplified with the primers CidB_QPCR_spe_2_dir3 (5'-GGG-AAT-AGT-GCT-TTT-GAT-AGA-GAG-TA) and CidB_QPCR_spe_rev1 (5'-GTT-AAA-CAT-CTT-AAA-CCC-TCA-TCA-CC), under the following PCR conditions: 5 min at 94°C, followed by 35 cycles of 94°C for 30 s, 58°C for 45 s, and 72°C for 30 s, with a final elongation for 5 min at 72°C.

Specific detection of the cidB_IV(a/2) variant

The *cidB_IV(a/2)* variant associated with the “incompatible” *mod* phenotype in natural populations from North Africa in the study by Bonneau et al. (2018)[22] was clearly identified only by the cloning and Sanger sequencing of the *cidB* gene repertoire. We therefore named a variant as *cidB_IV(a/2)* only in situations in which cloning and Sanger sequencing experiments were performed.

Real-time quantitative PCR

Quantification of Wolbachia density in male testes

We quantified the density of *Wolbachia* in the testes of males from the Michele26 and Mezzo9 lines, by quantitative PCR with the LightCycler 480 system (Roche). Specific primer pairs and procedures were described by Berticat et al. (2002)[29]. Each DNA template was obtained from pools of three pairs of testes from six-day-old males, with the DNeasy Blood & Tissue Spin-Column kit (Qiagen; bench protocol: animal tissues). Five independent DNA templates were used for each line. We estimated the number of *Wolbachia* bacteria per mosquito testis, by amplifying two different genes for each sample: the *C. pipiens*-specific *ace-2* locus [30] and the *Wolbachia*-specific single-copy *wsp* locus (Berticat et al. 2002)[29]. Standard curves were generated with dilutions of a pBluescriptKS vector containing single copies of the *ace-2* and *wsp* genes. Each DNA template was analyzed in triplicate for the quantification of both *wsp* and *ace-2*. As both genes were present as single copies per haploid genome, the ratio of the *wsp* and *ace-2* signals could be used to estimate the relative number of *Wolbachia* genomes per *Culex* genome, thus correcting for mosquito size and DNA quality.

Expression of the cidA and cidB genes

For the Mezzo9 and Michele26 lines, we extracted RNA from 10 six-day-old males with Trizol (Life Technologies). The RNA was then treated with DNase with the TURBO DNA-free™ Kit (Life Technologies), in accordance with the manufacturer's instructions. We subjected 2-5 µg of each total

RNA sample to reverse transcription with the SuperScript III Reverse Transcriptase Kit and 30 ng of random oligomer primers ((RP)10; Invitrogen, Life Technologies). Four different quantitative PCRs were performed on the resulting cDNA, according to the procedure described by Berticat et al. (2002)[29]. The first was specific for the *wsp* locus, as described by Berticat et al. (2002)[29]. The second was specific for a 189 bp fragment of the *cidA* gene conserved in all sequenced *wPip* strains, and was performed with the primers *wPip_0282_QPCR_2_Dir* (5'-AGG-TCC-TGT-ATT-TGA-TTT-CTG-GA) and *wPip_0282_QPCR_2_Rev* (5'-TGA-ACG-CGA-GAA-AGA-GCA-AG). The third was specific for a 135 bp fragment of the *cidB* gene conserved in all sequenced *wPip* strains and was performed with the primers *wPip_0283_QPCR_1_Dir* (5'-TGA-GTG-TTT-GGA-GAA-TGA-AGG-A) and *wPip_0283_QPCR_1_Rev* (5'-TTC-CCA-AAA-GCA-AAA-CCA-GTT). The fourth was specific for the 107 bp fragment of *cidB_IV(2)* dr described above. Each DNA template was analyzed in triplicate for *wsp*, *cidA*, *cidB* and *cidB_IV(2)* dr. Standard curves were generated for the *cidA*, *cidB*, *cidB_IV(2)* and *wsp* genes, by diluting the PCR products for these four genes. Expression levels for the *cidA*, *cidB* and *cidB_IV(2)* genes were estimated relative to that of the *wsp* gene.

Hatching rate estimation

Crosses were performed between males from the Michele26 and Luna8 isofemale lines and females from the Tunis line, and hatching rates were compared. We crossed 50 males from each of the Michele26 and Luna8 isofemale lines with 100 females from the Tunis line. The females were kept in cages for six days and were then fed turkey blood. Five days later, egg-rafts were collected and deposited in 24-well plates. Hatched larvae were allowed to die and pictures of both eggs and larvae were taken for 30 egg-rafts per cross, with a Leica M205 C stereomicroscope equipped with a Leica DFC420 C digital camera. Eggs and larvae were counted by eye with ImageJ [31]. Hatching rates were calculated as the number of larvae divided by the number of eggs.

Statistical analysis

All analyses were performed with R version 3.4.4 software [32]. Comparisons between the real-time quantitative variables of the Michele26 and Mezzo9 isofemale lines and of hatching rate between the Michele26 and Luna8 isofemale lines were performed with the nonparametric Wilcoxon rank-sum test [33].

Supplementary information

See Annexes at the end of the thesis manuscript

Data accessibility

The nucleotide and amino-acid sequences of the *cidA-cidB* variants were deposited in GenBank and accession numbers are provided in Supplementary Table 1. The authors declare that all other data supporting the findings of this study are available within the article and its supporting information files.

Authors' contributions

M.B., M.S., and M.W. conceptualized and designed the study; M.B., B.C., D.A., M.S. and M.W. field sampled the mosquitoes; M.B., A.L., R.C., S.U. performed the experiments; M.B., M.S., and M.W. analyzed and interpreted the data; M.B., B.C., R.C., M.S., and M.W. wrote the manuscript.

Competing interest

We have no competing interest.

Funding

This work was funded by the French ANR (project “CIAWOL” ANR-16- CE02-0006-01).

Acknowledgments

We would like to thank Bertrand Lelièvre for his help with mosquito sampling and Nicole Pasteur for helpful comments on the manuscript. We thank Justy Fabienne for her help in RNA extraction and Dr Laurent Marivaux for the use of the stereomicroscope funded by PALASIAFRICA ANR/ERC. We also thank Dr Philippe Clair for his help with the real-time quantitative PCR experiments performed at the technical facility of the “qPCR Haut Debit (qPHD) Montpellier génomiX” platform. Sequencing data were generated on the GENSEQ platform of the technical facilities of the LabEX "Centre Méditerranéen de l'Environnement et de la Biodiversité". The CI status of crosses was determined at the CytoEvol facilities of UMR ISEM - CBGP of the LabEx CeMEB.

References

1. Werren JH, Baldo L, Clark ME. 2008 *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**, 741–751. (doi:10.1038/nrmicro1969)
2. Taylor MJ, Bandi C, Hoerauf A. 2005 *Wolbachia* bacterial endosymbionts of filarial nematodes. In *Advances in Parasitology*, pp. 245–284. (doi:10.1016/S0065-308X(05)60004-8)
3. Ferri E *et al.* 2011 New insights into the evolution of *Wolbachia* infections in filarial nematodes inferred from a large range of screened species. *PLoS One* **6**. (doi:10.1371/journal.pone.0020843)
4. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. 2008 How many species are infected with *Wolbachia*? a statistical analysis of current data. *FEMS Microbiol. Lett.* **281**, 215–220. (doi:10.1111/j.1574-6968.2008.01110.x)

5. Zug R, Hammerstein P. 2012 Still a Host of Hosts for *Wolbachia*: Analysis of Recent Data Suggests That 40% of Terrestrial Arthropod Species Are Infected. *PLoS One* **7**, e38544. (doi:10.1371/journal.pone.0038544)
6. Weinert LA, Araujo-Jnr E V., Ahmed MZ, Welch JJ. 2015 The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc. R. Soc. B Biol. Sci.* **282**, 20150249–20150249. (doi:10.1098/rspb.2015.0249)
7. Yen JH, Barr AR. 1973 The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *J. Invertebr. Pathol.* **22**, 242–250. (doi:10.1016/0022-2011(73)90141-9)
8. Callaini G, Riparbelli MG, Giordano R, Dallai R. 1996 Mitotic Defects Associated with Cytoplasmic Incompatibility in *Drosophila simulans*. *J. Invertebr. Pathol.* **67**, 55–64. (doi:10.1006/jipa.1996.0009)
9. Laven H. 1967 Speciation and evolution in *Culex pipiens*. In *Genetics of Insect Vectors of Disease* (eds J Wright, R Pal), pp. 251–275. Elsevier: Amsterdam.
10. O'Neill SL, Karr TL. 1990 Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature* **348**, 178–180. (doi:10.1038/348178a0)
11. Breeuwer JAJ, Werren JH. 1990 Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature* **346**, 558–560. (doi:10.1038/346558a0)
12. Bordenstein SR, O'Hara FP, Werren JH. 2001 *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* **409**, 707–710. (doi:10.1038/35055543)
13. Duron OI, Bernard C, Unal S, Berthomieu A, Berticat C, Weill M. 2006 Tracking factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. *Mol. Ecol.* **15**, 3061–3071. (doi:10.1111/j.1365-294X.2006.02996.x)
14. Atyame CM, Duron O, Tortosa P, Pasteur N, Fort P, Weill M. 2011 Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito

- populations. *Mol. Ecol.* **20**, 286–298. (doi:10.1111/j.1365-294X.2010.04937.x)
15. Atyame CM, Labbé P, Dumas E, Milesi P, Charlat S, Fort P, Weill M. 2014 *Wolbachia* Divergence and the Evolution of Cytoplasmic Incompatibility in *Culex pipiens*. *PLoS One* **9**, e87336. (doi:10.1371/journal.pone.0087336)
 16. Hurst LD. 1991 The evolution of cytoplasmic incompatibility or when spite can be successful. *J. Theor. Biol.* **148**, 269–277. (doi:10.1016/S0022-5193(05)80344-3)
 17. Werren JH. 1997 Biology of *Wolbachia*. *Annu. Rev. Entomol.* **42**, 587–609.
 18. Beckmann JF, Ronau JA, Hochstrasser M. 2017 A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2**, 17007. (doi:10.1038/nmicrobiol.2017.7)
 19. LePage DP *et al.* 2017 Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **543**, 243–247. (doi:10.1038/nature21391)
 20. Lindsey ARI, Rice DW, Bordenstein SR, Brooks AW, Bordenstein SR, Newton ILG. 2018 Evolutionary Genetics of Cytoplasmic Incompatibility Genes *cifA* and *cifB* in Prophage WO of *Wolbachia*. *Genome Biol. Evol.* **10**, 434–451. (doi:10.1093/gbe/evy012)
 21. Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. 2018 One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **115**, 4987–4991. (doi:10.1073/pnas.1800650115)
 22. Bonneau M, Atyame C, Beji M, Justy F, Cohen-Gonsaud M, Sicard M, Weill M. 2018 *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat. Commun.* **9**. (doi:10.1038/s41467-017-02749-w)
 23. Atyame CM *et al.* 2015 Stable coexistence of incompatible *Wolbachia* along a narrow contact zone in mosquito field populations. *Mol. Ecol.* **24**, 508–521. (doi:10.1111/mec.13035)
 24. Dumas E, Atyame CM, Milesi P, Fonseca DM, Shaikevich E V, Unal S, Makoundou P, Weill M, Duron O. 2013 Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol. Biol.* **13**, 181. (doi:10.1186/1471-2148-13-181)

25. Altinli M, Gunay F, Alten B, Weill M, Sicard M. 2018 *Wolbachia* diversity and cytoplasmic incompatibility patterns in *Culex pipiens* populations in Turkey. *Parasites and Vectors* **11**, 1–9. (doi:10.1186/s13071-018-2777-9)
26. Duron O, Lagnel J, Raymond M, Bourtzis K, Fort P, Weill M. 2005 Transposable element polymorphism of *Wolbachia* in the mosquito *Culex pipiens*: Evidence of genetic diversity, superinfection and recombination. *Mol. Ecol.* **14**, 1561–1573. (doi:10.1111/j.1365-294X.2005.02495.x)
27. Duron O, Weill M. 2006 *Wolbachia* infection influences the development of *Culex pipiens* embryo in incompatible crosses. *Heredity.* **96**, 493–500. (doi:10.1038/sj.hdy.6800831)
28. Gouy M, Guindon S, Gascuel O. 2010 SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol. Biol. Evol.* **27**, 221–224. (doi:10.1093/molbev/msp259)
29. Berticat C, Rousset F, Raymond M, Berthomieu A, Weill M. 2002 High *Wolbachia* density in insecticide-resistant mosquitoes. *Proc. R. Soc. B Biol. Sci.* **269**, 1413–1416. (doi:10.1098/rspb.2002.2022)
30. Weill M, Berticat C, Raymond M, Chevillon C. 2000 Quantitative Polymerase Chain Reaction to Estimate the Number of Amplified Esterase Genes in Insecticide-Resistant Mosquitoes. *Anal. Biochem.* **285**, 267–270. (doi:10.1006/abio.2000.4781)
31. Schneider CA, Rasband WS, Eliceiri KW. 2012 NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675. (doi:10.1038/nmeth.2089)
32. R Core Team. 2018 R: A language and environment for statistical computing.
33. Bauer DF. 1972 Constructing confidence sets using rank statistics. *J. Am. Stat. Assoc.* **67**, 687–690.

II. Résumé des résultats & Discussion

a. L'opéron *cidA/cidB* est amplifié et polymorphe dans les génomes des *wPips*

Le séquençage Illumina de trois souches de *wPips* du laboratoire et le mapping des reads issus de ces séquençages sur le génome de référence *wPipPel* ont permis de mettre en évidence que les gènes *cidA* et *cidB* chez les *wPips* séquencées possèdent une couverture au minimum deux fois plus élevée que celle des gènes de ménage présents en une seule copie dans le génome. Ce résultat suggère que chez les *wPips* séquencées ces gènes sont amplifiés au sein d'un même génome. Deuxièmement, le mapping a permis de révéler chez ces deux gènes la présence de polymorphisme entre les *wPips* mais aussi au sein d'une même souche. Ces observations suggèrent que les différentes copies respectives de *cidA* et *cidB* présentes dans les génomes des *wPips* sont différentes au sein d'un même génome et entre les différents génomes. Ces résultats révèlent une diversité insoupçonnée de copies de *cidA* et *cidB* résultant probablement d'événements de duplication et de divergence qui pourraient être à l'origine de la diversité des phénotypes d'IC observés chez *C. pipiens*. En revanche, aucun polymorphisme n'a été détecté dans les gènes *cinA* et *cinB*, que ce soit au sein d'un génome de *wPip* ou que ce soit entre souches, ce qui suggère que ces gènes potentiellement impliqués dans l'IC chez *C. pipiens* ne sont pour autant pas impliqués dans la diversité des phénotypes d'IC. Par conséquent, la suite des analyses des bases génétiques de la diversité de phénotypes de l'IC chez *C. pipiens* portera uniquement sur l'opéron *cidA/cidB*.

b. La variabilité de l'opéron *cidA/cidB* est très probablement responsable de la diversité des phénotypes d'IC chez *C. pipiens*

Pour que les gènes *cidA* et *cidB* soient effectivement impliqués dans la diversité des phénotypes d'IC chez *C. pipiens*, il faut que des souches de *wPips* induisant des phénotypes d'IC différents possèdent des répertoires de variants de *cidA* et *cidB* différents. L'analyse par qPCR du nombre de copies de *cidA* et *cidB* dans les génomes des *wPips* appartenant à des groupes phylogénétiques (I à IV) de *wPips* différents et induisant quatre phénotypes d'IC différents a révélé que ces gènes sont bien présents en plusieurs copies dans chaque génome. Cependant, ces gènes étant localisés dans le génome de *wPipPel* dans une zone de prophage, il est possible qu'ils soient contenus dans des phages en phase lytique (*i.e.* sous forme de particule virale), ce qui pourrait fausser l'estimation du nombre de copie.

Pour étudier la diversité des gènes *cidA* et *cidB* au sein de ces quatre souches, une stratégie classique de clonage et séquençage a été employée. Ces quatre souches de *wPips* présentent toutes des répertoires de variants *cidA* et *cidB* différents ce qui conforte l'implication de ces deux gènes dans la diversité de l'IC. De plus, aucun variant de *cidA* et *cidB* n'a été trouvé en commun entre les quatre groupes de *wPip* suggérant que les variants de *cidA* et *cidB* suivent la phylogénie des groupes de *wPips*. Cependant, les

analyses phylogénétiques en réseau des séquences de ces variants montrent que les différences qui les discriminent se présentent sous la forme de combinaisons de blocs de polymorphisme partagés pour certains entre les différents groupes alors que d'autres sont spécifiques à un groupe. C'est par exemple le cas des variants observés dans le groupe *wPipIV* qui présentent, que ce soit pour *cidA* ou pour *cidB*, des blocs de polymorphisme non observés dans les autres groupes. Ce patron de polymorphisme est très probablement dû à des événements de recombinaison de blocs de séquences au sein et entre les différents génomes de *wPips*. Ces événements de recombinaison pourraient résulter d'une coïnfection par différentes souches de *wPips* bien que de telles coïnfections n'aient jamais été décrites chez *C. pipiens*. Une deuxième hypothèse pour expliquer ces événements de recombinaison fait intervenir la présence de particules virales de phages lytiques du fait de l'association de ces gènes avec le prophage WO (Wright *et al.*, 1978). En effet, si les particules virales contiennent ces gènes (ce qui semble être le cas d'après des communications scientifiques lors du congrès *Wolbachia* 2018), la recombinaison entre les particules virales pourrait permettre l'apparition de nouveaux variants au sein d'une *Wolbachia* mais aussi entre *Wolbachia* lors de phases d'infection.

c. Le groupe *wPipIV* lève le voile sur les rôles de *cidA* et *cidB* dans l'IC chez *C. pipiens*

Afin d'associer des répertoires de *cidA* et *cidB* plus spécifiquement à des phénotypes d'IC, nous avons étudié une situation de phénotypes d'IC plus simple au sein du groupe *wPipIV*. Nous possédons au laboratoire des lignées de *C. pipiens* infectés par des *wPipIV* qui induisent le même phénotype *resc* mais deux phénotypes *mod* différents : les mâles de certaines lignées vont être incompatibles avec n'importe quelles femelles infectées par des *wPipI*, II ou III alors que d'autres vont être compatibles avec ces mêmes femelles. Ainsi pour que *cidA* et/ou *cidB* soient impliqués dans le *mod*, il faut que les lignées « incompatibles » et « compatibles » présentent des répertoires de variants *cidA* et *cidB* différents et que certains de ces variants *cidA* et/ou *cidB* soient uniquement associés à l'un ou l'autre de ces deux phénotypes *mod*.

L'analyse des répertoires de 6 de ces isofemelles provenant de Chine, de Turquie, d'Algérie et de Tunisie a montré la présence de plusieurs variants de *cidA* et *cidB* dans chacun de leurs répertoires et les lignées isofemelles « compatibles » et « incompatibles » présentent des répertoires de *cidA* et de *cidB* différents. Que ce soit pour le gène *cidA* ou le gène *cidB*, le polymorphisme qui discrimine les variants est divisé en deux zones : une zone en amont (déterminée par une lettre grecque pour *cidA* ou latine pour *cidB*) et une zone en aval du gène (déterminée par un chiffre). Ainsi, les variants *cidA_IV*(α /1) et *cidA_IV*(α /2) présentent la même séquence en amont mais deux séquences différentes en aval. Inversement, les variants *cidB_IV*(a/1) et *cidB_IV*(b/1) présentent la même séquence en aval mais deux séquences différentes en amont. La comparaison des variants de *cidA* et de *cidB* avec les phénotypes d'IC révèle que ce sont les zones polymorphes en amont du gène de *cidA* (α , β , γ et δ) et en aval du gène

de *cidB* (1, 2 et 3) qui sont corrélées aux phénotypes d'IC. En effet, les séquences « 1 » et « 2 » en aval de *cidA* sont retrouvées aussi bien chez les lignées « incompatibles » que « compatibles », en revanche seules les lignées « incompatibles » possèdent la région en amont « δ ». Ainsi, le variant *cidA_IV*(δ) a uniquement été détecté chez les ω PipIV infectant des lignées « incompatible ». Inversement pour *cidB*, les séquences « a » et « b » en amont sont retrouvées aussi bien chez les lignées « incompatibles » que « compatibles », en revanche seules les lignées « incompatibles » possèdent la région en aval « 2 ». Ainsi le variant *cidB_IV*(2) a uniquement été détecté chez les ω PipIV infectant des lignées « incompatibles ». Plus spécifiquement, seul le variant *cidB_IV*(a/2) a été observé chez toutes les lignées incompatibles car *cidB_IV*(b/2) n'est présent que chez la lignée Istanbul.

Les variants *cidA_IV*(δ) et *cidB_IV*(a/2) ont uniquement été détectés chez les lignées « incompatibles », ce qui suggère leur implication dans le phénotype *mod* « incompatible » et par conséquent l'implication de *cidA* et de *cidB* dans la fonction *mod*. Par ailleurs, le variant *cidA_IV*(α) a été observé chez les 6 ω PipIV analysées, indépendamment de leur origine géographique. La présence d'un tel variant ubiquitaire est attendu dans le cas du facteur *resc* car il pourrait être responsable de la compatibilité toujours observée lorsque deux moustiques infectés par des ω Pips du même groupe sont croisés (Atyame *et al.*, 2014, 2015b). Ce résultat suggère l'implication de *cidA* dans la fonction *resc*.

Afin de conforter l'ubiquité du variant *cidA_IV*(α) et l'association des variants *cidA_IV*(δ) et *cidB_IV*(a/2) avec le phénotype *mod* « incompatible », des analyses de populations naturelles de 174 lignées isofemelles provenant de Tunisie et d'Algérie ont été réalisées. Ces populations comportent des *C. pipiens* infectés par des ω PipIV qui comme nos lignées de laboratoire présentent soit le phénotype *mod* « compatible » soit le phénotype *mod* « incompatible ». Pour analyser ces populations, deux tests PCR-RFLP permettant de discriminer les variants de *cidA* et *cidB* basés sur la zone de polymorphisme associée au phénotype *mod* dans chacun de ces deux gènes ont été mis en place. Le test PCR-RFLP associé au gène *cidA* permet donc de discriminer *cidA_IV*(α), *cidA_IV*(β) et *cidA_IV*(δ) indépendamment de la zone polymorphe en aval du gène. Ainsi la détection de *cidA_IV*(α) dans une isofemelle indique la présence de *cidA_IV*(α /1) et/ou *cidA_IV*(α /2). Le test PCR-RFLP associé au gène *cidB* permet donc de discriminer *cidB_IV*(1), *cidB_IV*(2) et *cidB_IV*(3) indépendamment de la zone polymorphe en amont du gène. Ainsi, la détection de *cidB_IV*(2) chez une isofemelle indique la présence de *cidB_IV*(a/2) et/ou *cidB_IV*(b/2).

L'utilisation de ces deux tests dans les populations naturelles de Tunisie et d'Algérie a confirmé la présence chez toutes les ω PipIV infectant ces lignées isofemelles du variant *cidA_IV*(α), supportant son rôle dans la fonction *resc*. De plus, les variants *cidB_IV*(2) et *cidA_IV*(δ) ont été détectés chez toutes les lignées « incompatibles » de ces deux populations naturelles confortant leur implication dans la fonction *mod* et le phénotype *mod* « incompatible ». En outre, les variants *cidB_IV*(2) et *cidA_IV*(δ) ont été détectés chez quelques isofemelles présentant le phénotype *mod* « compatible », ce qui suggère que

soit la caractérisation au laboratoire du phénotype d'IC « compatible » de ces lignées n'est peut-être pas bonne, soit des phénomènes génétiques ou épigénétiques viendraient perturber le lien entre génotype et phénotype de l'IC chez ces souches. Cependant, l'absence de telles lignées tunisiennes ou algériennes « discordantes » en élevage au laboratoire ne nous a pas permis d'investiguer cette incongruence.

L'analyse des variants *cidA_IV* et *cidB_IV* montre que la zone polymorphe associée aux phénotypes d'IC est restreinte à un domaine spécifique d'environ 30 acides aminés sur chaque protéine. Ces domaines présentent des motifs typiquement associés à des domaines d'interaction protéines-protéines (domaine ankyrine ou Heat repeat) et leur variabilité pourrait très probablement être impliquée dans l'interaction spécifique entre un variant de CidA et un variant de CidB. De telles variabilités d'interaction entre variants CidA et CidB pourraient être à l'origine de la très grande diversité de phénotypes d'IC observée chez *C. pipiens*.

Grâce à cette analyse populationnelle nous avons pu conforter le rôle de *cidB* dans la fonction *mod* et la diversité des phénotypes d'IC. De plus, nos résultats semblent indiquer que *cidA* pourrait jouer à la fois un rôle dans le *mod* et dans le *resc* comme proposé par Shropshire et al. (2018) chez la drosophile. En revanche, il existe une hypothèse plus parcimonieuse des rôles de *cidA* et *cidB* dans les fonctions *mod* et *resc* où *cidB* serait impliqué dans le *mod* et *cidA* uniquement dans le *resc*. Ainsi, la corrélation de *cidA_IV*(δ) avec le phénotype *mod* « incompatible » dans les populations naturelles de Tunisie et d'Algérie infectées par des *wPipIV* pourrait résulter de l'interaction spécifique mise en évidence entre les protéines CidA et CidB (Beckmann *et al.*, 2017). En effet, du fait de leur interaction spécifique il est possible qu'elles aient co-diversifié entraînant ainsi localement l'association entre *cidA_IV*(δ) et *cidB_IV*(a/2) indépendamment de l'implication de *cidA* dans la fonction *mod*

Afin de clarifier le rôle de *cidA* dans l'IC chez *C. pipiens* mais également dans le but d'élever une lignée « discordante » au laboratoire, nous avons mené le même type d'expérience dans des populations naturelles d'Italie du nord précédemment décrites comme infectées par des *wPipIV* (Dumas *et al.*, 2013). Nous sommes donc allés collecter des populations naturelles de *C. pipiens* en Italie du nord dans les caniveaux et autres gîtes à moustiques des localités San Michele, Rovere Della Luna, Mezzocorona et Zambana. Nous avons, au laboratoire, obtenus 67 lignées isofemelles dont 5 présentent le phénotype d'IC « compatible » et 62 le phénotype d'IC « incompatible ». Le séquençage Sanger des gènes *cidB* des *wPipIV* infectant 4 lignées « incompatibles » et 2 lignées « compatibles » a mis en évidence la présence chez les 4 lignées « incompatibles » des variants *cidB_IV*(a/2) et *cidB_IV*(b/2) et leur absence dans les lignées « compatibles », ce qui indique que dans ces populations italiennes ces deux variants sont associés au phénotype *mod* « incompatible ». Ce résultat en Italie du nord conforte le fait que c'est la partie polymorphe en aval du gène *cidB* (*i.e.* la séquence « 2 ») qui est associée au phénotype *mod*. L'analyse par PCR-RFLP du gène *cidB* des 67 lignées isofemelles a permis de confirmer la présence des variants *cidB_IV*(2) [*cidB_IV*(a/2) et *cidB_IV*(b/2) n'étant pas discriminables avec ce test] dans toutes

les lignées incompatibles et sa présence dans une lignée isofemelle « compatible » : Michele26 (lignée « discordante » tant espérée que nous avons réussie à élever au laboratoire). L'association du variant *cidB_IV(2)* avec le phénotype *mod* « incompatible » conforte les résultats de transgénèse chez *D. melanogaster* qui impliquaient *cidB* dans la fonction *mod* (Beckmann *et al.*, 2017; LePage *et al.*, 2017).

Le séquençage du gène *cidA* a révélé la présence chez toutes les *wPipIV* du variants *cidA_IV(α)* et l'absence totale du variant *cidA_IV(δ)*. De plus, aucun autre variant de *cidA* n'a été observé comme associé avec l'un ou l'autre des deux phénotypes *mod*. L'analyse par PCR-RFLP de ce même gène dans les 67 isofemelles a confirmé la présence chez chacune d'entre elles du variant *cidA_IV(α)* et l'absence de *cidA_IV(δ)*. La présence du variant *cidA_IV(α)* chez toutes les isofemelles infectées par des *wPipIV* d'Italie du nord soutient l'implication de *cidA* dans la fonction *resc* chez *C. pipiens*, ce qui est en accord avec les résultats chez *D. melanogaster* (Shropshire *et al.*, 2018). En revanche, l'absence d'association du variant *cidA_IV(δ)* avec le phénotype *mod* « incompatible » en Italie du nord ainsi que l'absence d'association de quelques variants *cidA* que ce soit avec un des deux phénotypes *mod* suggèrent que *cidA* n'est pas impliqué dans la fonction *mod* chez *C. pipiens* contrairement à ce qui a été observé chez *D. melanogaster* (LePage *et al.*, 2017). Ainsi, il semble que l'opéron *cidA/cidB* ne fonctionne pas exactement de la même manière chez *D. melanogaster* où *cidA* est impliqué dans la fonction *mod* et dans la fonction *resc* en fonction du sexe chez lequel il est exprimé alors que chez *C. pipiens* *cidA* ne semble pas impliqué dans la fonction *mod*.

Nous nous sommes ensuite intéressés à la lignée « discordante » Michele26 afin de comprendre comment une lignée infectée par une souche *wPipIV* portant le variant associé au phénotype *mod* « incompatible » pouvait tout de même être compatible. Nous avons donc i) séquencé les répertoires de variants *cidA* et *cidB* de la souche de *wPipIV* infectant cette lignée, ii) étudié l'expression de *cidA*, de *cidB* et celle du variant *cidB_IV(2)* ainsi que quantifié la densité de *Wolbachia* dans les testicules de cette lignée et de la lignée « incompatible » Mezzo9 et enfin iii) estimé le degré de compatibilité de cette souche (taux d'éclosion) en comparaison avec la lignée « compatible » Luna 8. Premièrement, le degré de compatibilité de Michele26 est identique à celui d'une souche « compatible » normale [*i.e.* ne possédant pas *cidB_IV(2)*]. Deuxièmement, la souche *wPipIV* qui infecte Michele26 possède un répertoire de variant de *cidA* qui ne présente aucune particularité par rapport aux lignées « compatibles » et « incompatibles ». En revanche, dans son répertoire *cidB* deux variants qui n'avaient jamais été détectés chez aucune autre *wPipIV* auparavant ont été observés. Il est possible que la présence de nouveaux variants de *cidB* puisse entraîner un effet de titration empêchant la protéine CidB_IV(2) d'interagir avec ses cibles et donc d'induire l'IC. De plus, contrairement à toutes les *wPipIV* infectant des lignées incompatibles qui ont été séquencées en Italie du nord, Michele26 ne possède pas le variant *cidB_IV(b/2)*. On aurait pu penser que ce soit l'association des variants *cidB_IV(a/2)* et *cidB_IV(b/2)* qui permette d'être incompatible. Or, en Tunisie et Algérie, la présence du variant *cidB_IV(a/2)* seul suffit à induire l'IC, ce qui réfute cette hypothèse. L'analyse de l'expression des gènes *cidA* et *cidB* ne

révèle aucune différence de niveau d'expression entre la souche « discordante » Michele26 et la souche « incompatible » Mezzo9. En revanche, l'analyse de l'expression du variant *cidB_IV(2)* qui comprend l'expression de *cidB_IV(a/2)* et *cidB_IV(b/2)* révèle que ce variant est exprimé chez Michele26, mais deux fois moins que chez la lignée isofemelle « incompatible » Mezzo9. Cette différence est probablement due au fait que contrairement à Mezzo9 qui possède *cidB_IV(a/2)* et *cidB_IV(b/2)*, Michele26 ne possède que le variant *cidB_IV(a/2)*. Ainsi, les *wPipIV* infectant Michele26 produisent deux fois moins de protéines *cidB_IV(2)* que celles de Mezzo9. De plus, les testicules des mâles Michele26 sont infectés par significativement moins de *Wolbachia* que ceux des mâles de la lignée « incompatible » Mezzo9. Ainsi, la présence de moins de *Wolbachia* dans les testicules des mâles Michele26 qui expriment deux fois moins *cidB_IV(2)* pourrait être responsable d'une production totale de protéine *CidB_IV(2)* plus faible chez Michele26 que chez Mezzo9. Cette quantité pourrait être trop faible pour induire l'IC. Cependant, l'acquisition d'autres lignées « discordantes » de ce type est nécessaire pour confirmer cette hypothèse.

La très grande diversité de phénotypes d'IC observée chez les moustiques *C. pipiens* est gouvernée par la présence dans les génomes des *Wolbachia* qui les infectent de l'opéron *cidA/cidB* présent en plusieurs copies polymorphes. L'association dans des populations naturelles de *C. pipiens* infectées par des *wPipIV* du variant *cidB_IV(2)* avec le phénotype *mod* « incompatible » supporte le rôle de *cidB* dans la fonction *mod*. La présence chez chacune des 247 isofemelles infectées par des *wPipIV* du variant *cidA_IV(a)* qui pourrait être responsable de la compatibilité réciproque toujours observée quand des *C. pipiens* infectés par des *wPipIV* sont croisés suggère l'implication de *cidA* dans la fonction *resc*. Chez *C. pipiens*, *cidB* serait impliqué dans la fonction *mod* et *cidA* dans la fonction *resc*. Il apparaît ainsi que l'opéron *cidA/cidB* semble chez *C. pipiens* fonctionner comme un simple système toxine-antidote comme proposé par Beckmann et al. (2013 & 2017) et par nos prédécesseurs Hurst (1991) et Werren (1997) qui avaient vu juste il y a déjà plus de 20 ans.

Partie II :

**Les mécanismes cytologiques
sous-jacents à l'IC et sa diversité chez
*Culex pipiens***

I. Introduction

La découverte de l'opéron *cidA/cidB* a permis une grande avancée dans la compréhension des bases génétiques de l'incompatibilité cytoplasmique induite par *Wolbachia*. En particulier, la découverte de la fonction déubiquitineuse de la protéine CidB ouvre de nouvelles voies de recherches sur les mécanismes moléculaires sous-jacents à l'IC. Cette découverte donne de nouveaux indices sur les protéines de l'hôte ciblées par *Wolbachia* pour induire l'IC et sur les voies moléculaires potentiellement impliquées. En effet, jusqu'à présent peu de choses étaient connues sur les mécanismes moléculaires et cellulaires de l'IC et aucune réelle avancée n'avait été faite sur ce sujet depuis près de 10 ans. Les pistes de recherche de ces protéines cibles et des mécanismes moléculaires étaient basées sur les observations en microscopie des défauts de développement embryonnaires, responsables de la mortalité des embryons issus de croisements incompatibles, chez *D. melanogaster*, *D. simulans* et *Nasonia spp.*. Le phénotype cellulaire responsable de la mort des embryons dans les croisements incompatibles (*i.e.* problème de condensation et de ségrégation de la chromatine paternelle) est connu depuis les années 60 (Ryan & Saul, 1968) et ce n'est que très récemment (de 2002 à 2009) que des avancées sur des mécanismes en amont de ces perturbations développementales ont été faites (Tram & Sullivan, 2002; Tram *et al.*, 2003, 2006; Landmann *et al.*, 2009). Le défaut de développement le plus précoce mis en évidence est un problème de dépôt des histones H3.3 lors du remodelage de la chromatine paternelle suite au retrait des protamines juste après la fécondation (Landmann *et al.*, 2009). Ce remodelage défectueux serait responsable d'une mauvaise réplication et condensation de la chromatine paternelle avant l'entrée en première division embryonnaire, ce qui empêcherait la ségrégation de celle-ci lors de l'anaphase. Dans certain cas, la chromatine paternelle peut être totalement exclue lors de l'anaphase entraînant la ségrégation de la chromatine maternelle en deux noyaux haploïdes. Ces noyaux peuvent continuer à se diviser et donner un embryon haploïde qui sera non viable chez les espèces diploïdes comme *D. melanogaster* ou *C. pipiens* (Callaini *et al.*, 1997; Duron & Weill, 2006), mais peut donner des mâles viables chez des espèces haplodiploïdes comme *N. vitripennis* ou *L. heterotoma* (Ryan & Saul, 1968; Vavre *et al.*, 2001; Bordenstein *et al.*, 2003; Tram *et al.*, 2006). Il a été proposé que ce soit l'intensité des défauts embryonnaires impactant la chromatine paternelle (*i.e.* degré de condensation) qui soit à l'origine de la production ou non d'embryons haploïdes. Les défauts embryonnaires les plus sévères entraîneraient une totale exclusion des chromosomes paternels lors de la première division embryonnaire et donc la production d'un embryon haploïde. Ainsi, la proportion d'embryons haploïdes provenant d'un croisement incompatible constitue un proxy de la fréquence de l'exclusion totale de la chromatine paternelle (*i.e.* un proxy de l'intensité de l'IC à l'échelle cellulaire).

Bien que l'IC et *Wolbachia* aient été découvertes chez *C. pipiens*, aucune étude n'a été menée sur les mécanismes cytologiques responsables de la mort des embryons dans les croisements incompatibles. Les particularités de l'IC chez *C. pipiens* suggèrent que des mécanismes cytologiques différents de ceux observés chez *Drosophila spp.* et *Nasonia spp.* pourraient être à l'œuvre chez cette

espèce. En effet, contrairement à *Drosophila spp.* et *Nasonia spp.*, *C. pipiens* présente une très grande diversité de phénotypes d'IC qui pourrait reposer sur une diversité de mécanismes cytologiques. De plus, alors que des embryons haploïdes sont observés dans les croisements incompatibles entre mâles infectés et femelles non infectées chez *Drosophila spp.* (Callaini *et al.*, 1997) et *Nasonia spp.* (Tram *et al.*, 2006), chez *C. pipiens*, aucun développement haploïde n'est observé lorsque les femelles ne sont pas infectées (Duron & Weill, 2006). Ainsi, il est possible que chez *C. pipiens* les mécanismes cellulaires responsables de la mort des embryons soient différents suivant que ces embryons soient issus de femelles infectées (par une souche incompatible) ou non infectées par *Wolbachia*.

Dans la partie I, nous avons mis en évidence la présence de plusieurs variants de *cidA* et de *cidB* dans tous les génomes des *wPips* que nous avons étudiés et leur implication dans la diversité des phénotypes d'IC chez *C. pipiens*. Cette diversité de variants pourrait également être à la base de la diversité de phénotypes d'IC à l'échelle cellulaire. Par ailleurs, l'amplification et la diversification des variants de *cidA* et *cidB* (mais plus probablement de *cidB* car impliqué dans la fonction *mod*) pourraient avoir une influence sur l'intensité de l'IC à l'échelle cellulaire.

En partant de ces hypothèses, nous nous sommes intéressés, dans cette partie, aux mécanismes cellulaires à l'origine de la mort des embryons issus de croisements incompatibles chez *C. pipiens*. Ceci afin de déterminer :

- i) Si les mécanismes de l'IC à l'échelle cytologique différaient de ceux déjà décrits chez *Drosophila spp.* et *Nasonia spp.* chez qui la diversité d'IC est réduite ;
- ii) S'il existait des variations dans ces mécanismes qui pourraient participer à la diversité des phénotypes d'IC chez *C. pipiens* ;
- iii) Si ces mécanismes étaient différents lorsque les femelles étaient infectées ou non par *Wolbachia* ;
- iv) Si l'amplification et la diversification des variants *cidA* et *cidB* pouvaient influencer sur l'intensité de l'IC à l'échelle cellulaire.

Pour ce faire, nous avons étudié différentes étapes (précoces et tardives) du développement d'embryons issus de croisements incompatibles impliquant des lignées de *C. pipiens* infectées (ou non) par des souches de *Wolbachia* différentes appartenant à des groupes de *wPip* différents et induisant des phénotypes d'IC différents.

**Les phénotypes cellulaires de l'incompatibilité
cytoplasmique chez *Culex pipiens* au regard de
la diversité de *cidB***

Article n°3 publié dans Plos Pathogens

The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity

Manon Bonneau¹, Frédéric Landmann², Pierrick Labbé¹, Fabienne Justy¹, Mylène Weill¹, Mathieu Sicard^{1*}

¹ ISEM, University of Montpellier, CNRS, IRD, Montpellier, France

² CRBM, University of Montpellier, CNRS, Montpellier, France

*corresponding author

Email: mathieu.sicard@umontpellier.fr (MS)

Abstract

Wolbachia are maternally inherited endosymbiotic bacteria, widespread among arthropods thanks to host reproductive manipulations that increase their prevalence into host populations. The most commonly observed manipulation is cytoplasmic incompatibility (CI). CI leads to embryonic death in crosses between i) infected males and uninfected females and ii) individuals infected with incompatible *Wolbachia* strains. CI can be conceptualized as a toxin-antidote system where a toxin deposited by *Wolbachia* in the sperm would induce embryonic death unless countered by an antidote produced by *Wolbachia* present in the eggs. In *Drosophila melanogaster*, transgenic expression of *Wolbachia* effector *cidB* revealed its function of CI-inducing toxin. Moreover in *Culex pipiens*, the diversity of *cidB* variants present in *wPip* strains accounts for the diversity in crossing-types. We conducted cytological analyses to determine the CI mechanisms that lead to embryonic death in *C. pipiens*, and assess whether diversity in crossing-types could be based on variations in these mechanisms. We revealed that paternal chromatin condensation and segregation defects during the first embryonic division are always responsible for embryonic death. The strongest observed defects lead to an exclusion of the paternal chromatin from the first zygotic division, resulting in haploid embryos unable to hatch. The proportion of unhatched haploid embryos, developing with only maternal chromatin, which reflects the frequency of strong defects can be considered as a proxy of CI intensity at the cellular level. We thus studied the putative effect of variations in crossing types and *cidB* diversification on CI defects intensity. Incompatible crosses involving distinct *wPip* strains revealed that CI defects intensity depends on the *Wolbachia* strains hosted by the males and is linked to the diversity of *cidB* genes harbored in their genomes. These results support that, additionally to its implication in *C. pipiens* crossing type variability, *cidB* diversification also influences the strength of CI embryonic defects.

Author summary

In some crosses, mosquito males belonging to the species *Culex pipiens* prevent their females from having live progenies. This phenomenon called cytoplasmic incompatibility (CI) is caused by intracellular bacteria named *Wolbachia*. CI occurs when males infected with *Wolbachia* fertilize females

infected with genetically distinct incompatible *Wolbachia* resulting in the death of all the embryos. At the world scale, crossing relationships between *C. pipiens* are quite puzzling. Despite this complexity in crossing relationships and the diversity of *cidB* genes involved in CI mechanisms in *C. pipiens*, we demonstrate a single shared CI cellular phenotype leading to the death of the embryos: the paternal chromatin exclusion from the first embryonic division. If paternal chromatin is fully excluded, embryos developed with haploid set of chromosomes. We show that the frequency of haploid development varies according to the *Wolbachia* strains hosted by the males which differ in the *cidB* variants harbored in their genomes. Absence of *Wolbachia* in the eggs totally block haploid development showing that maternal *Wolbachia* presence interplays with CI mechanisms in a way that allows haploid development to occur. Understanding CI mechanism in mosquitoes is the corner stone to build new sustainable and adaptable *Wolbachia* based strategies for vector control.

Introduction

Wolbachia are maternally-inherited endosymbionts, widespread among arthropods and filarial nematodes [1,2], and the most frequent endocytobiotic bacteria detected in arthropods [3]. This high prevalence is attributed to their ability to manipulate their host reproduction to spread within arthropod populations [1]. The main reproductive manipulation strategy used by *Wolbachia* is named cytoplasmic incompatibility (CI) [4]. CI is a form of conditional sterility resulting in embryonic lethality [5]. In most of the host species, CI occurs when males infected with *Wolbachia* fertilize uninfected females whereas the reciprocal cross remains compatible. This difference in the production of viable offspring between infected and uninfected female reproduction enhances the spread of *Wolbachia* in host's populations [6]. CI can also occur between males and females both infected with different and incompatible *Wolbachia* strains [7–12]. In such situations, incompatibility can be either unidirectional (one cross direction is compatible while the reciprocal one is incompatible) or bidirectional (both cross directions are incompatible) [7–9]. The penetrance of CI, *i.e.* the number of embryos affected by CI in a cross, varies depending on the *Wolbachia* strain and the host involved in the interaction. Indeed, in the same host *Drosophila simulans*, wRi induces complete CI (*i.e.* crosses in which all the embryos are affected

by CI), while *wNo* and *wHa* strains induce lower levels of CI, *i.e.* some embryos can develop [13]. Complete CI penetrance was also described in *Nasonia spp.* depending on the *Wolbachia* strain involved and in all *Culex pipiens* incompatible crosses [14,15]. The variability of CI penetrance has been correlated to different factors such as the *Wolbachia* density in the sperm and eggs [16,17], host age [18] and host nuclear genotype [19].

Cellular consequences of *Wolbachia*-induced CI have been monitored during embryogenesis in *D. simulans*, *D. melanogaster* and *N. vitripennis* [5,20–23]. These studies revealed common cellular defects in these three species: a delay in paternal chromatin condensation and segregation defect during the first mitotic division of the embryo [23–25]. In *D. melanogaster*, a delay in histone H3.3 deposition after protamine removal on the paternal chromatin was observed and linked to chromatin remodeling defects [22]. This remodeling defect was associated with the persistence of the DNA replication factor PCNA (Proliferating Cell Nuclear Antigen) during mitosis, reflecting incomplete replication of paternal DNA [22]. It has been envisioned that these DNA replication defects might be responsible for the paternal chromatin bridges and segregation failure during the first mitotic division, which result in early embryonic arrest [22,24,25]. Nevertheless, some embryos reaching late development stages have been reported in CI crosses [21,25,26]. Late development in CI embryos was interpreted as resulting from a complete paternal chromatin exclusion during the first division, which allows successful maternal chromatin segregation and the formation of two haploid nuclei [25]. These haploid nuclei which further divide, lead to gynogenetic development (*i.e.* haploid development with only maternal genetic material) until late embryonic stages [27]. These haploid embryos are never viable in diploid species such as *D. simulans* [25]. However, in the haplodiploid parasitoid wasps *Leptopilina heterotoma* and *N. vitripennis*, CI-induced paternal chromosome defects can lead either i) to the death of the embryos or ii) to the production of healthy males [19,21,28,29]. It has been proposed that these two CI developmental outcomes could result from different degree of paternal chromatin defect (improper condensation) before the first division [29–31]. Severe defects would lead to complete elimination of male chromosomes from the first zygotic division resulting in haploidization and male development, whereas less severe defects would lead in partial exclusion of the paternal chromatin resulting in incomplete

elimination of male chromosomes and early arrest of the aneuploid development [29–31]. In diploid species such as *C. pipiens* and *D. melanogaster*, the proportion of unhatched developed haploid embryos observed in fully incompatible CI crosses would be a proxy of the frequency of total paternal chromatin exclusion during embryogenesis due to strong CI intensity at the cellular level. However, this hypothesis is counter intuitive as one could expect that strong CI defects would prevent any development to occur while soft CI defect would allow development.

The molecular mechanism underlying CI can be conceptualized as a toxin-antidote system in which i) a toxin produced by *Wolbachia* in the testes, more generally called a “*mod* factor”, and introduced in the sperm during spermatogenesis would interfere (“modify”) with the paternal chromatin and induce embryonic perturbations, and ii) an antidote released by *Wolbachia* in the egg, more generally called “*resc* factor”, would “rescue” these paternal chromatin defects to allow normal embryogenesis to occur [32,33]. The recent discoveries of i) *Wolbachia* genes *cidA* and *cidB* ability to recapitulate the CI phenotypes when expressed in transgenic *Drosophila* [34,35], and ii) the link between specific allelic *cidA*^{wPip}/*cidB*^{wPip} variations in worldwide natural *C. pipiens* populations and the capacity of males to sterilize females [36], open new paths into understanding CI mechanisms. *CidA* and *cidB* genes are syntenic genes within the WO phage region (S6 Table) [34,35,37–39]. *CidB* encodes a deubiquitylating enzyme (DUB) and when a *cidB*^{wPip} construct bearing this catalytically inactivated DUB domain was expressed in *D. melanogaster* males, CI was no longer observed, showing the implication of the deubiquitylating activity in the *mod* function [34]. The role of *cidA* in the CI mechanism is more debated because i) both *cidA*^{wMel} and *cidB*^{wMel} are required to induce CI in transgenic *Drosophila* [35] and ii) in natural populations of *C. pipiens*, specific *cidA*^{wPip} allelic variations were found to be linked to *mod* variations [36]. However, the implication of *cidA* in the *resc* function is supported by the capacity of *cidA* to prevent *cidB* toxicity in yeast [34] and the capability of transgenic uninfected females expressing *cidA*^{wMel} throughout oogenesis to rescue the effect of *cidB*^{wMel} [40].

In *C. pipiens*, all individuals are infected with different *Wolbachia* strains belonging to the monophyletic wPip group, but divided in five subgroups wPipI to wPipV. MLST (Multi Locus Sequence Typing) genes from Baldo et al. (2006) [41] were not polymorphic between wPip strains, thus a wPip

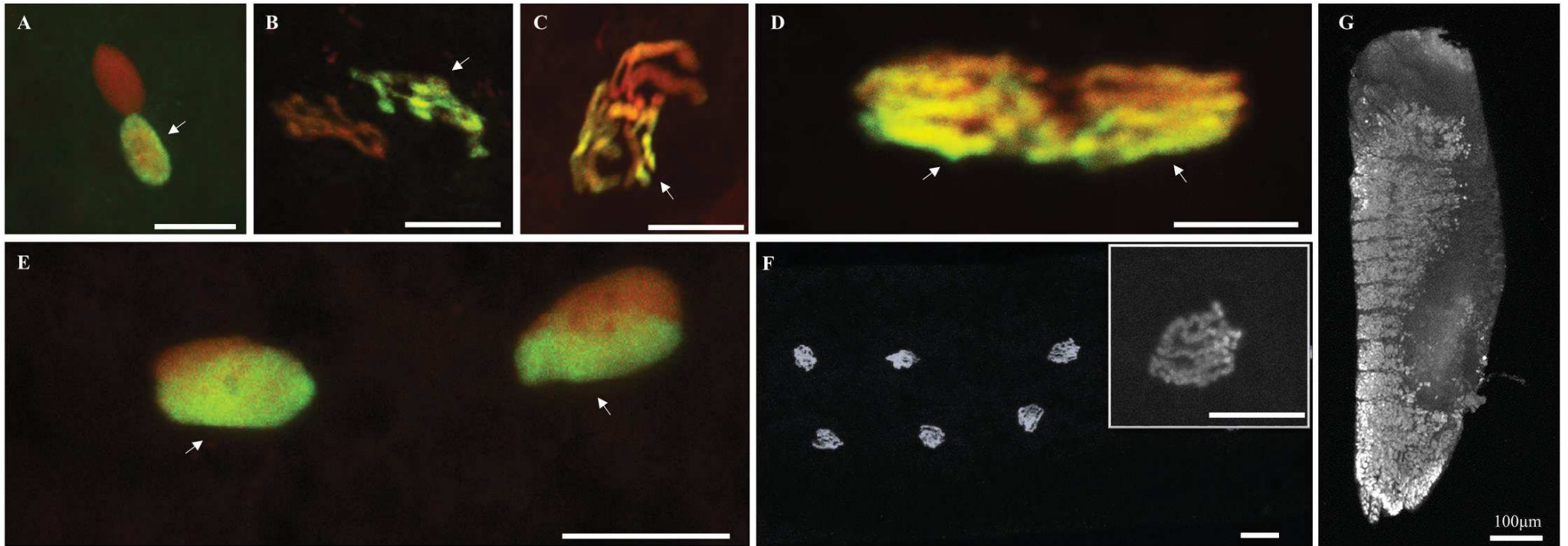


Fig 1. Normal embryogenesis in *C. pipiens*: from first nuclear divisions to segmentation.

Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant). (A) apposition of maternal and paternal pronuclei, (B) chromatin under condensation, (C) condensed chromatin, (D) first mitotic division anaphase (maternal and paternal chromosome segregate independently), (E) two nuclei following the first division, (F) normal diploid development 1 hour after oviposition, six diploid nuclei are visible after 1h development, (G) segmented embryo after 24 hours of development. White arrows indicate the paternal chromatin. Confocal stacks were obtained on embryos from several fertile intra-line crosses due to the difficulty to obtain all the early embryonic stages from each cross (S2 Table). Green dots are background noises likely due to the presence of residual antibodies.

Scale bar is 10μm.

specific MLST with more polymorphic genes MutL, ank2, pk1, pk2, GP12, GP15, and RepA was used to resolve *wPip* phylogeny (S6 Table) [12]. Mosquitoes hosting *wPip* from the same group are likely to be compatible with each other but incompatible with mosquitoes infected with *Wolbachia* from other *wPip* groups [42]. This diversity of *wPip* strains distributed all around the world is responsible for the unique complexity of CI crossing types described in this host species [12,43]. Unlike *wVitA* and *wMel*, which harbors only one *cidA/cidB* copy, and *wRi*, which harbors two identical copies of *cidA/cidB*, high intra and inter-genomic diversities of *cidA^{wPip}/cidB^{wPip}* genes were uncovered between and within all *wPip* strains studied [36]. This diversity certainly explains the unrivaled diversity of crossing types described in *C. pipiens* [36]. This *cidA^{wPip}/cidB^{wPip}* genes amplification and diversification within the same *Wolbachia* genome may also account for the impressive CI penetrance described in *C. pipiens*. Indeed, expression of multiple *cidA^{wPip}* and *cidB^{wPip}* variants in males could i) be responsible for differences in CI cellular phenotype(s) and ii) influence the penetrance of CI. Here, we investigated the putative impact of crossing type variations and *cidA^{wPip}/cidB^{wPip}* diversification on CI cellular phenotypes and CI intensity during *C. pipiens* embryogenesis. To this end, we monitored the development of embryos derived from various incompatible crosses involving males from *C. pipiens* lines infected with *Wolbachia* strains from distinct *wPip* groups and exhibiting different crossing types.

Results

A single cellular phenotype of CI in *C. pipiens*

Three different types of crosses were performed using different laboratory mosquito lines: i) fertile crosses between individuals from the same line, representing our control to monitor normal embryonic early development, ii) sterile crosses between mosquito lines harboring different *wPip* strains, and iii) sterile crosses between infected males and uninfected females (TC lines), to test the effect of *Wolbachia* absence on embryo development and CI cellular mechanism (S1 and S2 Tables).

The cellular phenotype during embryogenesis in fertile intra-line crosses is illustrated in Fig 1. To differentially visualize the paternal from the maternal chromatin, we used propidium iodide to mark both

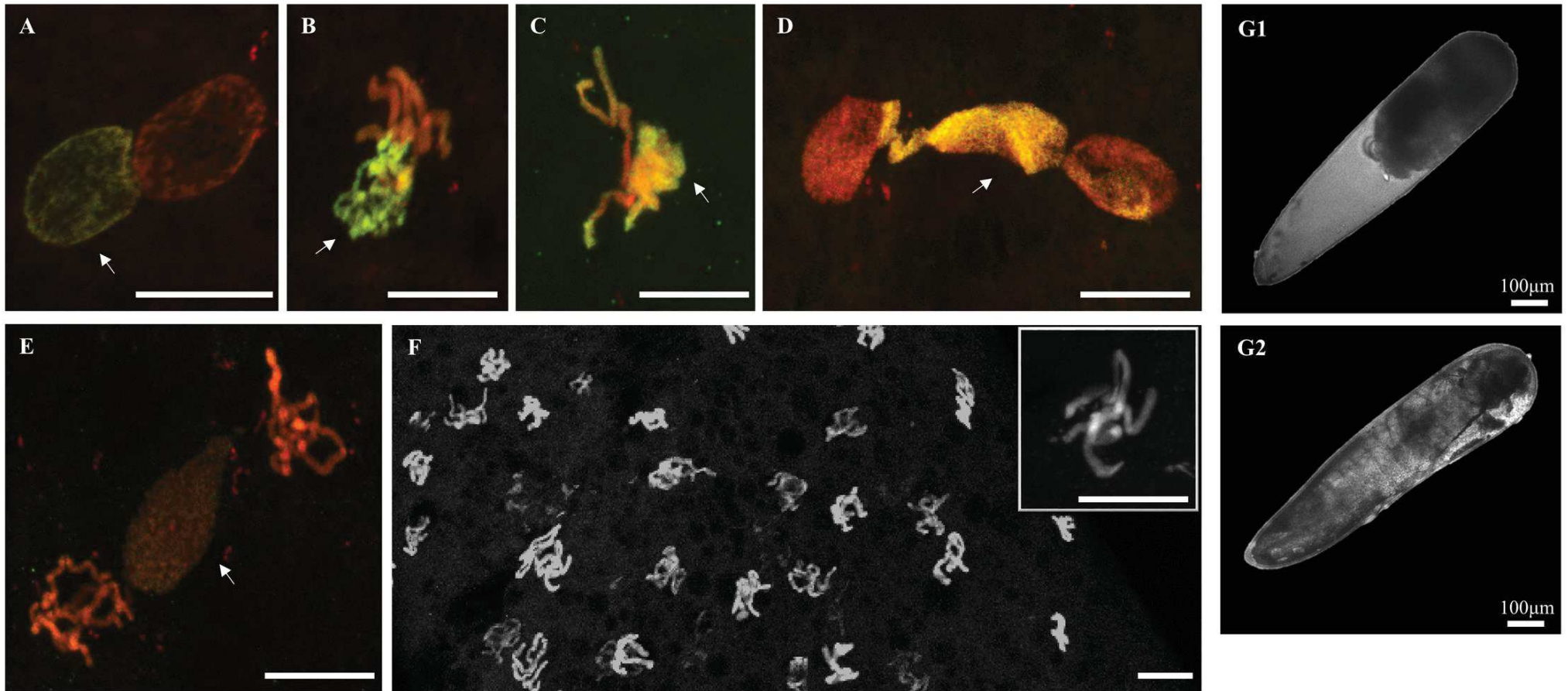


Fig 2. CI embryos from incompatible infected parents in *C. pipiens*: from first nuclear divisions to segmentation.

Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant) (A) apposition of maternal and paternal pronuclei, (B) delay in paternal chromatin condensation, (C) condensed maternal chromatin and under-condensed paternal chromatin, (D) paternal chromatin failed to segregate and form a chromatin bridge between segregating maternal chromatin, (E) two nuclei containing mainly maternal chromatin while paternal chromatin do not segregate, (F) haploid development 2 hours after oviposition, (G) the two possible fates of development after 48 hours (1) non-viable embryo with no visible development, and (2) unhatched developed embryo with visible segments. White arrows indicate the paternal chromatin. Confocal stacks (panels A,B,C,D,E,F) and optical images (panels G1 and G2) were obtained on embryos from several CI crosses between infected male and female due to the difficulty to obtain all the early embryonic stages for each cross (S2 Table). Green dots are background noises likely due to the presence of residual antibodies.

Scale bar is 10 μ m.

maternal and paternal chromatin and an anti-acetylated histone H4 labelling that preferentially marks the *de novo* assembled paternal chromatin after protamine removal [22]. Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant). After fertilization, maternal and paternal pronuclei migrated toward each other and apposed (documented embryos with confocal microscopy images $n=4$, Fig 1A). Then, paternal and maternal chromatins condensed and entered into first mitotic division ($n=3$, Fig 1B). During the first division, paternal and maternal chromosomes aligned in separate region at the **metaphase plate ($n=1$, Fig 1C). Both sets of chromosomes segregated equally during anaphase ($n=3$, Fig 1D) to produce two diploid nuclei ($n=1$, Fig 1E) that proliferate mitotically ($n=16$, Fig 1F). After 24 hours of development, organogenesis was ongoing and segmentation was clearly visible ($n=2$, Fig 1G).**

In sterile crosses between two infected incompatible *C. pipiens* lines (Fig 2) as well as in crosses between infected males and uninfected females (Fig 3), paternal and maternal pronuclei migrated and apposed normally ($n=2$, Figs 2A and 3A). However, during the early prophase, paternal chromatin appeared under-condensed compared to maternal chromatin ($n=2$, Figs 2B and 2C). Then the paternal chromatin failed to segregate properly during anaphase ($n=16$, Figs 2D, 2E, 3B and 3C). In telophase paternal chromatin can either i) formed chromatin bridges between the two maternal nuclei ($n=10$, Figs 2D and 3B), certainly causing the early arrest of embryogenesis and production of undeveloped embryos (Figs 2G1 and 3E), or ii) appeared fully excluded ($n=6$, Figs 2E and 3C), allowing maternal chromatin to successfully segregate and eventually formed unhatched haploid developed embryos presenting eyes and segments (Fig 2G2).

Eight distinct CI crosses were done between males and females infected with *wPip* strains belonging to different *wPip* groups, and two distinct CI crosses were done between infected males and uninfected females (S2 Table). However, despite this diversity of CI crosses, condensation and segregation defects of the paternal chromatin were the only observed cellular defects resulting in embryonic death, and were never observed in any embryo resulting from fertile crosses (seven intra-line fertile crosses, S2 Table).

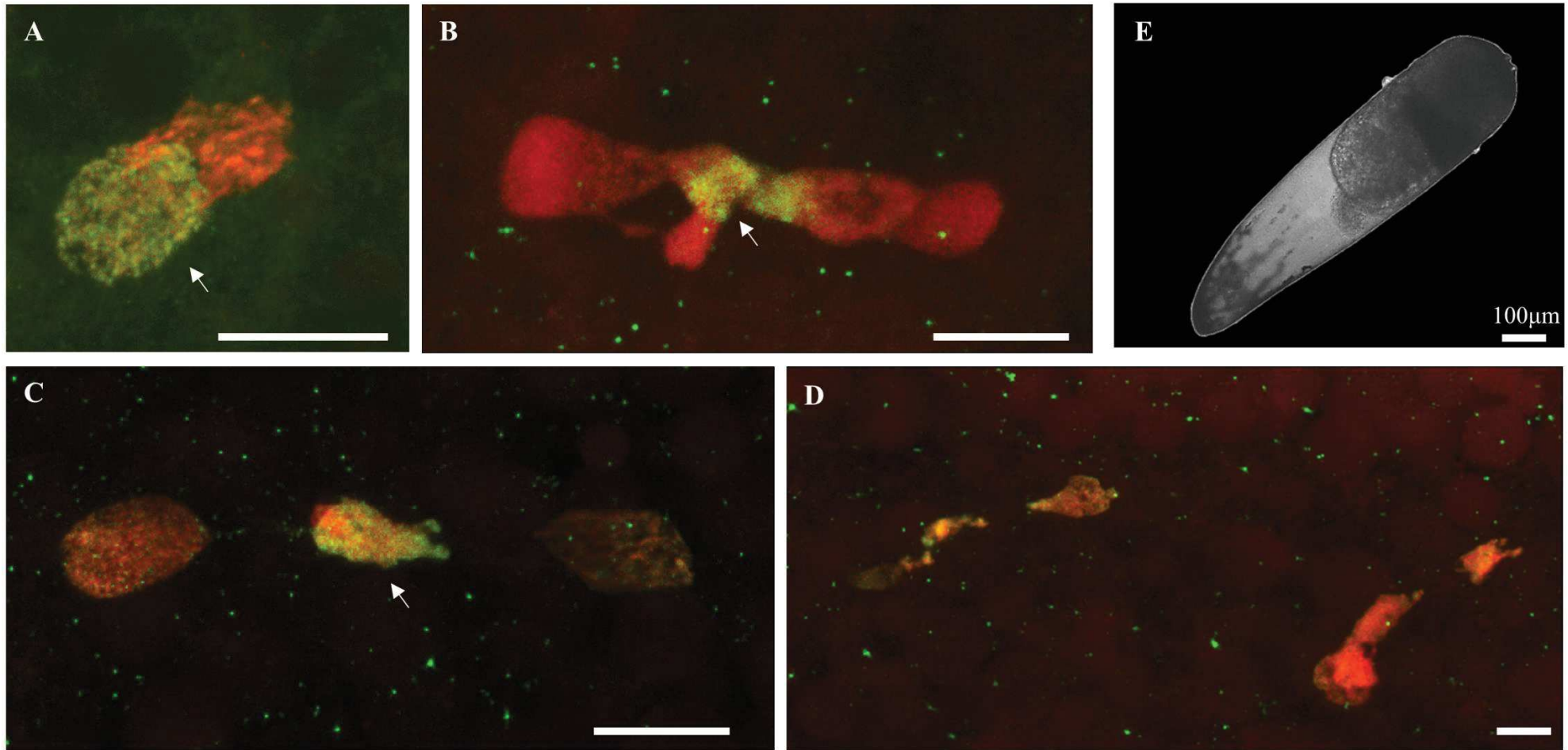


Fig 3. CI embryos from infected fathers and uninfected mothers: an arrest after the first nuclear divisions.

Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant) (A) apposition of maternal and paternal pronuclei, (B) paternal chromatin failed to segregate and form a chromatin bridge between segregated maternal chromatin, (C) two nuclei containing mainly maternal chromatin while paternal chromatin do not segregate, (D) abnormal development at 2 hours after oviposition: only few aborted divisions are observed, (E) none of the embryos from such crosses exhibited any visible development under microscope 48 hours post-oviposition. White arrows indicate the paternal chromatin. Confocal stacks (panels A,B,C,D,) and optical images (panels E) were obtained on embryos from several CI crosses between infected male and uninfected female due to the difficulty to obtain all the early embryonic stages for each cross (S2 Table). Green dots are background noises likely due to the presence of residual antibodies.

Scale bar is 10 μ m.

Absence of *Wolbachia* in the oocytes blocks embryogenesis in CI embryos

2 hours after oviposition, some embryos resulting from CI crosses between infected lines pursued their embryogenesis (n=2, Fig 2F), and after 48 hours these embryos exhibited visible development, as segmentation was clearly observable under optical microscope (Fig 2G2). However, more than 99.9% of these developed embryos did not hatch. In the sterile cross between ♂ Slab x ♀ Ichkeul 13, unhatched developed embryos only displayed maternal markers (see Material and methods, "Ploidy determination in CI developed embryos", S1 Fig), showing that they were composed of only haploid maternal DNA, as previously described in Duron and Weill (2006)[44].

All the seven different crosses performed between males infected with different *wPip* strains and uninfected females from different TC-treated lines produced 100% of non-developed embryos (Fig 3E and S3 Table). Confocal observations of these embryos showed that only few and abnormal nuclei were observed in the cytoplasm 2 hours post oviposition (n=5, Fig 3D), indicating an early arrest of the embryogenesis (S2 Table).

Males infected by different *wPip* with distinct *mod* profiles induced different CI defects intensities

It has been previously proposed that the production of haploid or aneuploid embryos in CI crosses represented a proxy of intensity of CI defects that leads to more or less complete paternal chromatin exclusion [29–31,44]. Severe defects would lead to the complete exclusion of the paternal chromatin during the first embryonic division (*i.e.* strong cellular CI intensity), which would allow maternal chromatin successful segregation and the production of a developed haploid embryo. Thus, unhatched developed haploid embryos reflect the occurrence during the first zygotic division of strong CI defects while unhatched non-developed embryos illustrate the occurrence of weak CI defects. We used this link between the degree of paternal chromatin exclusion (*i.e.* weak or strong cellular CI) and the proportion of unhatched developed embryos in eggs-rafts from incompatible crosses to investigate the variability of CI intensity (*i.e.* frequency of strong *versus* weak CI defects).

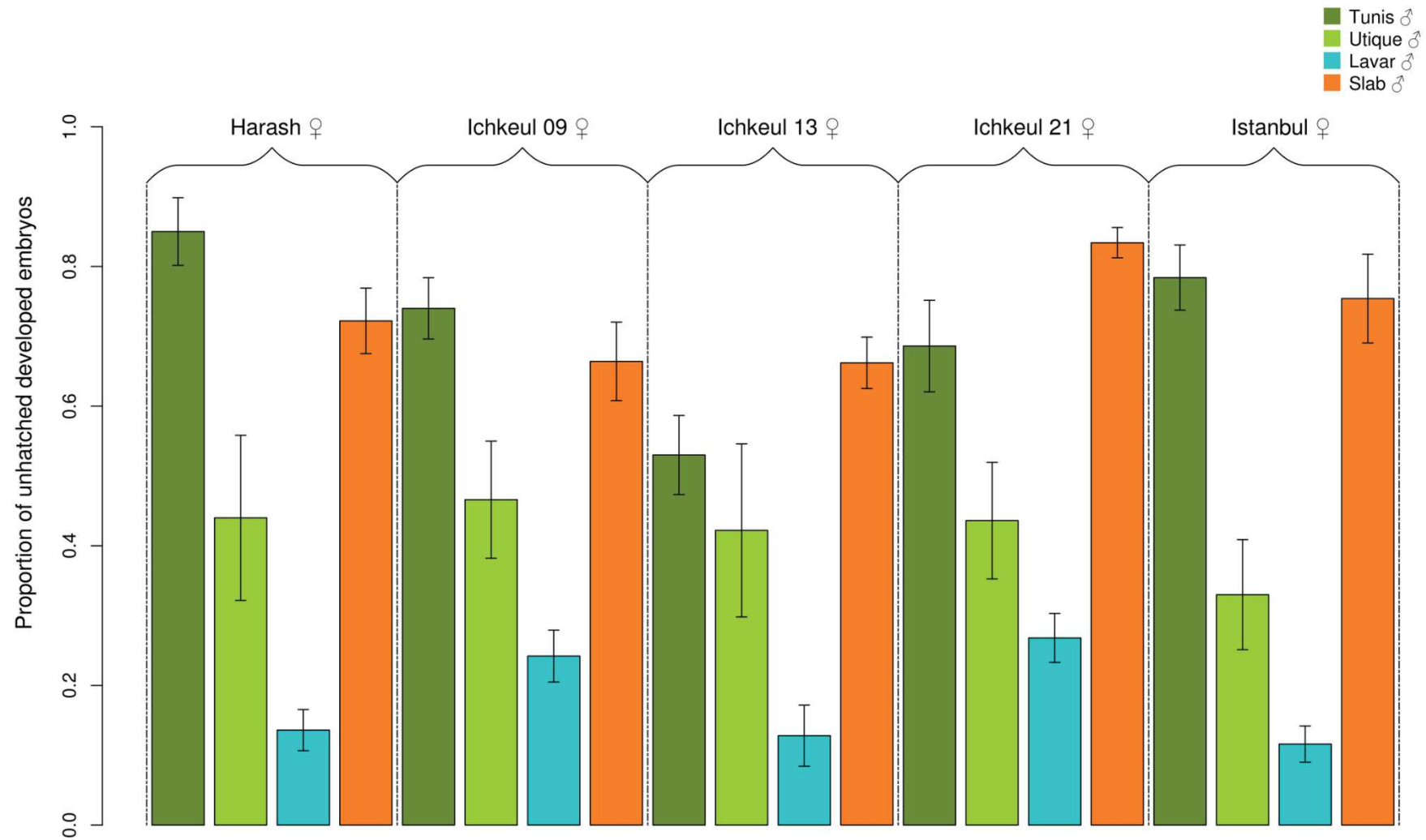


Fig 4. Variations in cellular CI intensity.

Infected males from four lines exhibiting different *mod* profiles (MAL lines) and females from five lines with the same *resc* profile (FEM lines) were crossed and resulted in the 20 studied CI crosses (*i.e.* >99,9% of the embryos died before hatching). The vertical bars represent the proportion of unhatched developed embryos in each of these crosses, the four different colors represent the four different MAL lines, each group of four bars corresponding to the same FEM line. Error bars represent the standard error. The proportion of unhatched developed embryos was significantly different between crosses involving males infected with distinct *wPip* responsible for different *mod*; no significant effect of the FEM lines was detected.

Using this proxy, we studied the variation in CI intensity between 20 incompatible crosses between infected lines (S3 Table). These 20 crosses involved i) males from four different isofemale lines (MAL lines) infected with *wPip* strains from different *wPip* groups all exhibiting distinct *mod* profiles, and ii) females from five isofemale lines (FEM lines) all harbouring *wPip* strains from the *wPipIV* group and exhibiting the same *resc* profile [36,42] (S4 Table). Significant differences were found regarding the proportion of unhatched developed embryos between these incompatible crosses (generalized linear model (GLM), $\chi^2=245.695$, $df=19$, $p<0.001$, Fig 4, Table 1 and S3 Table). While no effect of FEM lines was detected on this proportion (GLMM, $\chi^2=2.508$, $df=4$, $p=0.643$, Fig 4), the MAL lines involved in the crosses had a significant effect (GLMM, $\chi^2=16.211$, $df=3$, $p=0.001$, Table 1 and Fig 4). Males from Tunis (*wPipI mod ii*) and Slab (*wPipIII mod iii*) lines induced the highest proportion of developed embryos (72% and 73%, respectively) but were not significantly different from one another (GLMM, $\chi^2=0.002$, $df=1$, $p=0.968$); males from Utique (*wPipI mod iv*) and Lavar (*wPipII mod vi*) lines induced significantly different and lower proportions of unhatched developed embryos (respectively 42% and 18%, Table 1). The nuclear genetic background of the males seems not to be involved in the variability of CI defects intensity: males from backcrossed line SI(*wPipI-Tunis*) and males from the Tunis line, which host the same *wPipI* strain in different genetic backgrounds, indeed induced similar unhatched developed embryos proportions when crossed with the five FEM lines (0.71 ± 0.22 and 0.72 ± 0.19 respectively; GLMM, $\chi^2=0.008$, $df=1$, $p=0.927$). Consequently variability in CI defects intensity appears to be only dictated by the *wPip* strain harbored by the different males.

Genetic investigations of cellular CI intensity variation

The results from the previous section indicate that the proportion of unhatched developed embryos in CI crosses likely depended on variations in the males' *mod* profiles. To investigate the sources of such variation in CI defects, we tested the putative influence of several variables: i) the density of *Wolbachia* in the testes, ii) the copy numbers of *cidA* and *cidB* genes in the different *wPip* genomes, iii) the expression levels of *cidA* and *cidB*, and iv) the *cidA* and *cidB* variants repertoires in the genomes of the different *wPip* strains hosted by the males.

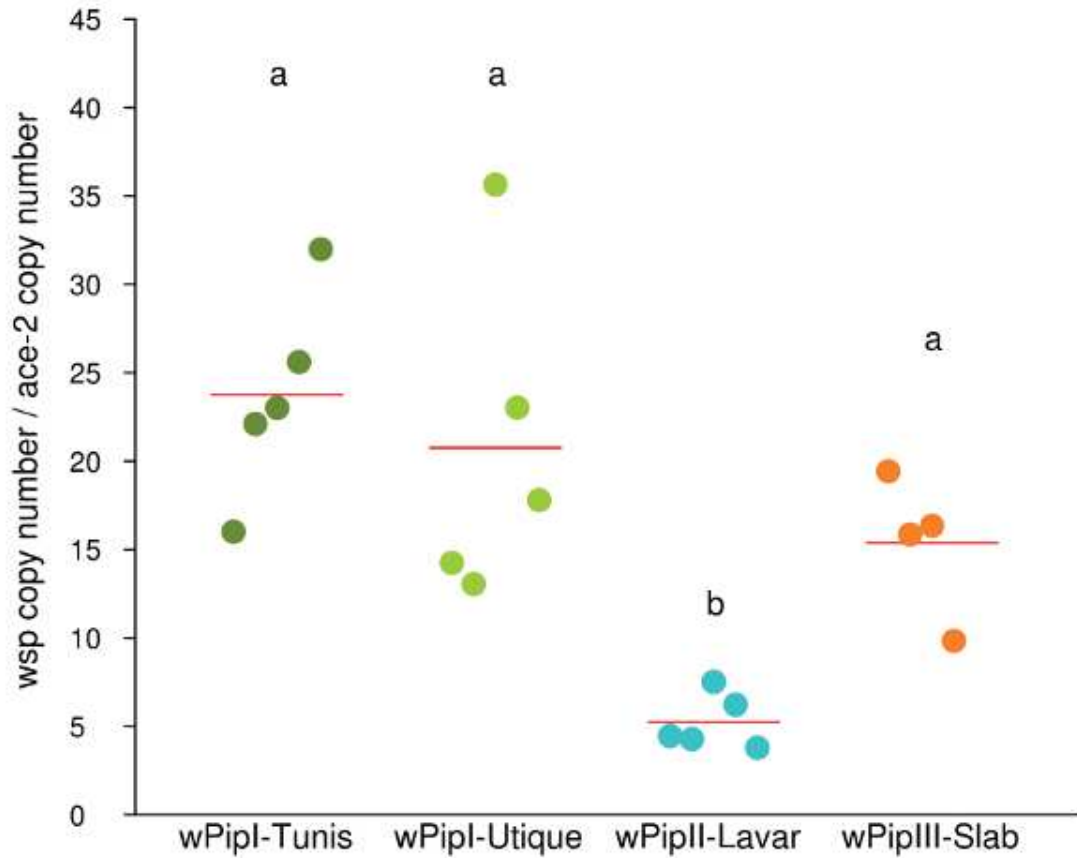


Fig 5. *Wolbachia* density in testes.

Wolbachia densities in mosquito testes were measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia wsp* gene and the *C. pipiens ace-2* nuclear gene. The colored dots represent the average density of *Wolbachia* in a pool of three pairs of testes and the red strips represent the average *Wolbachia* density for each line/*wPip* strain. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different), showing that Lavar males harbored significantly less *Wolbachia* in their testes than the other males.

***Wolbachia* density was significantly lower in Lavar males' testes.** Testicular *Wolbachia* densities were not significantly different between males from Tunis, Utique and Slab lines (GLM, $F=3.919$, $df=1$, $p=0.065$), but significantly lower in Lavar males (GLM, $F=9.337$, $df=3$, $p<0.001$, Fig 5).

***CidA* and *cidB* copy numbers differed between the wPip strains.** Number of genomic copies of *cidA* per wPip strain (*i.e.* *Wolbachia* cells from the same *C. pipiens* line) varied between MAL lines from 4.68 ± 0.42 to 5.98 ± 0.73 copies, wPipI-Tunis displaying a significantly higher *cidA* copy number than the other strains (GLM, $F=8.077$, $df=3$, $p<0.001$, Table 1, S5 Fig). Number of genomic copies of *cidB* also varied from 4.07 ± 0.47 to 5.57 ± 0.56 copies (GLM, $F=9.142$, $df=3$, $p<0.001$), and was found significantly higher in wPipI-Tunis than in wPipI-Utique, wPipIII-Slab and wPipII-Lavar displaying significantly lower but similar copy numbers (Table 1, S6 Fig). Despite these differences, the *cidA/cidB* copy number ratios were not significantly different and close to one, for all wPip strains (GLM, $F=1.504$, $df=3$, $p=0.230$, Table 1, S7 Fig).

Variation in *cidA* and *cidB* relative expression levels. The expression levels of both *cidA* and *cidB* were not different between the four *C. pipiens* MAL lines (GLM, *cidA*: $F=2.409$, $df=3$, $p=0.083$, *cidB*: $F=1.239$, $df=3$, $p=0.310$, Table 1, S8 and S9 Figs). *CidA* was found about 1.5 times more expressed than *cidB*, except for wPipII-Lavar which *cidA/cidB* expression level ratio appeared significantly higher (GLM, $F=5.447$, $df=3$, $p=0.003$, Table 1, S10 Fig).

***CidA* and *cidB* variants repertoires were variable between wPip strains.** Cloning and Sanger sequencing revealed that the four MAL lines with different *mod* were infected with *Wolbachia* strains that harbored different *cidA* and *cidB* variant repertoires (S2 and S3 Figs). No *cidA* or *cidB* nucleotide sequence variant were shared between the three wPip groups. However, the CidA_II($\alpha/1$) variant of wPipII-Lavar and the CidA_III($\beta/8$) variant of wPipIII-Slab presented the same amino-acid sequence (S2 Fig). wPipIII-Slab exhibited ten variants of *cidA* based on their nucleotide sequences (however only

Table 1. Males with different *mod* profiles: proportions of unhatched developed embryos, *cidA-cidB* copy number and expression.

Line	Proportion of unhatched developed embryo	<i>cidA</i> copy number	<i>cidB</i> copy number	<i>cidA/cidB</i> copy number	<i>cidA</i> expression	<i>cidB</i> expression	<i>cidA/cidB</i> expression
wPipI-Tunis	0.72 ± 0.19* (a)	5.98 ± 0.73 (a)	5.57 ± 0.56 (a)	1.08 ± 0.14 (a)	0.75 ± 0.18 (a)	0.50 ± 0.16 (a)	1.55 ± 0.30 (a)
wPipI-Utique	0.42 ± 0.31 (b)	5.02 ± 0.41 (b)	4.76 ± 1.00 (b)	1.08 ± 0.16 (a)	0.76 ± 0.27 (a)	0.49 ± 0.13 (a)	1.53 ± 0.32 (a)
wPipII-Lavar	0.18 ± 0.12 (c)	4.98 ± 0.83 (b)	4.14 ± 0.76 (c)	1.22 ± 0.22 (a)	0.97 ± 0.24 (a)	0.47 ± 0.15 (a)	2.14 ± 0.58 (b)
wPipIII-Slab	0.73 ± 0.16 (a)	4.69 ± 0.42 (b)	4.07 ± 0.47 (c)	1.17 ± 0.18 (a)	0.94 ± 0.26 (a)	0.60 ± 0.20 (a)	1.62 ± 0.32 (a)

*The average proportion of unhatched developed embryos, the number of copies of *cidA* and *cidB* and their ratio, as well as the expression levels of *cidA* and *cidB* and their ratio are indicated as the MAL line means ± standard deviations.

a, b, c letters represent statistical groups (*i.e.* means with the same letter are not significantly different).

seven of them differed in their amino-acid sequences), wPipII-Lavar three, wPipI-Tunis four and wPipI-Utique seven (S2 Fig). Both wPipII-Lavar and wPipI-Utique carried two different variants of *cidB*, while wPipI-Tunis and wPipIII-Slab carried four (S3 Fig).

Correlations between CI defects intensity and wPip genetic variations

Correlative analyses were conducted to assess the potential links between variations in CI defects intensity and genetic variations. We found no significant correlations between the proportion of unhatched developed embryos in CI crosses and i) *Wolbachia* density in the testes (Spearman, $\rho=0.4$, $p=0.750$), ii) *cidA* copy number (Spearman, $\rho=-0.2$, $p=0.917$), iii) *cidB* copy number (Spearman, $\rho=-0.2$, $p=0.917$), iv) *cidA/cidB* copy number ratio (Spearman, $\rho=-0.4$, $p=0.750$), v) *cidA* expression levels (Spearman, $\rho=-0.4$, $p=0.750$), vi) *cidB* expression levels (Spearman, $\rho=1$, $p=0.083$), vii) *cidA* over *cidB* expression levels (Spearman, $\rho=-0.2$, $p=0.917$) and viii) the number of different *cidA* variants in the repertoires (Spearman, $\rho=0.8$, $p=0.333$). However, males infected with wPip strains with 4 *cidB* variants induced significantly higher proportions of unhatched developed embryos (wPipI-Tunis and wPipIII-Slab mean : 0.72 ± 0.17) than males infected with wPip strains with only 2 *cidB* variants (wPipII-Lavar and wPipI-Utique; mean : 0.30 ± 0.26 , Wilcoxon, $W = 1159$, $p<0.001$, S4 Fig).

Discussion

To investigate whether the high diversity of *cidA/cidB* variants within wPip could be responsible for variations in the cellular phenotype of CI, we studied the development of *C. pipiens* embryos resulting from various incompatible crosses. The early embryogenesis was assessed using fluorescence confocal microscopy in i) fertile intra-line crosses, ii) incompatible crosses between infected males and infected females, and iii) incompatible crosses between infected males and uninfected females. Despite the diversity of performed crosses between males and females infected with wPip strains harboring different *cidA/cidB* variants repertoires or uninfected female, a unique and recurrent embryonic phenotype was detected, consisting in paternal chromatin condensation and segregation defects during

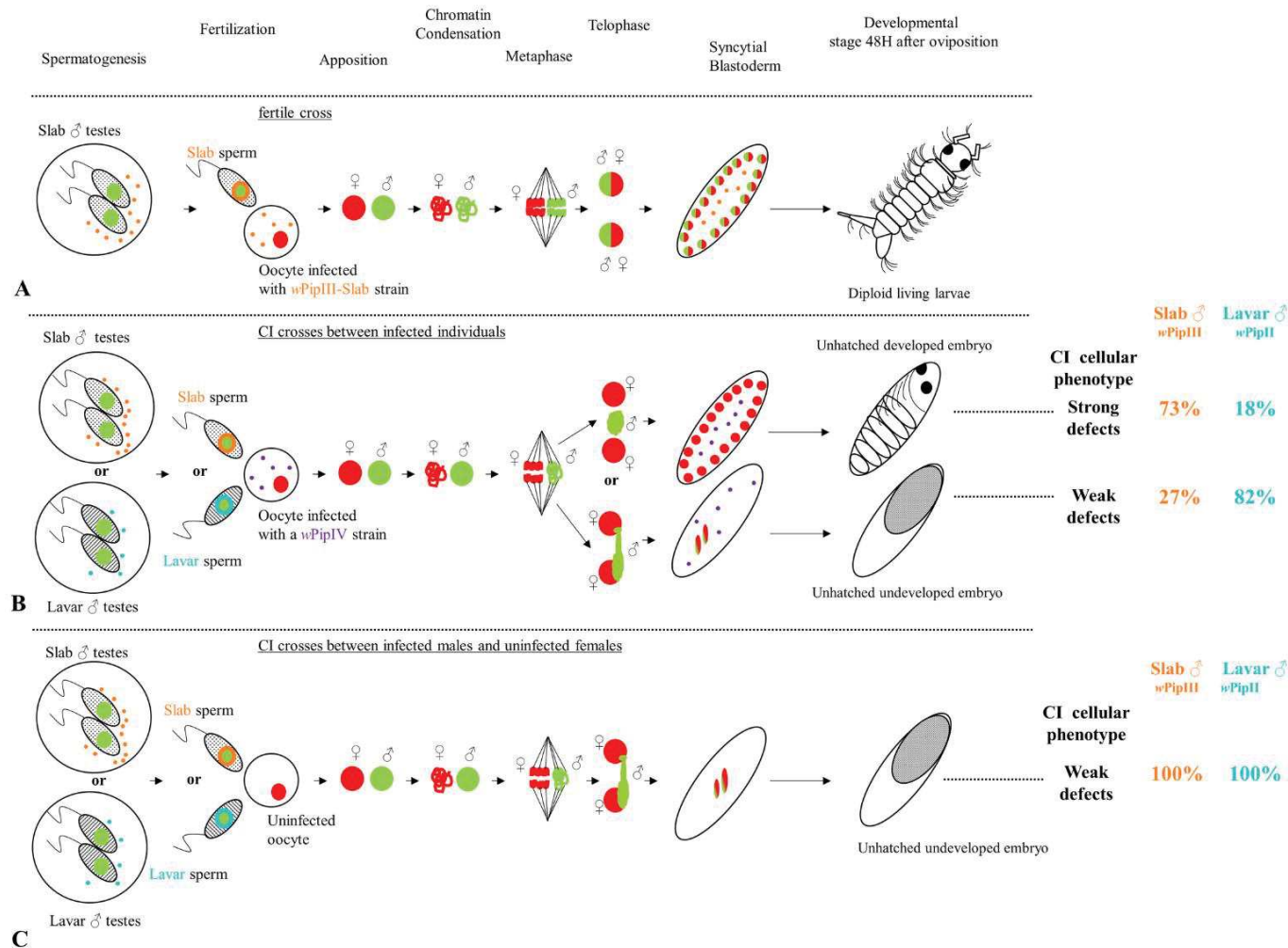


Fig 6. Cellular bases of different developmental fates in *C. pipiens*: from spermatogenesis to late development stages.

Paternal and maternal DNA are represented in green and red respectively and *Wolbachia* cells are represented by the small dots (blue for wPipII; orange for wPipIII and purple for wPipIV). The two MAL lines Slab and Lavar have been chosen to illustrate the variability in CI defects intensity (*i.e.* frequency of strong and weak cellular CI). During spermatogenesis, the *Wolbachia* in male testes produce a toxin (*mod* factor). After fertilization, if females and males are infected with the same *Wolbachia* (Panel A), they can prevent the toxicity (*resc* function) and allow normal embryogenesis and the production of living diploid larvae. However, if the *Wolbachia* in the females are incompatible with the *Wolbachia* in the males (Panel B) or absent (Panel C) then the toxicity would not be prevented and paternal chromatin condensation delay and segregation defects occur. Two outcomes of the first zygotic division are possible regarding the degree of paternal chromatin exclusion. If the paternal chromatin is partially excluded, chromatin bridges would be formed resulting in aneuploid nuclei which might go through few mitotic divisions but will arrest the embryogenesis at early stages producing embryos with no visible development. If the paternal chromatin is fully excluded from the first zygotic division, maternal chromatin can segregate and produce two haploid nuclei which will be divided further to produce non-viable haploid embryos exhibiting advanced development stages with eyes and segmentation clearly visible. The proportion of unhatched haploid embryos is influenced by different wPip with distinct *mod* profiles in the MAL lines involved in the crosses. However, when the egg is not infected by *Wolbachia* (Panel C), all embryos exhibited an absence of further development after the first zygotic divisions whatever the *mod* induced by the wPip hosted by the MAL lines.

the first embryonic division (Figs 2B-E, 3B and 3C). This phenotype was never detected in any embryos derived from intra-line crosses (Fig 1). Hence the diversity of *cidA/cidB* variants repertoires describes in *C. pipiens* does not seem to influence the CI mechanism itself, which is consistent with all CidB variants carrying a conserved DUB domain [36]. Similar defects were already reported in both *Drosophila* and *Nasonia* [23,25], suggesting an universality of *Cid* induced-cellular CI mechanism whenever *cid* genes are diversified or not in *Wolbachia* genome.

An unsolved question is the molecular pathway(s) targeted by CidA and CidB. Most protein domains within CidA and CidB remain to be characterized and how they interact with each other and host targets to induce CI remains unclear. However, a first tangible element is that the catalytically active DUB domain (involved in deubiquitination) in CidB proteins, which is considered as involved in the *mod* function, is necessary to induce CI in transgenic *Drosophila* [34]. Ubiquitination pathways have been shown to be crucial for many essential cellular processes, such as the regulation of the chromatin dynamics and the cell cycle progression [45]. Changes in ubiquitination could for instance directly or indirectly affect H3.3 histone incorporation after protamine removal and PCNA persistence on the paternal chromatin due to DNA replication defects [22], which would result in an asynchronous mitotic entry of paternal and maternal pronuclear chromatin [30]. Interestingly, *Cardinium*, an endosymbiont phylogenetically distant from *Wolbachia*, induces CI with quite similar embryogenesis defects in the hymenoptera *Encarsia suzanna* [46]. Moreover, an ubiquitin specific protease USP classified as a DUB protein has also been detected in *Cardinium* genome, suggesting a convergent implication of DUB in CI induced by insect endosymbionts [47]. However, some *Wolbachia* strains able to induce CI do not carry DUB domain (*i.e.* no *Cid*) in their genomes, but display instead a paralog gene with a nuclease domain called *cinB* [34,35,38]. DUB (Cid) and Nuclease (Cin) domains do not have the same predicted functions suggesting that distinct molecular pathways may be responsible for CI [34,38]. The CI cellular defects caused by *Wolbachia* strains harboring only *cin* genes remain unknown and could differ from the one induced by *cid* genes. Our study showed that wPip strains, which carry both *cid* and *cin* genes in their genomes, induce similar defects during embryogenesis as wMel, which carries only a *cid* gene. This suggests that the association of *cid* and *cin* does not change the cellular phenotype of CI, but the

molecular mechanism induces by DUB and Nuclease which must be different due to the biochemical nature of the proteins might converge on a similar cellular defect (*i.e.* paternal chromatin condensation defect). However, the presence of DUB and Nuclease domains in the same *Wolbachia* genome could still contribute to CI by modifying its penetrance: *w*Ri (*D. simulans*) and *w*Pip (*C. pipiens*), which harbor both paralogs, have indeed a strong CI penetrance (almost no hatched embryos), while *w*No and *w*Ha (*D. simulans*), which carry either *cin* or *cid* genes, respectively, induce lower CI penetrance [13,38].

Our cytological investigation in *C. pipiens* evidenced a link between the paternal chromatin exclusion degree during the first zygotic division and the existence of two developmental fates following first-division defects. In fact, unhatched embryos can either reach advanced developmental stages, exhibiting segments and visible eyes, or display no visible development (Figs 2G and 6) [26,44]. We confirmed Duron and Weill (2006)[44] findings that the unhatched developed embryos resulting from CI were haploid, and carried genetic material from maternal origin only (Fig 6 and S1 Fig). Confocal observations showed that such haploid development likely occurred when paternal chromatin was fully excluded during the first zygotic division, allowing the successful segregation of the isolated maternal chromatin (Fig 6). In contrast, unhatched non-developed embryos would be due to partial exclusion of the paternal chromatin, which would result in aneuploid nuclei and early arrest of embryogenesis (Fig 6). It has been previously proposed for other arthropod models that the participation of paternal chromatin to the first division would depend on the intensity of paternal chromatin defects (*i.e.* improper condensation) [29,30,44,48]. Severe defects would lead to complete paternal exclusion (*i.e.* strong cellular CI) and to the production of haploid developed embryos, while less severe defects would lead to a partial paternal chromatin exclusion (*i.e.* weak cellular CI) and to the production of aneuploid non-developed embryos. We used this link between the degree of paternal chromatin exclusion and the ratio of unhatched developed and non-developed embryos in eggs-rafts from incompatible crosses to investigate the variability of cellular CI intensity between different incompatible crosses.

We first studied the variability of CI intensity using males and females both infected with incompatible *w*Pip strains. Developed embryos were observed in all these incompatible crosses, with two possible outcomes: i) less than one per thousand of these embryos were apparently not affected by

CI and hatched into diploid larvae [14,44], and ii) from 11% to 85% of the unhatched embryos, depending on the crosses, reached late embryonic developmental stages showing that they experienced strong CI defects (Figs 4 and 6). We then studied the influence of the absence of *Wolbachia* in the oocytes on the cellular CI intensity. As in Duron and Weill (2006)[44], i) we confirmed that not a single larvae was produced in such crosses, and ii) all the seven CI crosses between infected males and uninfected females (TC lines) resulted in 100% of non-developed embryos suggesting that in such crosses, CI phenotype was always weak (Fig 6).

In crosses between infected individuals, it clearly appeared that MAL lines harboring *Wolbachia* from different *wPip* groups (*wPip* I, II, III) and displaying distinct *mod* induced significant variation in CI defects intensity when crossed with females harboring distinct *wPip*IV strains displaying the same *resc* (Figs 4 and 6). Variation in CI defects intensity has already been reported in *Nasonia* species, where the production of haploid viable males in *N. vitripennis* was interpreted as resulting from severe paternal chromatin defects, while the production of unviable aneuploid embryos in *N. longicornis* and *N. giraulti* was interpreted as resulting from weak paternal chromatin defects. However, variation of CI intensity in these host species was not associated with the different *Wolbachia* strains, but to variation in host genetic backgrounds [19]. The backcross experiment performed in the present study suggests that CI intensity is not impacted by nuclear genetic variations in *C. pipiens*. While it was already established that *Wolbachia* drives alone the observed variation in crossing types in *C. pipiens* [11,49,50], *Wolbachia* also seems to dictate the intensity of CI defects. Consequently, the variation in CI intensity observed when two infected individuals are crossed seems to be under the major influence of the *wPip* strain infecting the MAL line *via* the degree of paternal chromatin exclusion they trigger.

In *C. pipiens*, when tetracycline-cleared females (TC females) were crossed with the four MAL lines, 100% of unhatched non-developed embryos only exhibiting few degenerated nuclei were observed, even 2 hours after oviposition (Figs 3D and 6). Such CI phenotype suggests that the defects caused by the *wPip* infecting all the MAL lines are always weak (Fig 6). This result is counter intuitive because one would expect that when *Wolbachia* is absent from the eggs CI should be always strong and many haploid embryos should be produced. We mentioned above that all the MAL lines can induce

strong CI defects in variable proportion of the embryos when crossed with infected females. Consequently, the constant weak CI phenotype observed when females are not infected is linked to the absence of *Wolbachia* during egg maturation. Our results suggest that in incompatible crosses between infected *C. pipiens* individuals, the presence of maternal *Wolbachia* somehow interferes with early embryogenesis allowing haploid development to occur. It seems very unlikely that the presence of incompatible *Wolbachia* in the egg would enhance the mechanisms leading ultimately to paternal chromosome condensation defects (*i.e.* accentuate the *mod* function) to result in its total exclusion during the first embryonic division. Instead, the presence of incompatible *Wolbachia* in the eggs may have an additive effect on the incompatibility between pronuclei, not by directly affecting the paternal chromatin but by influencing the cell cycle timing. For instance, maternal *Wolbachia* could modulate the maternal kinetics for DNA replication or the mitotic entry during early development, increasing the incompatibility between pronuclei and therefore favoring the haploid development. Thus, while paternal *Wolbachia*-induced CI defects always occur regardless of the infection status of the eggs, the absence of incompatible maternal *Wolbachia* would block haploid development resulting in weak CI phenotype.

We then investigated the putative genetic determinism of CI intensity variation in embryos derived from infected parents. We assessed whether it could result from difference in *Wolbachia* density, *cidA-cidB* gene expression, copy numbers, or variant diversity between the wPip strains. As previously described in *Drosophila* [35,38], we found in *C. pipiens* that *cidA* was always significantly more expressed than *cidB*, whatever the wPip strain (Table 1). This is in accordance with the hypothesis that *cidA* and *cidB* form a toxin-antidote system where CidA is the antidote of CidB [34,36]. Indeed, in such system the antidote was always found more expressed than the toxin to prevent the host from toxicity [51]. No significant difference between MAL lines was found for *cidA* and *cidB* expression levels per *Wolbachia* cell (Table 1), suggesting that the *cidA* and *cidB* expression does not influence CI defects intensity. However, while the *cidA* and *cidB* expression levels per *Wolbachia* cell did not significantly vary between *C. pipiens* lines, the total amount of CidA and CidB proteins in the host mainly depends on the density of *Wolbachia*. Since the *mod* factors are most likely deposited on the sperm in the testes during spermatogenesis [32,33], we measured the density of *Wolbachia* in the male gonads. We found

that Lavar males hosted significantly less *Wolbachia* in their testes than males from the three other lines (Fig 5); Lavar males were also those that generated the lowest proportion of unhatched developed embryos in their offspring, whatever the FEM lines (Fig 4). Due to lower *Wolbachia* density in the testes, the global amount of CidB protein could be lower in Lavar line compared to the other lines. This low dosage of CidB would more likely result in weak CI defects leading to only few haploid development. However, this hypothesis relies on a single line and requires more *C. pipiens* lines with distinct testicular *Wolbachia* densities to be confirmed. Lavar was also the line with the highest *cidA* expression relatively to *cidB* (Table 1); as CidA has been proposed as the CidB antidote [34,36], its overexpression could reduce CidB-induced CI defects, and contribute to the low frequency of developed haploid embryos observed in crosses involving males from Lavar line.

We previously demonstrated that the amplification followed by the diversification of *cidA* and *cidB* variants in *wPip* certainly constitutes the source for CI diversity profiles in *C. pipiens* while *cinA* and *cinB* did not exhibit any polymorphism [36]. Indeed, specific variations in *cidA* and *cidB* repertoires (number and/or nature of the variants) clearly seemed to determine the compatibility outcome of crossings between *wPipIV*-infected males and any infected females, pointing out the putative role of these variations in the prodigious CI complexity recorded in this species [36,42]. Here, we tested the putative consequence of *cidA* and *cidB* gene amplification (*i.e.* number of copies per genome) on variation of CI defects intensity, and demonstrated no significant correlation between the two parameters. When the quantification of genomic copies obtained by q-PCR are put in relation to the number of different variants in the same isofemale line obtained by cloning-sequencing, some of *cidA* results might appear discordant. This is especially true for the Slab line, which exhibits ten distinct *cidA* variants for ~5 copies per genomes quantified (Table 1 and S2 Fig). Even taking into account technical limits of q-PCR to quantify high level of gene amplification, this discordance suggests that, at least in the Slab line, some of the *Wolbachia* cells do not harbor the same variants.

We found that the different *wPip* strains carried by the four MAL lines exhibiting different *mod* profiles harbored distinct *cidB* variants. Any variant of this gene could certainly trigger CI alone, as the DUB domain is perfectly conserved between all variants [36]. However, their diversity can modulate CI

defects intensity. We thus tested whether *cidB* repertoire diversity could play a role in CI intensity variability. Supporting this hypothesis, we found that males from the two *C. pipiens* lines harboring wPip strains with four different *cidB* variants induced higher proportions of unhatched developed embryos compared to lines harboring wPip with only two different *cidB* variants (S4 Fig). Each distinct *cidB* variants could differentially impact the paternal chromatin (*i.e.* like different locks), putatively leading to an additive *mod* effect: the more different *cidB* variants present in a wPip strain, the more likely strong CI defects. However, more wPip strains varying in their diversity of *cidB* are required to further test this hypothesis.

In conclusion, despite the diversity of crossing types observed in *C. pipiens*, linked to the diversity of *cidA/cidB* variants repertoires, a single cellular phenotype of CI, was observed in this species. In all crosses (*i.e.* uni-bidirectional), CI results in early developmental defects in the paternal chromatin condensation and segregation during the first zygotic division similar to that observed in other insects. Our study demonstrates that in CI crosses between two infected individuals, the CI intensity (*i.e.* frequency of strong and weak CI defects) is influenced by the male-carried wPip. However, when the female is not infected, and despite the variability of the distinct wPip strains carried by the males, no unhatched developed embryos (strong cellular CI) were ever found, suggesting that the weak CI phenotype observed in such crosses is instead due to the absence of *Wolbachia* in the eggs. Genetic investigation reveals that the variability of CI defects intensity may be linked to *cidB* variant diversity in wPip strains. While the functional role and the singularity of *CidB* amplification and diversification in wPip remains yet to be fully solved, it clearly appears that it deeply modifies the wPip-induced CI phenotype at different scales, from crossing types [36] to its intensity at the cellular level.

Materials and Methods

Culex pipiens lines

***C. pipiens* lines hosting different *Wolbachia* strains or without *Wolbachia*.** Ten laboratory isofemale lines belonging to *C. pipiens s.l.*, from our laboratory, were used; they differ in their

geographical origins and in the *Wolbachia* strain hosted (S1 Table). To determine the *mod* and *resc* profiles of the different mosquito lines, crosses with four reference lines (4-ref cytotypes) were realized similarly to Atyame *et al.* (2014)[42] (S4 Table). To study the effect of the absence of *Wolbachia* on embryogenesis, tetracycline-treated *Wolbachia*-free lines (TC lines) SlabTC, IstanbulTC, Ichkeul 21TC, and Ichkeul 13TC were obtained respectively from Slab, Istanbul, Ichkeul 21 and Ichkeul 13 *wPip* infected lines, as described in Duron *et al.* (2006)[11]. After TC treatment, PCR amplifications of a fragment of *wsp* gene using the primers designed in Berticat *et al.* (2002)[52] allowed controlling the absence of *Wolbachia* in DNA extracted from a larvae pool (Dneasy Blood & Tissue Spin-Column protocol Kit; Qiagen; Bench protocol: Animal Tissues). To prevent possible side-effects of the treatment, TC-treated lines were raised in standard laboratory conditions for at least four generations without tetracycline before the beginning of experiments.

***C. pipiens* lines with the same host genetic background but different *Wolbachia* strains.**

wPipI strain from Tunis line was introduced into Slab line nuclear genetic background through 8 backcrosses. For the first generation, 200 females from Tunis line were crossed with 100 males of the SlabTC line. Then, for each of the seven following generations, 200 females from the previous generation were crossed with 100 males from the SlabTC line. This led to a progressive replacement (over 97%) of the maternal nuclear genome (Tunis) by the paternal nuclear genome (SlabTC), with retention of the maternal cytoplasm, including the Tunis *Wolbachia* strain [this line was called Sl(*wPipI*-Tunis)].

Cellular study of early embryogenesis

To characterize CI cellular phenotype(s) in *C. pipiens*, several crosses were performed (S2 Table). For every crosses, to avoid confounding age effects, two-day old adults were released in cages. Cages containing 100 females and 50 males were then put into a closet at 25°C where day-night cycle was inverted to allow collection of early developmental stage eggs during the day. After six days in these cages, females were fed with turkey blood in heparin sodium (bcl Wholly Wild World) using a Hemotek

membrane feeding system (Discovery Workshops, United Kingdom). Five days after blood meal, water-pots were placed into the cages to collect the eggs-raft. For *C. pipiens* eggs, at 25°C, the meiosis is approximately completed 30 minutes after the oviposition and the first mitotic nucleus division 15 minutes after the end of the meiosis, while four hours after oviposition the embryos normally reach the syncytial blastoderm stage [53]. Since, the CI defects described in *D. simulans* [25] and *N. vitripennis* [23,24] occurred during the first nucleus mitotic division, we mainly collected eggs aged from 30 minutes to 1h. Older eggs were also harvested to monitor further developmental stages in both fertile and sterile crosses. Eggs-rafts were then placed into commercial bleach (active ingredient, 9.6% of sodium hypochlorite) to dissociate eggs, and then washed in distilled water. They were then fixed by being shaken for 2 hours in a solution of 3.2% para-formaldehyde in PBS 1X with Tween 0,02% (PBS-T) and washed with PBS 1X. For each fixed egg, the chorion was removed manually with a needle under an optical microscope (Leica MZ 8). Dechorionated embryos were then collected and treated with RNase A (10 mg/mL, Sigma) overnight.

To differentially visualize the paternal from the maternal chromatin, we used propidium iodide to mark both chromatin and an anti-acetylated histone H4 labelling that preferentially marks the *de novo* assembled paternal chromatin after protamine removal [22]. Thus maternal and paternal chromatin will be respectively predominantly marked with propidium iodide (mostly red fluorescence) and with anti-acetylated histone H4 antibodies (mostly green fluorescence). For immunolabeling, embryos were first incubated overnight at 4°C with primary antibodies (Polyclonal anti-acetylated histone H4 primary antibody (1:1000, Upstate)), washed during one day with PBS-T 1X, then incubated overnight at 4°C with the secondary antibody (Alexa Fluor 488 goat anti-rabbit IgG secondary antibodies (1:250, Invitrogen)) then washed with PBS-T 1X. Embryos were then incubated in PBS-T 1X for 20 minutes with propidium iodide a DNA intercalating agent (Molecular Probes, 10 μ L/1mL). Finally, embryos were washed for 5 minutes and mounted between slide and coverslip in Fluoroshield Mounting Medium (Vector). Confocal microscope images were captured on an inverted photoscope (DMIRB; Leitz) equipped with a laser confocal imaging system (TCS SP5; Leica) using an HCX PL APO 1.4 NA 63 oil objective (Leica). Images from fixed, immunostained embryos are merged confocal z-stacks taken sequentially in the green and red channels for the anti-acetylated histone H4 labelling and the propidium

iodide signal respectively. Crosses from which confocal microscope images were obtained (Figs 1-3) are listed in S3 Table.

Proportion of unhatched developed embryos in CI crosses

To study the proportion of unhatched developed embryo in CI crosses, we performed a total of 32 crosses: 20 crosses involving four lines for the males (MAL lines) and five lines for the females (FEM lines), 5 involving Sl(wPipI-Tunis) for the males and the five FEM lines, and 7 involving the four MAL lines and females from different TC lines (S2 Table). All these crosses were performed using 50 females and 25 males. After 6 days in the cages, females were blood-fed and after five days eggs-rafts were collected in water pot and deposited into 24 wells plates. As hatching normally occurs approximately 48 hours after oviposition, developmental status in non-viable rafts was characterized at least two days after eggs-rafts collection. To attribute a developmental status to each egg, eggs-rafts were mounted between slide and coverslip, observed and documented with an optic microscope (Axiophot2 equipped with a CCD camera, Zeiss). Two developmental statuses were discriminated i) unhatched embryos harboring no visible development (Fig 2G1), or ii) unhatched embryos with visible development (Fig 2G2). For each cross, we calculated the proportion of embryos showing development for 50 embryos per eggs-raft in 10 eggs-rafts (total of 500 eggs observed per cross).

Ploidy determination in CI developed embryos

To assess the ploidy status in unhatched developed embryos, we used a PCR/RFLP diagnosis *kdr/RsaI* that allowed discriminating between *C. pipiens* and *C. quinquefasciatus* lines, as previously described in Duron and Weill (2006)[44]. Slab (*C. quinquefasciatus*) and Ichkeul 13 (*C. pipiens*) were chosen because they exhibit an unidirectional sterile cross: fertile in the direction (♂Ichkeul 13 x ♀Slab) and sterile in the other direction (♂Slab x ♀Ichkeul 13). This PCR/RFLP test was performed on DNA extracted as describe above from i) a pool of larvae from Slab and Ichkeul 13 parental lines and ii) from eggs-rafts resulting from the two reciprocal crosses between those two lines.

Real Time Quantitative PCR

Quantification of *Wolbachia* density in male testes. In order to test the influence of *Wolbachia* densities in testes on the CI cellular intensity, we quantified them with Real Time Quantitative PCR using the LightCycler 480 system (Roche). Specific primers and procedures were described in Berticat et al. (2002)[52]. Testes from 6-day old males of Tunis, Utique, Lavar and Slab lines were sampled. Each DNA template were obtained from pools of three testis pairs and extracted as described above. Five independent DNA templates were realized for each line. To estimate the number of *Wolbachia* per mosquito testes, we amplified two different genes on each sample, the *C. pipiens* specific *ace-2* locus [54] and the *Wolbachia* specific monocopy *wsp* locus [52]. Standard curves were performed using dilutions of a pBluescriptKS vector containing a unique *ace-2* and *wsp* gene copy. Each DNA template was analyzed in triplicate for both *wsp* and *ace-2* quantification. As both genes are present as single copies per haploid genome, the ratio of *wsp* over *ace-2* signals allowed estimating the relative number of *Wolbachia* genomes per *Culex* genome, thus correcting for mosquito size and DNA quality.

Amplification of *cidA* and *cidB* genes within *wPip* genome. For each *C. pipiens-Wolbachia* line (Tunis, Utique, Lavar and Slab hosting a different *wPip* strains belonging to group I, II or III) and exhibiting different *mod* profiles, quantitative PCRs were carried out to estimate the number of copies of *cidA* and *cidB* genes per *wPip* genome. Three different quantitative PCRs were performed on DNA samples extracted from ten six-day old males per line following the procedure described in Berticat et al. (2002)[52]: i) specific of the locus *wsp*, ii) specific of a 189bp fragment of the *cidA* gene conserved between all *wPip* strains using primers *wPip_0282_QPCR_2_Dir* (5'-AGG-TCC-TGT-ATT-TGA-TTT-CTG-GA) and *wPip_0282_QPCR_2_Rev* (5'-TGA-ACG-CGA-GAA-AGA-GCA-AG), and iii) specific of a 135bp fragment of the *cidB* gene conserved between all *wPip* strains using primers *wPip_0283_QPCR_1_Dir* (5'-TGA-GTG-TTT-GGA-GAA-TGA-AGG-A) and *wPip_0283_QPCR_1_Rev* (5'-TTC-CCA-AAA-GCA-AAA-CCA-GTT). Standard curves of *cidA*, *cidB* and *wsp* genes were performed using dilution of the PCR product of these three genes previously quantified using the flurorometre-QuBit (Invitrogen). Each DNA template was analysed in

triplicate for *wsp*, *cidA* and *cidB* locus. *CidA* and *cidB* copy numbers were estimated using the ratio of *cidA* or *cidB* estimated copy number over *wsp* estimated copy number, to obtain a copy number per *Wolbachia* genome since *wsp* is present in one copy in *Wolbachia* genome.

Expression of *cidA* and *cidB* genes. For each line Tunis, Utique, Lavar and Slab, ten six-day old males were used for RNA extraction with Trizol (Life Technologies) and treated with DNase with the TURBO DNA-free™ Kit (Life Technologies), in accordance with the manufacturer's instructions. 2-5 µg of each total RNA sample were reverse-transcribed into cDNA with the SuperScript III Reverse Transcriptase Kit and 30 ng of random oligomer primers ((RP)10; Invitrogen, Life Technologies). Three different quantitative PCRs were performed: i) specific of *wsp* locus, ii) specific of *cidA* locus, and iii) specific of *cidB* locus as describe above. Each DNA template was analysed in triplicate for *wsp* *cidA* and *cidB* locus. Levels of expression of *cidA* and *cidB* genes were estimated relatively to *wsp* genes by using the ratio of expression of these two genes over *wsp*.

Determination of *cidA/cidB* variants repertoire in the different *wPip* hosted by males

To describe the diversity of *cidA*^{wPip}/*cidB*^{wPip} repertoires for the two *C. pipiens* lines Utique and Slab not yet investigated, cloning and Sanger sequencing of the *cidA* and *cidB* variants were performed as described in Bonneau et al. (2018)[36] on DNA from pools of larvae extracted as described above. Variant sequences were aligned, using the Muscle algorithm implemented in Seaview 6.4.1 software.

Statistical Analysis

Variability of unhatched developed embryo proportion in sterile crosses was analyzed using a generalized linear model (GLM): $UDEP = CROSS + \epsilon$, with UDEP the proportion of unhatched developed embryos for each cross (CROSS, which represent the interaction between the MAL and FEM lines) and ϵ the error parameter, following a binomial distribution. To test the specific effect of the four MAL lines and the five FEM lines separately, GLMs with mixed effects (GLMM) were used: $UDEP = MALE +$

FEMALE + 1|CROSS + ϵ with MALE and FEMALE respectively the MAL and FEM lines involved in each cross as fixed effects, with CROSS as the interaction between MAL and FEM lines as a random effect (as crosses to produce embryos necessary require an interaction between females and males), and ϵ the error parameter, following a binomial distribution. To test for a specific effect of the host genetic background in crosses involving males from SI(*w*PipI-Tunis) and Tunis lines which host the same *Wolbachia* in two different genetic background we used a GLMM : UDEP = MALBACK + FEMALE + 1|CROSS + ϵ with UDEP the unhatched developed embryos proportion for each cross involving males from SI(*w*PipI-Tunis) and Tunis lines (MALBACK) and the five FEM lines (FEMALE) as fixed effects, with CROSS as a random effect, and with ϵ the error parameter, following a binomial distribution.

For several variables (*Wolbachia* density in testes, *cidA* and *cidB* expressions and copy number) obtained with q-PCR, variability between the four MAL lines was analyzed using GLMs in the form VAR = MALE + ϵ , with VAR one of the estimated variable of the MAL line (MALE) and ϵ the error parameter, following a Gaussian distribution.

Spearman correlation tests [55] were used to test for correlation between these variables (*Wolbachia* density in testes, *cidA* and *cidB* expressions and copy number) and the proportion of unhatched developed embryos for each MAL line. We did the same for the relation between the number of different *cidA* variants and the proportion of unhatched developed embryos for each MAL line. Finally, Wilcoxon test [56] was used to compare mean proportions of unhatched developed embryos between the two MAL lines harboring only two different *cidB* variants and the two MAL lines harboring four different *cidB* variants.

All computations were performed using the R version 3.4.4 [57]. Computed models were simplified by testing the significance of the different terms using likelihood ratio tests (LRT) and starting from the higher-order terms, as described in Crawley [58]. Factor levels of qualitative variables that were not different in their estimates (using LRTs) were grouped as described by Crawley [58]. The normality of the residuals was tested using Shapiro test for models with Gaussian error [59]. For models with Binomial error, overdispersion was calculated using the “dispersion_glmer” function from the

package *blme* for GLMM, and by dividing the residual deviance by the residuals degree of freedom of the model for GLM [60]; when detected, overdispersion was taken into account in the LRTs [61,62].

Supporting information

See Annexes at the end of the thesis manuscript.

Data availability

Nucleotide and amino acid sequences of *cidA- cidB* variants were deposited in GenBank and accession numbers are available in the S5 Table. The authors declare that all other data supporting the findings of this study are available within the article and its supporting information files.

Acknowledgements

We thank Sandra Unal, Adeline Toro and Marco Perriat-Sanguinet for technical support in the crossing experiments, Dr Rudy Caparros for technical support in cloning and Sanger sequencing and Mine Altinli for providing us with some of the TC-treated lines. We also thank Dr Philippe Clair for his help in the Real Time Quantitative PCR experiments which were performed through the technical facility of the “qPCR Haut Debit (qPHD) Montpellier génomiX” platform. Sequencing data were produced through the technical facilities of the LabEX "Centre Méditerranéen de l'Environnement et de la Biodiversité" in the GENSEQ platform. Confocal microscopy imaging was performed in the MRI-CRBM platform. Optical observations of embryos 48 hours after oviposition were performed at the CytoEvol facilities of UMR ISEM - CBGP of the LabEx CeMEB.

References

1. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol*. 2008;6: 741–751. doi:10.1038/nrmicro1969
2. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. How many species are infected with *Wolbachia*? a statistical analysis of current data. *FEMS Microbiol Lett*. 2008;281: 215–220. doi:10.1111/j.1574-6968.2008.01110.x
3. Zug R, Hammerstein P. Still a Host of Hosts for *Wolbachia*: Analysis of Recent Data Suggests That 40% of Terrestrial Arthropod Species Are Infected. *PLoS One*. 2012;7: e38544. doi:10.1371/journal.pone.0038544
4. Yen JH, Barr AR. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *J Invertebr Pathol*. 1973;22: 242–250. doi:10.1016/0022-2011(73)90141-9
5. Callaini G, Riparbelli MG, Giordano R, Dallai R. Mitotic Defects Associated with Cytoplasmic Incompatibility in *Drosophila simulans*. *J Invertebr Pathol*. 1996;67: 55–64. doi:10.1006/jipa.1996.0009
6. Engelstädter J, Telschow A. Cytoplasmic incompatibility and host population structure. *Heredity*. 2009;103: 196–207. doi:10.1038/hdy.2009.53
7. Laven H. Speciation and evolution in *Culex pipiens*. In: Wright J, Pal R, editors. *Genetics of Insect Vectors of Disease*. Elsevier: Amsterdam; 1967. pp. 251–275.
8. O'Neill SL, Karr TL. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature*. 1990;348: 178–180. doi:10.1038/348178a0
9. Breeuwer JAJ, Werren JH. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature*. 1990;346: 558–560. doi:10.1038/346558a0
10. Bordenstein SR, O'Hara FP, Werren JH. *Wolbachia*-induced incompatibility precedes other

- hybrid incompatibilities in *Nasonia*. *Nature*. 2001;409: 707–710. doi:10.1038/35055543
11. Duron OI, Bernard C, Unal S, Berthomieu A, Berticat C, Weill M. Tracking factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. *Mol Ecol*. 2006;15: 3061–3071. doi:10.1111/j.1365-294X.2006.02996.x
 12. Atyame CM, Delsuc F, Pasteur N, Weill M, Duron O. Diversification of *Wolbachia* Endosymbiont in the *Culex pipiens* Mosquito. *Mol Biol Evol*. 2011;28: 2761–2772. doi:10.1093/molbev/msr083
 13. Merçot H, Charlat S. *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. *Genetica*. 2004. pp. 51–59. doi:10.1007/978-94-007-0965-2_4
 14. Duron O, Fort P, Weill M. Influence of aging on cytoplasmic incompatibility, sperm modification and *Wolbachia* density in *Culex pipiens* mosquitoes. *Heredity*. 2007;98: 368–374. doi:10.1038/sj.hdy.6800948
 15. Bordenstein SR, Werren JH. Bidirectional incompatibility among divergent *Wolbachia* and incompatibility level differences among closely related *Wolbachia* in *Nasonia*. *Heredity*. 2007;99: 278–287. doi:10.1038/sj.hdy.6800994
 16. Bourtzis K, Nirgianaki A, Markakis G, Savakis C. *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. *Genetics*. 1996;144: 1063–73.
 17. Breeuwer JA, Werren JH. Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. *Genetics*. 1993;135: 565–74.
 18. Bressac C, Rousset F. The Reproductive Incompatibility System in *Drosophila simulans*: Dapi-Staining Analysis of the *Wolbachia* Symbionts in Sperm Cysts. *J Invertebr Pathol*. 1993;61: 226–230. doi:10.1006/jipa.1993.1044
 19. Bordenstein SR, Uy JJ, Werren JH. Host genotype determines cytoplasmic incompatibility type in the haplodiploid genus *Nasonia*. *Genetics*. 2003;164: 223–33.

20. Lassy CW, Karr TL. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. *Mech Dev*. 1996;57: 47–58. doi:10.1016/0925-4773(96)00527-8
21. Ryan SL, Saul GB. Post-fertilization effect of incompatibility factors in *Mormoniella*. *MGG Mol Gen Genet*. 1968;103: 29–36. doi:10.1007/BF00271154
22. Landmann F, Orsi GA, Loppin B, Sullivan W. *Wolbachia*-Mediated Cytoplasmic Incompatibility Is Associated with Impaired Histone Deposition in the Male Pronucleus. *PLoS Pathog*. 2009;5: e1000343. doi:10.1371/journal.ppat.1000343
23. Tram U, Sullivan W. Role of Delayed Nuclear Envelope Breakdown and Mitosis in *Wolbachia*-Induced Cytoplasmic Incompatibility. *Science*. 2002;296: 1124–1126. doi:10.1126/science.1070536
24. Reed KM, Werren JH. Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria (*Wolbachia*): A comparative study of early embryonic events. *Mol Reprod Dev*. 1995;40: 408–418. doi:10.1002/mrd.1080400404
25. Callaini G, Dallai R, Riparbelli MG. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. *J Cell Sci*. 1997;110: 271–80.
26. Jost E. Untersuchungen zur Inkompatibilit t im *Culex pipiens*-Komplex. *Wilhelm Roux Arch Entwickl Mech Org*. 1970;166: 173–188. doi:10.1007/BF00576992
27. Tram U, Fredrick K, Werren JH, Sullivan W. Paternal chromosome segregation during the first mitotic division determines *Wolbachia*-induced cytoplasmic incompatibility phenotype. *J Cell Sci*. 2006;119: 3655–3663. doi:10.1242/jcs.03095
28. Vavre F, Fleury F, Varaldi J, Fouillet P, Bouleatreau M. Evidence for female mortality in *Wolbachia*- mediated cytoplasmic incompatibility in haplodiploid insects: epidemiologic and evolutionary consequences. *Evolution*. 2000;54: 191–200. doi:10.1111/j.0014-

3820.2000.tb00019.x

29. Vavre F, Dedeine F, Quillon M, Fouillet P, Fleury F, Boulétreau M. Within- species diversity of *Wolbachia*- induced cytoplasmic incompatibility in haplodiploid insects. *Evolution*. 2001;55: 1710–1714. doi:10.1111/j.0014-3820.2001.tb00691.x
30. Tram U, Ferree PM, Sullivan W. Identification of *Wolbachia*–host interacting factors through cytological analysis. *Microbes Infect*. 2003;5: 999–1011. doi:10.1016/S1286-4579(03)00192-8
31. Breeuwer JAJ. *Wolbachia* and cytoplasmic incompatibility in the spider mites *Tetranychus urticae* and *T. turkestanii*. *Heredity*. 1997;79: 41–47. doi:10.1038/sj.hdy.6881810
32. Werren JH. Biology of *Wolbachia*. *Annu Rev Entomol*. 1997;42: 587–609.
33. Hurst LD. The evolution of cytoplasmic incompatibility or when spite can be successful. *J Theor Biol*. 1991;148: 269–277. doi:10.1016/S0022-5193(05)80344-3
34. Beckmann JF, Ronau JA, Hochstrasser M. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat Microbiol*. 2017;2: 17007. doi:10.1038/nmicrobiol.2017.7
35. LePage DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, et al. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature*. 2017;543: 243–247. doi:10.1038/nature21391
36. Bonneau M, Atyame C, Beji M, Justy F, Cohen-Gonsaud M, Sicard M, et al. *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat Commun*. 2018;9: 319. doi:10.1038/s41467-017-02749-w
37. Beckmann JF, Fallon AM. Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: Implications for cytoplasmic incompatibility. *Insect Biochem Mol Biol*. 2013;43: 867–878. doi:10.1016/j.ibmb.2013.07.002
38. Lindsey ARI, Rice DW, Bordenstein SR, Brooks AW, Bordenstein SR, Newton ILG. Evolutionary Genetics of Cytoplasmic Incompatibility Genes cifA and cifB in Prophage WO of *Wolbachia*. *Genome Biol Evol*. 2018;10: 434–451. doi:10.1093/gbe/evy012

39. Klasson L, Walker T, Sebaihia M, Sanders MJ, Quail M a., Lord A, et al. Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol Biol Evol.* 2008;25: 1877–1887. doi:10.1093/molbev/msn133
40. Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc Natl Acad Sci.* 2018;115: 4987–4991. doi:10.1073/pnas.1800650115
41. Baldo L, Hotopp JCD, Jolley K a., Bordenstein SR, Biber S a., Choudhury RR, et al. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol.* 2006;72: 7098–7110. doi:10.1128/AEM.00731-06
42. Atyame CM, Labbé P, Dumas E, Milesi P, Charlat S, Fort P, et al. *Wolbachia* Divergence and the Evolution of Cytoplasmic Incompatibility in *Culex pipiens*. *PLoS One.* 2014;9: e87336. doi:10.1371/journal.pone.0087336
43. Dumas E, Atyame CM, Milesi P, Fonseca DM, Shaikevich E V, Unal S, et al. Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol Biol.* 2013;13: 181. doi:10.1186/1471-2148-13-181
44. Duron O, Weill M. *Wolbachia* infection influences the development of *Culex pipiens* embryo in incompatible crosses. *Heredity.* 2006;96: 493–500. doi:10.1038/sj.hdy.6800831
45. Atanassov BS, Koutelou E, Dent SY. The role of deubiquitinating enzymes in chromatin regulation. *FEBS Lett.* 2011;585: 2016–2023. doi:10.1016/j.febslet.2010.10.042
46. Gebiola M, Giorgini M, Kelly SE, Doremus MR, Ferree PM, Hunter MS. Cytological analysis of cytoplasmic incompatibility induced by *Cardinium* suggests convergent evolution with its distant cousin *Wolbachia*. *Proc R Soc B Biol Sci.* 2017;284: 20171433. doi:10.1098/rspb.2017.1433
47. Penz T, Schmitz-Esser S, Kelly SE, Cass BN, Müller A, Woyke T, et al. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in *Cardinium*

- hertigii*. PLoS Genet. 2012;8. doi:10.1371/journal.pgen.1003012
48. Breeuwer JAJ, Werren JH. Effect of genotype on cytoplasmic incompatibility between two species of *Nasonia*. Heredity. 1993;70: 428–436. doi:10.1038/hdy.1993.60
 49. Atyame CM, Duron O, Tortosa P, Pasteur N, Fort P, Weill M. Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito populations. Mol Ecol. 2011;20: 286–298. doi:10.1111/j.1365-294X.2010.04937.x
 50. Duron O, Bernard J, Atyame CM, Dumas E, Weill M. Rapid evolution of *Wolbachia* incompatibility types. Proc R Soc B Biol Sci. 2012;279: 4473–4480. doi:10.1098/rspb.2012.1368
 51. Aakre CD, Herrou J, Phung TN, Perchuk BS, Crosson S, Laub MT. Evolving New Protein-Protein Interaction Specificity through Promiscuous Intermediates. Cell. 2015;163: 594–606. doi:10.1016/j.cell.2015.09.055
 52. Berticat C, Rousset F, Raymond M, Berthomieu A, Weill M. High *Wolbachia* density in insecticide-resistant mosquitoes. Proc R Soc B Biol Sci. 2002;269: 1413–1416. doi:10.1098/rspb.2002.2022
 53. Clements AN. The Biology of Mosquitoes: Development, Nutrition and Reproduction. Wallingford, UK: CABI publishing; 1992.
 54. Weill M, Berticat C, Raymond M, Chevillon C. Quantitative Polymerase Chain Reaction to Estimate the Number of Amplified Esterase Genes in Insecticide-Resistant Mosquitoes. Anal Biochem. 2000;285: 267–270. doi:10.1006/abio.2000.4781
 55. Best D, Roberts D. Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. J R Stat Soc. 1975;24: 377–379.
 56. Bauer DF. Constructing confidence sets using rank statistics. J Am Stat Assoc. 1972;67: 687–690.
 57. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R

Foundation for Statistical Computing; 2018.

58. Crawley M. *The R Book*. John Wiley, editor. Chichester, UK: John Wiley & Sons, Ltd; 2007.
59. Royston P. An extension of Shapiro and Wilk's W test for normality to large samples. *Appl Stat.* 1982;31: 115–124.
60. Korner-Nievergelt F, Roth T, Felten Sv, Guelat J AB and K-NP. *Bayesian Data Analysis in Ecology using Linear Models with R, BUGS and Stan*. Press A, editor. Elsevier: New York; 2015.
61. Lebreton J-D, Burnham KP, Clobert J, Anderson DR. Modeling Survival and Testing Biological Hypotheses Using Marked Animals: A Unified Approach with Case Studies. *Ecol Monogr.* 1992;62: 67–118. doi:10.2307/2937171
62. Anderson DR, Burnham KP, White GC. AIC Model Selection in Overdispersed Capture-Recapture Data. *Ecology.* 1994;75: 1780–1793. doi:10.2307/1939637

II. Résumé des résultats & Discussion

a. Un seul phénotype cellulaire à l'origine de l'IC chez *C. pipiens*

L'observation en microscopie confocale des embryons issus de croisements incompatibles a révélé que ce sont des défauts de condensation et de ségrégation de la chromatine paternelle lors de l'anaphase de la première division qui sont à l'origine d'un arrêt du développement et de la mort des embryons chez *C. pipiens*. Seuls ces défauts embryonnaires ont été observés malgré la diversité de croisements réalisée, suggérant que la diversité des phénotypes d'IC ne repose pas sur une diversité de mécanismes cytologiques. De plus, ces mêmes défauts ont été observés dans les embryons issus de croisement où les femelles étaient non infectées, ce qui suggère que l'absence ou la présence de *Wolbachia* chez les femelles ne modifie pas les premières étapes de l'IC. Ces défauts de développement sont similaires à ceux précédemment décrits chez *Drosophila* et *Nasonia* (Callaini *et al.*, 1997; Tram & Sullivan, 2002). Cette similarité est probablement due au fait que les *Wolbachia* infectant *D. melanogaster*, *D. simulans*, *N. vitripennis* et *C. pipiens*, possèdent toutes les gènes *cidA* et *cidB*. La détection d'un seul et même mécanisme cellulaire responsable de la mort des embryons suggère que la diversité des répertoires de variants *cidA* et *cidB* présents chez les *wPips* impliquées dans les croisements étudiés n'est pas responsable d'une variété de mécanismes cellulaires. Ceci est probablement dû au fait que tous les variants de *cidB* possèdent un domaine catalytique DUB (impliqué dans la déubiquitination) identique, interagissant probablement de la même manière avec les protéines hôtes ciblées. Les voies d'ubiquitination/déubiquitination ont été montrées comme impliquées dans des processus cellulaires essentiels tels que la régulation de la structure de la chromatine ou la synchronisation des cycles cellulaires (Atanassov *et al.*, 2011). Ainsi, des changements d'ubiquitination de protéines pourraient par exemple jouer un rôle direct ou indirect dans la restructuration de la chromatine lors de l'incorporation des histones H3.3 après l'enlèvement des protamines.

La déubiquitination ne semble pas être la seule voie moléculaire employée par les *Wolbachia* pour induire l'IC car il existe des souches de *Wolbachia* capables d'induire l'IC sans *cidA/cidB*. En revanche, ces souches de *Wolbachia* possèdent dans leur génome un paralogue de *cidA/cidB* nommé *cinA/cinB*. La protéine *cinB* possède un domaine catalytique de type nucléase (Beckmann *et al.*, 2017; LePage *et al.*, 2017; Lindsey *et al.*, 2018). DUB et nucléase ne possèdent pas les mêmes fonctions catalytiques, ce qui suggère que des voies moléculaires différentes pourraient mener à l'induction de l'IC chez les insectes. A l'heure actuelle, aucune étude des mécanismes cytologiques de l'IC induite par une souche de *Wolbachia* ne possédant que l'opéron *cinA/cinB* n'a été réalisée. En revanche, les mécanismes cellulaires de l'IC induits par *wPip* qui possède les deux types d'opérons sont similaires à ceux induits par *wMel* qui possède uniquement l'opéron *cidA/cidB*. Ces résultats suggèrent que l'association des deux opérons ne change pas le phénotype cellulaire de l'IC. Il est cependant possible que les phénotypes

cellulaires d'IC induit par les opérons *cin* et *cid* soient très similaires bien que les voies moléculaires empruntées par les deux opérons pour l'induction soient différentes.

b. L'exclusion complète de la chromatine paternelle à l'origine du développement haploïde

L'étude de l'embryogénèse précoce chez des embryons de *C. pipiens* issus de croisements incompatibles a permis de mettre en évidence un lien entre le degré d'exclusion de la chromatine paternelle lors de la première division embryonnaire et la présence d'embryons montrant des signes de développement avancés. En effet, chez *C. pipiens*, bien que 99,99% des embryons d'un croisement incompatible meurent, 48 heures après oviposition, certains présentent des caractéristiques d'un développement embryonnaire avancé (présence de segments, d'yeux, de soies...) alors que d'autres ne présentent aucun développement visible en microscopie optique (Duron & Weill, 2006). Ces embryons montrant un développement marqué sont haploïdes et uniquement constitués de chromatine maternelle (Duron & Weill, 2006). Ici, nous montrons, grâce à nos observations en microscopie confocale, que ce développement haploïde peut avoir lieu lorsque la chromatine paternelle est complètement exclue de la première division embryonnaire permettant à la chromatine maternelle de se diviser en deux noyaux haploïdes. A l'inverse, les embryons ne présentant aucun développement visible 48 heures après oviposition proviendraient d'une exclusion partielle de la chromatine paternelle entraînant la production de noyaux aneuploïdes responsables d'un arrêt précoce de l'embryogénèse.

Nous avons utilisé le lien entre le degré d'exclusion de la chromatine paternelle et la production d'embryons haploïdes pour étudier la variabilité de l'intensité cellulaire de l'IC entre différents croisements incompatibles. La proportion d'embryons haploïdes issus d'un croisement incompatible correspond à la fréquence à laquelle une exclusion complète de la chromatine paternelle a eu lieu. Une proportion d'embryons haploïdes élevée traduit donc une exclusion totale très fréquente et donc une forte intensité cellulaire de l'IC.

c. L'intensité cellulaire de l'IC sous la dépendance des *Wolbachia* présentes chez les mâles et de la diversité de *cidB*

Nous avons analysé la proportion d'embryons haploïdes issus de croisements incompatibles impliquant des mâles infectés par des *Wolbachia* de différents groupes (*wPipI*, II et III) induisant différents phénotypes *mod* avec des femelles infectées par différentes souches de *wPipIV* possédant le même phénotype *resc*. Nous avons observé que cette proportion variait en fonction des croisements et était corrélée aux types de mâle impliqués dans le croisement. Des expériences de backcross ont permis de montrer que le fond génétique hôte n'était pas impliqué dans la variation de l'intensité cellulaire de

l'IC. Alors que des études précédentes (Duron *et al.*, 2006b; Atyame *et al.*, 2011c; b, 2015b) avaient démontré que les *wPips* étaient seules maîtresses de la diversité des phénotypes d'IC observée chez *C. pipiens*, nos résultats montrent qu'elles dictent également la variation de l'intensité de l'IC à l'échelle cellulaire.

Nous avons par la suite étudié chez les différentes souches de *wPip* infectant les mâles, différents facteurs pouvant être responsables de la variation de l'intensité cellulaire de l'IC. Nous avons étudié : i) l'expression de *cidA* et *cidB*, ii) la densité des *Wolbachia* dans les testicules des différents mâles et iii) le nombre de copies et de variants différents de *cidA* et *cidB* dans les différentes souches de *wPips*.

(i) Ni l'expression de *cidA* ni celle de *cidB* n'est significativement différente entre les souches de *wPips* infectant les mâles, ce qui indique que ce n'est pas un différentiel de niveau d'expression qui influence l'intensité cellulaire de l'IC. En revanche, les deux gènes ne sont pas exprimés de la même manière : comme décrit précédemment chez *D. melanogaster* (LePage *et al.*, 2017; Lindsey *et al.*, 2018), *cidA* est significativement plus exprimé que *cidB* quelle que soit la souche de *wPip*, ce qui est en accord avec l'hypothèse que *cidA* et *cidB* forment un système toxine-antidote où *cidA*, l'antidote, serait transcrit plus fortement que *cidB*, la toxine (Beckmann *et al.*, 2017).

(ii) Bien que les niveaux d'expression respectifs de *cidA* et de *cidB* ne varient pas entre les différentes *wPips*, les quantités totales de protéines CidA et CidB dans l'hôte dépendent de la quantité de *Wolbachia*. Puisque les facteurs *mod* sont très probablement déposés dans les spermatozoïdes au moment de la spermatogenèse, nous avons quantifié la densité de *Wolbachia* dans les testicules des mâles. Les mâles de la lignée Lavar infectés par des *wPipII* présentent une densité de *Wolbachia* dans leurs testicules significativement inférieure aux autres mâles. Ces mâles Lavar étaient aussi ceux induisant la proportion d'embryons haploïdes la plus faible (*i.e.* l'intensité cellulaire d'IC la plus faible). A cause d'une densité de *Wolbachia* plus faible, la quantité globale de protéines CidB produite par les *Wolbachia* chez les mâles Lavar est plus faible que chez les autres mâles. Cette faible quantité de protéines CidB pourrait être responsable d'une plus faible perturbation de la condensation de la chromatine paternelle entraînant la production de très peu d'embryons haploïdes. Cependant, ces observations ne sont valables que pour la lignée Lavar et d'autres lignées présentant des quantités de *Wolbachia* différentes sont nécessaires pour confirmer cette hypothèse.

(iii) Le nombre de copies des gènes *cidA* et de *cidB* n'est pas corrélé avec l'intensité cellulaire de l'IC. En revanche, le nombre de variants différents présents dans les répertoires *cidB* des *wPips* infectant les mâles est significativement corrélé à la proportion d'embryons haploïdes. En effet, les deux lignées dont les mâles induisent la plus forte proportion d'embryons haploïdes, sont les deux lignées qui possèdent le plus de variants de *cidB* différents. L'hypothèse formulée pour expliquer cette corrélation serait que chaque variant de *cidB* impacterait différemment la chromatine paternelle (*i.e.* comme plusieurs serrures). Plus le nombre de variants différents de *cidB* serait élevé dans un génome de *wPip*,

plus il y aurait de protéines différentes de CidB introduites dans l'œuf qui pourraient modifier différemment la chromatine paternelle. L'étude d'autres lignées de *C. pipiens* montrant une amplification et une diversification variables de CidB est également nécessaire pour confirmer cette hypothèse.

d. L'absence de *Wolbachia* chez les femelles empêche tout développement haploïde de se produire

Lorsque des femelles non infectées, issues de lignées dont les *Wolbachia* ont été retirées par un traitement antibiotique à la tétracycline, sont croisées avec les mâles infectés par les différentes *wPip*I, II ou III, aucun embryon haploïde n'est produit. Les observations en microscopie confocale d'embryons issus de ces croisements, même plusieurs heures après l'oviposition, ne présentent que quelques noyaux dégénérés. Ces observations suggèrent que l'IC, à l'échelle cellulaire, induite par les *wPips* infectant ces mâles est toujours faible. Cependant, nous avons montré lors de croisements avec des femelles infectées par des *wPips* incompatibles que ces mêmes mâles étaient tout à fait capables d'induire de fortes intensités cellulaires d'IC. Par conséquent, le phénotype d'intensité faible observé systématiquement lorsque les mères ne sont pas infectées viendrait de l'absence de *Wolbachia* dans les œufs plutôt que de l'incapacité des mâles à induire une forte IC. Ainsi, les *Wolbachia* présentes dans les œufs influenceraient l'issue des premières étapes de l'embryogenèse permettant le développement haploïde. Leur absence empêcherait ce type de développement. Plusieurs hypothèses peuvent être proposées pour expliquer pourquoi les *Wolbachia* maternelles sont nécessaires au développement haploïde. Une première hypothèse serait que les *Wolbachia* maternelles présentes dans l'œuf, accentueraient dans les croisements incompatibles les mécanismes responsables de la mauvaise condensation de la chromatine paternelle résultant en une exclusion totale de celle-ci et en la production d'embryons haploïdes. Cette hypothèse est peu probable car elle équivaldrait à accentuer la fonction *mod* alors que les *Wolbachia* maternelles sont impliquées dans la fonction *resc*. Une autre hypothèse serait que les protéines CidB présentes dans le sperme impacteraient aussi la chromatine maternelle et que les *Wolbachia* maternelles la « protégeraient », permettant sa ségrégation lors de la première division embryonnaire. Cependant, parmi toutes les études de cytologie des mécanismes de l'IC, aucun impact des facteurs *mod* sur la chromatine maternelle n'a été détecté (Ryan & Saul, 1968; Reed & Werren, 1995; Callaini *et al.*, 1996, 1997; Tram & Sullivan, 2002; Tram *et al.*, 2003, 2006; Landmann *et al.*, 2009). Une autre hypothèse serait que la présence d'une *Wolbachia* incompatible dans l'œuf pourrait avoir un effet additif sur l'incompatibilité entre les pronuclei paternel et maternel, pas directement en affectant la chromatine paternelle mais en influençant le timing du cycle cellulaire. Par exemple, les *Wolbachia* maternelles pourraient moduler la cinétique de réplication de la chromatine maternelle ou l'entrée en mitose,

augmentant ainsi l'incompatibilité entre les pronuclei et favorisant par conséquent le développement haploïde.

Des défauts de condensation puis de ségrégation de la chromatine paternelle sont responsables de la mort des embryons dans tous les croisements incompatibles analysés chez *C. pipiens*. Des défauts similaires ont été décrits chez *Drosophila spp.* et *Nasonia spp.* ce peut être attribué à la présence chez toutes ces *Wolbachia* de l'opéron *cidA/cidB* impliqué dans l'induction de l'IC. La grande diversité des phénotypes d'IC observée chez *C. pipiens*, est donc basée sur un seul et même mécanisme cellulaire.

Au sein des embryons issus de croisements incompatibles, donc tous voués à la mort, certains vont présenter un développement avancé sans pour autant éclore. Ces embryons sont haploïdes et résultent de l'exclusion totale de la chromatine paternelle lors de la première division embryonnaire. Cette exclusion est interprétée comme résultant de défauts plus importants de la chromatine paternelle (*i.e.* forte intensité de l'IC à l'échelle cellulaire). L'intensité des défauts varie selon les mâles impliqués dans les croisements et apparaît comme corrélée au nombre de variants *cidB* différents présents dans les génomes des *Wolbachia* les infectant. Enfin, l'absence de *Wolbachia* maternelles dans les œufs empêche tout développement haploïde de se produire. Ceci pourrait être dû au fait que lorsque les *Wolbachia* maternelles sont présentes, elles interfèrent avec le cycle cellulaire lors de la première division embryonnaire, aggravant l'incompatibilité entre le matériel génétique paternel et maternel et permettant ainsi le développement haploïde.

Ainsi, malgré les nombreuses spécificités de l'IC chez *C. pipiens* qui ont fait de cette espèce une exception parmi les hôtes chez lesquels *Wolbachia* induit l'IC, il s'avère que les mécanismes cellulaires responsables de la mort des embryons sont les mêmes que ceux décrits chez d'autres espèces. *C. pipiens* reste cependant une exception intrigante si l'on considère l'absence de développement haploïde lorsque les femelles ne sont pas infectées.

Discussion & Perspectives

L'identification d'une protéine issue d'un gène de *Wolbachia* dans le sperme de *Culex pipiens* (i.e. la protéine CidA) par Beckmann & Fallon (2013) a constitué une grande avancée dans la compréhension des bases génétiques et moléculaires de l'IC induite par *Wolbachia*. Bien que les premiers croisements incompatibles aient été découverts en 1937 (Marshall & Staley, 1937) et attribués à *Wolbachia* en 1971 (Yen & Barr, 1971), aucune étude jusqu'à récemment n'avait pu mettre en évidence des gènes bactériens sous-jacent à cette manipulation malgré de nombreuses tentatives par des approches génétiques et génomiques sur des souches variées de *Wolbachia* (Guillemaud *et al.*, 1997; Iturbe-Ormaetxe *et al.*, 2005; Sinkins *et al.*, 2005; Duron *et al.*, 2007a; Walker *et al.*, 2007; Sutton *et al.*, 2014; Newton *et al.*, 2016). Les récentes études de transgénèse ont par ailleurs démontré l'implication de l'opéron *cidA/cidB* dans un système « simple » d'IC entre mâles infectés et femelles non infectées (Beckmann *et al.*, 2017; LePage *et al.*, 2017; Shropshire *et al.*, 2018).

Cependant, la question de l'IC et de sa diversité entre individus infectés n'avait jusqu'ici pas été abordée. C'est pourquoi, au cours de ma thèse, nous avons confronté ces gènes au « modèle *C. pipiens* » qui présente la plus grande diversité de phénotype d'IC révélée à ce jour et chez lequel tous les individus sont infectés par *Wolbachia*. Les gènes *cidA*, *cidB* et *cinA*, *cinB* pour être impliqués dans la diversité des phénotypes d'IC doivent respecter un cahier des charges très clair et restrictif. En effet, les gènes candidats au déterminisme génétique de la diversité des phénotypes d'IC chez *C. pipiens*, doivent 1) être présents en plusieurs copies différentes au sein d'un même génome de *wPip* et 2) présenter des copies différentes entre des souches de *wPips* induisant des phénotypes d'IC différents.

Nos données récentes de séquençages NGS de différentes souches de *wPip* ont permis de révéler la présence de plusieurs copies polymorphes de *cidA* et *cidB* au sein d'un même génome de *wPip* et entre génomes de *wPips*. En revanche, aucun polymorphisme n'a été détecté entre les gènes *cinA* et *cinB* des différentes souches (Partie I, Chapitre I: Bonneau *et al.*, 2018a). Il semble donc que, même si les gènes *cinA* et *cinB* puissent être impliqués dans l'IC (ce qui est en cours de publication d'après la communication de Hochstrasser *et al.* au congrès *Wolbachia* 2018), ils ne participent pas à la diversité des phénotypes d'IC observés chez *C. pipiens*. Le séquençage Sanger de souches de *Wolbachia* induisant des phénotypes d'IC distincts appartenant à des groupes phylogénétiques *wPip* différents a révélé que ces souches possédaient des répertoires de variants de *cidA* et *cidB* différents. Les gènes *cidA* et *cidB* remplissent donc le cahier des charges initial pour tout gène potentiellement impliqué dans la diversité des phénotypes d'IC chez *C. pipiens*.

I. Les phages au cœur de l'IC

a. Une origine phagique de l'IC : une vieille histoire

L'association des gènes *cidA* et *cidB* avec le prophage WO chez de nombreuses souches de *Wolbachia* capables d'induire l'IC (Klasson *et al.*, 2008; LePage *et al.*, 2017; Lindsey *et al.*, 2018) conforte l'idée que l'origine de l'IC chez *Wolbachia* pourrait être d'origine phagique. Cette idée a plus de 20 ans et est basée sur quatre faits majeurs. Premièrement, chez de nombreuses bactéries, qu'elles soient libres ou endocellulaires, des gènes portés par les phages sont impliqués dans leur interactions pathogènes mais aussi mutualistes avec leurs hôtes (Miao & Miller, 1999; Boyd & Brüssow, 2002). C'est par exemple le cas chez *Vibrio cholerae* dont la « Cholera toxin » est codée par le gène *ctxAB* porté par le bactériophage CTXΦ (Waldor & Mekalanos, 1996) ou chez *Acyrtosiphon pisum* où les gènes permettant à la bactérie *Hamiltonella defensa* de protéger en partie ce puceron vis-à-vis du parasitoïde *Aphidius ervi*, sont présents dans le génome du bactériophage APSE inséré dans celui de *H. defensa* (Oliver *et al.*, 2009). Deuxièmement, ces bactériophages constituent les navettes de gènes de virulence les plus communes chez les bactéries (Miao & Miller, 1999; Boyd *et al.*, 2001; Boyd & Brüssow, 2002). Troisièmement, plus spécifiquement, les phages WO contiennent des gènes codant pour des protéines à domaine ankyrine (Wu *et al.*, 2004; Walker *et al.*, 2007; Klasson *et al.*, 2008, 2009). Ces protéines du fait de leurs interactions avec les protéines hôtes, notamment au niveau de la régulation du cycle cellulaire, qui est un des mécanismes cellulaires impacté par l'IC, constituent des candidats robustes pour les bases génétiques de l'IC (Sinkins *et al.*, 2005; Duron *et al.*, 2007a; Klasson *et al.*, 2008). Enfin, les prophages WO sont quasiment exclusivement observés dans les génomes des *Wolbachia* manipulatrices de la reproduction (Wu *et al.*, 2004; Foster *et al.*, 2005; Klasson *et al.*, 2008, 2009). La grande variabilité génétique présente dans les régions de phages entre les souches de *Wolbachia* infectant *C. pipiens* constituait une piste très prometteuse pour expliquer la diversité des patrons d'IC observée chez cette espèce, même si aucune corrélation parfaite n'avait été trouvée entre ces gènes et les phénotypes d'IC (Duron *et al.*, 2006a).

Bien que les gènes *cidA* et *cidB* soient associés aux phages, il est très peu probable que ce soient des protéines d'origine phagiques à proprement parler (*i.e.* impliquées dans la formation de la capsid de la queue, la tête, la plaque terminale, dans la lysogénie, dans l'introduction de l'ADN dans la bactérie ...). En effet, la localisation de ces deux gènes dans la région EAM (Eukaryotic Association Module) (voir Bordenstein & Bordenstein, 2016 pour revue) ainsi que la présence du domaine DUB de la protéine CidB qui est un domaine typiquement observé chez les protéines eucaryotes (Beckmann *et al.*, 2017) suggère que ces gènes proviendraient initialement de matériel génétique eucaryote.

b. Un système toxine-antidote à la base de l'IC

Denis Poinot dans sa thèse en 1997 proposait que l'IC trouve son origine dans un mécanisme toxine-antidote initialement « utilisé » par la bactérie pour se maintenir dans les cellules de son hôte. Selon lui, une bactérie endocellulaire, ancêtre des *Wolbachia* actuelles, aurait acquis la capacité de sécréter, dans ses cellules-hôtes, une toxine (poison) et un antidote (Figure 1a). Ce poison et cet antidote présents dans le cytoplasme des cellules seraient transmis aux cellules filles au cours de la division cellulaire. L'antidote étant une protéine moins stable que le poison, l'antidote hérité de la cellule mère est dégradé plus rapidement que le poison dont la toxicité est alors révélée. Si la transmission des bactéries aux cellules filles issues des divisions mitotiques est imparfaite, alors certaines d'entre elles ne posséderont pas la bactérie et se trouveront ainsi incapables de produire l'antidote permettant de contrer le poison hérité via le cytoplasme. A l'inverse, les cellules qui contiennent cette bactérie continueront à produire l'antidote, ce qui leur permettra de survivre. Ce système poison/antidote sélectionnerait ainsi pour la conservation de l'endosymbiote par les cellules hôtes : on parle de gènes d'addiction.

Ce système poison-antidote aurait une conséquence particulière lorsque les cellules en question sont des spermatozoïdes (Figure 1b). En effet, lors de la spermatogenèse, les *Wolbachia* sont systématiquement éliminées avec le surplus de cytoplasme, les spermatozoïdes seraient donc toujours impactés par le poison à moins qu'il n'y ait ajout d'antidote au moment de la fécondation. L'absence d'antidote dans l'œuf entraînerait l'activation de la fonction toxique sur la chromatine paternelle et la mort des embryons. L'absence d'antidote étant synonyme d'absence de *Wolbachia* dans les œufs, la mort n'aurait lieu que pour les œufs non infectés. On comprend que cela entraînerait une diminution du nombre de descendants des femelles non infectées, alors que les femelles infectées ne subissent pas cette diminution, permettant ainsi la propagation des *Wolbachia* dans les populations. Ce système, favorisant la propagation des *Wolbachia* dans les populations hôtes, aurait été sélectionné et aurait conduit à l'IC telle qu'on la connaît aujourd'hui.

Si l'on actualise la théorie de Denis Poinot (1997) avec nos connaissances sur les gènes *cidA* et *cidB* et leur localisation dans les phages, on peut constater qu'elle est superbement en accord avec les découvertes de ces deux dernières années. Nous pouvons d'ailleurs aller encore plus loin en proposant une origine phagique et eucaryote de l'IC. La valeur sélective des bactériophages, parasites des bactéries endosymbiotiques, dépend directement de la survie et de la multiplication de celles-ci au sein des populations d'arthropodes. Ainsi, tout gène ou autre élément génétique présent dans les bactériophages permettant à l'endosymbiote de ne pas être éliminé par l'hôte sera positivement sélectionné à la fois chez les bactériophages et chez les endosymbiotes. L'ancêtre de l'opéron *cidA/cidB*, probablement acquis par les bactériophages dans un génome eucaryote, est parfait pour ce rôle. La protéine CidB n'est, a priori, pas toxique pour la bactérie puisque le système d'ubiquitination/deubiquitination n'est pas

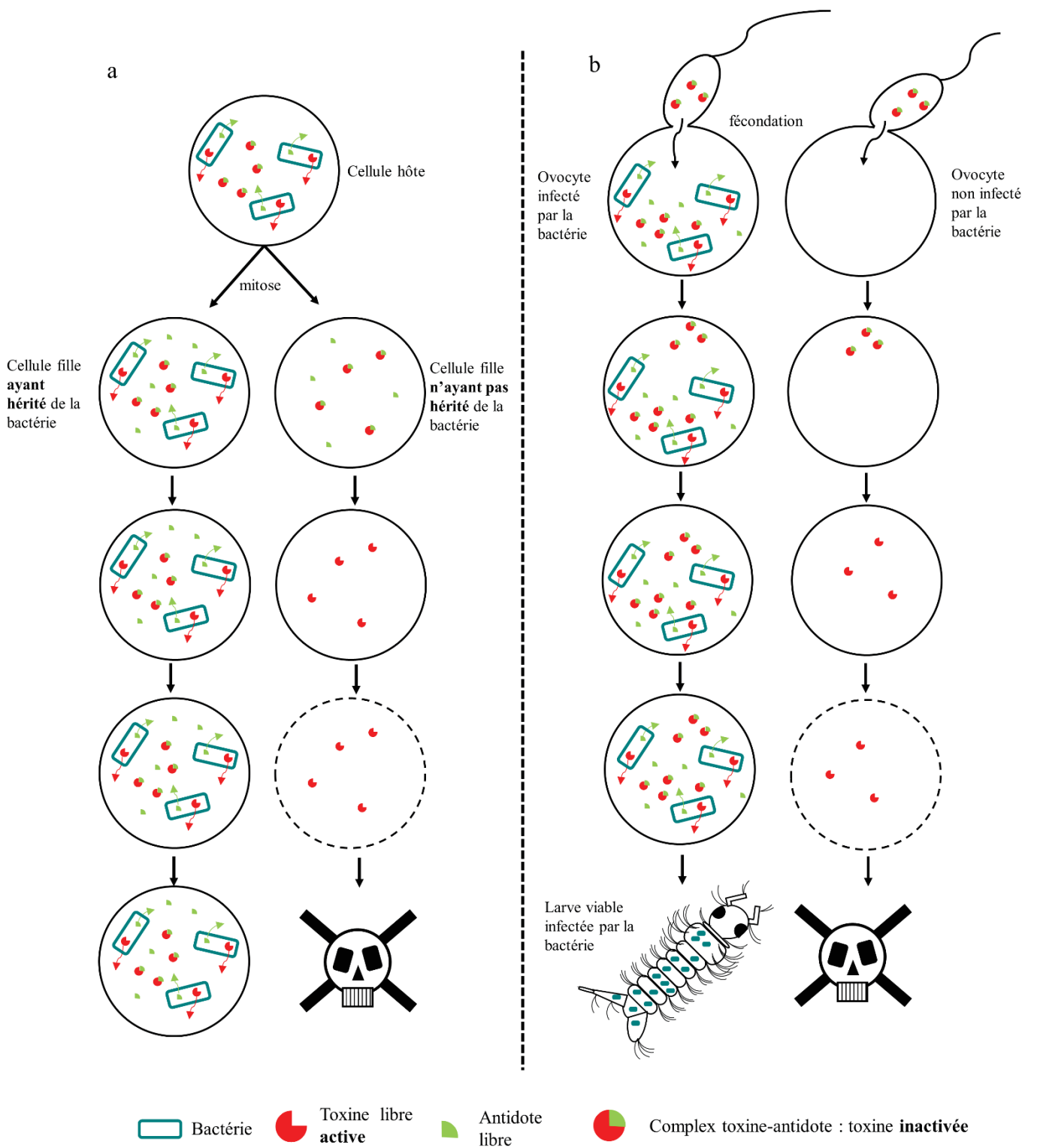


Figure 1 : Schéma explicatif de la théorie de Denis Poinot sur l'origine de l'IC.

a : Conséquences de l'absence d'une ségrégation parfaite d'une bactérie productrice d'une toxine et de son antidote (gènes d'addiction) dans une cellule hôte.

b : Conséquences de ce système d'addiction sur les gamètes et le développement embryonnaire.

présent chez celle-ci, en revanche elle est toxique pour l'hôte eucaryote. Si l'opéron *cidA/cidB* fonctionne effectivement comme un système toxine/antidote où *cidA* serait l'antidote de *cidB*, les cellules hôtes qui élimineraient les *Wolbachia* se suicideraient du fait de la présence, dans leur cytoplasme, de poison (hérité de leur « mère ») sans possibilité de produire l'antidote. Les lignées cellulaires hôtes conservant les endosymbiotes et donc les bactériophages qu'ils contiennent seraient par conséquent sélectionnées. Ce système permettrait à la fois le maintien du bactériophage dans l'endosymbiote et de l'endosymbiote dans les cellules-hôtes. L'apparition de l'IC étant secondaire comme explicité précédemment.

c. Les phages impliqués dans la diversité de l'IC : le rôle des transferts horizontaux

L'analyse du polymorphisme qui différencie les différents variants *cidA* et *cidB* entre les souches de *wPips* séquencées a révélé que ces variants sont en fait des combinaisons de blocs (Partie I, Chapitre I : Bonneau *et al.*, 2018a). La majorité des blocs sont partagés entre les variants des différentes souches appartenant aux différents groupes de *wPips* alors que certains sont spécifiques à un groupe. Une telle architecture de polymorphisme ainsi que l'incongruence entre la phylogénie des gènes *cidA* et *cidB* et celle des groupes de *wPip* suggèrent que des événements de recombinaison ont eu lieu entre les variants au sein d'un même génome de *wPip* mais également entre différents génomes.

Les mécanismes conduisant à de telles recombinaisons ne sont pas encore clairs mais l'hypothèse la plus probable fait intervenir les prophages WO. En effet, les gènes *cidA* et *cidB* que ce soit chez *wPip*, *wMel*, *wVitA*, *wSuzi*, ou encore *wBol* sont localisés au sein d'un prophage WO et plus spécifiquement dans la région EAM (eukaryotic association module) (Klasson *et al.*, 2008; Bordenstein & Bordenstein, 2016; LePage *et al.*, 2017; Lindsey *et al.*, 2018). Les prophages WO se propageraient donc d'une souche de *Wolbachia* à une autre, appartenant parfois même à des super-groupes différents via des transferts horizontaux intra et inter hôtes, et seraient impliqués dans des transferts horizontaux de gènes procaryotes et eucaryotes entre les souches de *Wolbachia* (Canchaya *et al.*, 2003; Bordenstein & Wernegreen, 2004; Gavotte *et al.*, 2004; Chafee *et al.*, 2010; Kent *et al.*, 2011a; b; Wang *et al.*, 2016a; b). Ainsi, il a été mis en évidence que des transferts horizontaux de phages ont eu lieu entre les souches *wVitA* et *wVitB* qui co-infectent *N. vitripennis* et que lors de ces transferts, des régions flanquantes de ces prophages contenant des gènes de *Wolbachia* ont été également transférées (Kent *et al.*, 2011b). De plus, si des gènes procaryotes ou eucaryotes sont localisés à l'intérieur du génome du phage, ils peuvent être transmis horizontalement via celui-ci (Bordenstein & Bordenstein, 2016; Wang *et al.*, 2016a). Ces transferts horizontaux seraient possible lorsque que les prophages insérés dans les génomes des *Wolbachia* produisent des particules de phages actives induisant la lyse des *Wolbachia* qui vont ensuite s'insérer dans les génomes d'autres *Wolbachia* (Bordenstein & Wernegreen, 2004; Bordenstein &

Reznikoff, 2005; Chafee *et al.*, 2010). De telles particules phagiques ont été observées en microscopie électronique chez *C. pipiens* (Wright *et al.*, 1978; Sanogo & Dobson, 2006), ce qui suggère que de tels transferts horizontaux sont possibles chez cette espèce. Cependant, ce n'est pas parce que les gènes *cidA* et *cidB* sont localisés dans une région de prophage qu'ils vont être présents dans les particules virales actives. En effet, lors de la réplication de l'ADN nécessaire à la formation des particules virales, ces gènes peuvent ne pas être encapsidés avec le reste du génome du phage. Les analyses des motifs protéiques des séquences des gènes *cidA* et *cidB* indiquent que ces gènes ne sont pas impliqués ni dans les gènes de « structure » du phage ni dans les gènes codant la machinerie menant à la lyse des cellules ou à l'insertion du phage dans l'ADN de son hôte. Cela suggère que les gènes *cidA* et *cidB* ne sont pas essentiels aux phages (Beckmann & Fallon, 2013; LePage *et al.*, 2017; Lindsey *et al.*, 2018). De récents travaux qui ne sont pas encore publiés, mais qui ont été communiqués à la conférence *Wolbachia* 2018 à Salem (Etats Unis), ont révélé que les particules virales produites par les souches de *Wolbachia* infectant *N. vitripennis* et *D. melanogaster* une fois purifiées et injectées dans des mâles non infectés leur confèrent la capacité d'induire de l'IC (Communication de Leigh *et al.*, à *Wolbachia* 2018). Chez ces deux espèces, ces particules virales contiendraient donc les gènes *cidA* et *cidB* ainsi que les protéines qu'ils codent.

Il est donc possible que les particules phagiques jouent un rôle dans la création de nouveau variant des gènes *cidA* et *cidB*. Afin de tester cette hypothèse expérimentalement, il sera nécessaire de purifier les particules virales des phages WO issues de différentes lignées de *C. pipiens*. Tout d'abord pour déterminer si elles contiennent les gènes *cidA* et *cidB*. Deuxièmement pour voir si la diversité des variants de *cidA* et *cidB* portée par ces phages est similaire en comparaison aux répertoires acquis sur ces mêmes lignées ou si, par exemple, des événements de recombinaison se produisent également au sein des phages. Il serait également possible de déterminer si ces particules sont capables d'induire de l'IC ou de la contrer en les injectant dans des mâles ou des femelles non infectés qui seraient par la suite croisés avec des individus qui leur sont normalement incompatibles. Enfin, nous pourrions également injecter ces phages purifiés dans un individu (au stade précoce de l'embryogenèse peu après l'oviposition) déjà infecté par une souche de *wPip* dont le répertoire est connu pour savoir si après quelques générations, on observe l'apparition de nouveaux variants chez cette souche de *wPip*.

d. Les phages, à l'origine de l'IC si particulière de *C. pipiens* ?

Culex pipiens fait figure d'exception parmi toutes les espèces chez lesquelles *Wolbachia* induit l'IC du fait de l'incroyable diversité de phénotypes d'IC décrite chez cette espèce. Cette diversité de phénotypes d'IC repose uniquement sur la présence d'une multitude de souches de *wPips* infectant les individus de cette espèce. C'est la présence dans les génomes de ces *wPips* de plusieurs copies polymorphes des gènes *cidA* et *cidB* qui sont responsables (du moins en partie) de cette diversité de phénotypes d'IC

(Partie I, Chapitre I: Bonneau *et al.*, 2018a & Chapitre II: Bonneau *et al.*, 2018b) . Une question demeure : pourquoi chez *C. pipiens* et uniquement chez *C. pipiens* retrouvons nous cette diversité si incroyable de phénotypes d'IC ?

Il est possible que *C. pipiens* ne soit pas la seule espèce chez laquelle une telle diversité existe. En effet, du fait de sa capacité vectorielle d'agents de maladies humaines, elle a attiré l'attention et a été très étudiée sur de nombreux aspects. Il est donc possible que d'autres espèces présentent une diversité de phénotype d'IC aussi grande que celle observée chez *C. pipiens*, mais qu'elles n'aient pas été trouvées.

A l'heure actuelle nous ne pouvons que proposer des hypothèses sur l'origine de cette incroyable diversité de phénotype d'IC chez *C. pipiens*. Il est par exemple possible que les souches de *wPips* infectant *C. pipiens* évoluent plus rapidement que les autres souches de *Wolbachia* infectant par exemple les *Drosophiles* ou les *Nasonia*. Ceci pourrait être dû à la présence d'éléments génétiques mobiles qui sont certes présents en grand nombre dans tous les génomes des souches de *Wolbachia* parasites du sexe mais qui pourraient être particulièrement actifs chez les *wPips*. Les prophages WO constituent de bons candidats à cette évolution rapide. En effet, comme nous l'avons discuté auparavant, grâce à leur cycle lytique, ils sont probablement en partie responsables de la diversification des variants de *cidA* et *cidB* via des transferts horizontaux. Si les phages WO présents dans les génomes des *wPips* ont des cycles lytiques plus fréquents chez *C. pipiens* que chez les autres espèces, les transferts horizontaux entre *wPips* pourraient avoir lieu plus fréquemment. De plus, du fait de la grande densité des larves de *C. pipiens*, en milieu aquatique, la transmission horizontale de ces phages pourrait être facilitée par le milieu de vie et les contacts répétés entre individus. Ce mode de vie aquatique est un facteur écologique majeur qui diffère entre les moustiques, les drosophiles et les guêpes parasitoïdes.

II. Lien entre diversité des variants *cidA-cidB* et diversité des phénotypes d'IC chez *C. pipiens*

a. A l'échelle cellulaire

Les variants multiples et polymorphes de *cidA* et *cidB* dans les génomes des différentes *wPips* que nous avons séquencés apparaissent impliqués dans la diversité des phénotypes d'IC observée chez *C. pipiens*. Une de nos hypothèses, au début de mon travail de thèse, était que cette diversité de phénotypes d'IC pouvait reposer sur une diversité de défauts cellulaires entraînant la mort des embryons dans les croisements incompatibles. Cette diversité de défauts cellulaires d'IC pourrait résulter, entre autres, de la présence de plusieurs variants de *cidA* et *cidB* dans les génomes des *wPips*. Grâce à notre étude des premières étapes du développement d'embryons de *C. pipiens* issus de croisements incompatibles, nous

avons observé des défauts cellulaires récurrents, à savoir un problème de condensation et de ségrégation de la chromatine paternelle, quels que soient les croisements incompatibles réalisés et les *wPips* impliquées (Partie II: Bonneau *et al.*, 2018c). Il apparaît donc que la diversité des phénotypes d'IC n'est pas basée sur une diversité de défauts cellulaires. En revanche, il est possible que les différents variants *cidA* et *cidB* observés dans les génomes de *wPips* induisent des défauts cellulaires similaires via des voies moléculaires différentes. Cependant, cela semble très peu probable car la zone catalytique DUB est conservée chez tous les variants *cidB*. Par conséquent, la diversité de phénotypes d'IC induite par *cidA* et *cidB* se joue à un autre niveau.

L'hypothèse que nous proposons sur la manière dont les différents variants *cidA* et *cidB* induisent la diversité des phénotypes d'IC fait appel, dans l'œuf, à des interactions spécifiques (affinités différentielles) entre les différentes protéines CidB produites par les *Wolbachia* paternelles et les protéines CidA produites par les *Wolbachia* maternelles. En effet, il a été montré que les protéines CidA et CidB interagissent spécifiquement l'une avec l'autre (Beckmann *et al.*, 2017). Ainsi, nous proposons le modèle suivant : si les *wPips* maternelles dans l'ovocyte possèdent dans leurs répertoires des variants codant des protéines CidA capables de se lier et d'empêcher ainsi la toxicité des protéines CidB issues des variants contenus dans les répertoires des *wPips* paternelles, alors le développement embryonnaire se poursuivra normalement (*i.e.* croisement compatible). En revanche, si dans l'œuf les variants de protéines CidA ne sont pas capables de se lier suffisamment intimement avec les variants de protéines CidB, alors la toxicité de CidB ne sera pas contrée et la chromatine paternelle sera impactée (*i.e.* croisement incompatible). L'analyse des séquences des variants *cidB* du groupe *wPipIV* a permis la mise en évidence d'une zone polymorphique associée au phénotype *mod* d'IC (zone 1/2/3) qui est localisée sur une portion d'environ 30 acides aminés. Sur les variants *cidA* du groupe *wPipIV*, la zone de polymorphisme potentiellement associée au phénotype *resc* d'IC (zone alpha, beta, gamma, delta, epsilon) est aussi localisée sur une zone d'environ 30 acides aminés. L'analyse des séquences d'acides aminés de ces zones a montré qu'elles présentent une homologie avec des motifs ankyrines ou « heat repeat » qui sont des motifs impliqués dans l'interaction protéine-protéine. La présence de ces domaines conforte l'hypothèse d'une interaction spécifique au niveau de ces zones polymorphes entre variants de *cidA* et de *cidB* du groupe IV.

Afin de tester cette hypothèse d'interaction spécifique entre différents variants, nous sommes en train de mettre au point, avec Frédéric Landmann et Kevin Terretaz du CRBM à Montpellier, dans le cadre du programme ANR CIAWOL que nous avons en commun, des expériences de Pull-Down. Ces expériences ont pour but d'étudier l'interaction entre deux protéines en fixant dans un ligand une protéine (par exemple CidA) qui sera mise en contact avec une solution purifiée de l'autre protéine (par exemple CidB) et leur interaction sera révélée par migration sur gel d'acrylamide. L'expérience consisterait donc à tester la liaison entre différents variants de *cidA* et de *cidB* que nous pensons pouvoir spécifiquement interagir ensemble et d'autres que nous estimons ne pas pouvoir interagir ensemble du

fait des relations d'incompatibilité des souches de *wPips* dont ils sont issus. Par ailleurs, des expériences de co-cristallisation de protéines CidA et CidB en complexe permettraient de caractériser plus précisément la zone d'interaction. Ces expériences seraient en cours (communication personnelle congrès *Wolbachia* 2018). Une fois les structures publiées, nous pourrions vérifier nos hypothèses de localisation de la zone d'interaction entre CidA et CidB puis modéliser les structures des protéines codées par nos différents variants pour prédire les affinités d'interaction entre les différents variants et ainsi potentiellement prédire les phénotypes d'IC.

b. Prédire et maîtriser les phénotypes d'IC : des analyses à différentes échelles

Depuis la découverte de la diversité des phénotypes d'IC induite par *Wolbachia* chez *C. pipiens*, l'enjeu a été de déterminer les gènes responsables de cette diversité mais aussi de pouvoir associer ces gènes (ou variants de ces gènes) à des phénotypes d'IC précis. Dans cette thèse, nous avons pu mettre en évidence une diversité de variants *cidA* et *cidB* et déterminer le répertoire de variants de nombreuses souches de *wPips* induisant des phénotypes d'IC différents (Partie I, Chapitre I: Bonneau *et al.*, 2018a & Chapitre II: Bonneau *et al.*, 2018b). La mise en relation entre les variants de ces souches et les phénotypes d'IC qu'elles induisent constitue un grand puzzle à reconstituer afin de connaître le rôle exact de chaque variant dans les différents phénotypes d'IC décrits. Cependant, l'association de phénotypes d'IC avec des variants spécifiques de *cidA* et *cidB* sur l'ensemble des différentes *wPips* infectant nos différentes lignées de *C. pipiens* du laboratoire constitue une matrice trop vaste et trop complexe pour mener à bien cette étude directement. Ainsi, il apparaît que pour pouvoir associer des variants précis de *cidA* et *cidB* à des phénotypes d'IC spécifiques il est nécessaire de commencer par analyser une sous-matrice de ce jeu de données.

C'est ce que nous avons fait en réduisant notre matrice au groupe *wPipIV* dont les mâles de différentes lignées sont infectés par des *wPipIV* étant soit « compatible » soit « incompatible » lorsque croisées avec des femelles infectées par des *wPipI*, II ou III. La présence de deux phénotypes *mod* différents entre des souches de *wPipIV* très proches génétiquement constituait un jeu de données pertinent pour commencer cette étude d'association entre variants et phénotypes d'IC. Grâce à l'étude restreinte de souches de *wPipIV* issues de populations naturelles de Tunisie, d'Algérie, de Chine, de Turquie et d'Italie, il a été possible d'associer le variant *cidB_IV(2)* avec le phénotype *mod* « incompatible » des mâles (Partie I, Chapitre I: Bonneau *et al.*, 2018a & Chapitre II: Bonneau *et al.*, 2018b). Nous sommes en train de mener une étude similaire avec des souches de *Wolbachia* du groupe *wPipI* de l'île de la Réunion dont certains mâles sont compatibles avec les femelles de la souche de *wPipIV* infectant la lignée Istanbul quand d'autres sont incompatibles (Atyame *et al.*, 2011b).

Nos études, dans différentes localités, des populations naturelles de *C. pipiens* infectées par des *wPipIV* ne permettaient pas de trouver un potentiel variant associé à une variation du phénotype *resc*

car ces lignées montraient toutes le même phénotype *resc*. Cependant, la découverte d'un variant *cidA* [*cidA*_IV(α)] ubiquitaire chez les *wPip*IV provenant de ces différentes régions du monde pourrait être la clef de leur compatibilité réciproque. En effet, ce variant *cidA*_IV(α) est le seul variant en commun à toutes les *wPip*IV, les autres variants trouvés dans ce groupe étant plutôt associés à des régions géographiques (*cidA*_IV(β) en Tunisie et Algérie, *cidA*_IV(γ) en Turquie et en Chine et *cidA*_IV(ϵ) en Italie). Il nous faudra donc tester la capacité de ce variant à « rescuer » les différents variants *cidB* du groupe IV en transgénése chez des drosophiles (*cf.* expérience de Shropshire *et al.*, 2018). Nous formulons l'hypothèse que *cidA*_IV(α) puisse être un « passe partout » des souches *wPip*IV, c'est-à-dire que ce variant serait un « antidote à large spectre » qui protège l'œuf de toutes les protéines CidB du groupe IV.

La première stratégie que nous avons appliquée sur le groupe *wPip*IV et que nous sommes en train d'appliquer à nouveau sur le groupe *wPip*I consiste à étudier des souches de *wPip* très proches génétiquement (*i.e.* avec des répertoires en variants *cidA* et *cidB* peu divergents) mais présentant deux phénotypes *mod* différents. Ainsi, on s'attend à ce que des *wPip*IV présentant deux *mod* différents aient des répertoires *cidB* différents. Il est aussi possible d'aborder le problème sous un angle plus large, à une échelle intergroupe. En effet, on pourrait chercher des variants *cidA* ou *cidB* en commun entre des souches de *wPips* beaucoup plus divergentes et responsables d'un phénotype d'IC commun. Pour le moment, dans notre jeu de données, uniquement les lignées de *C. pipiens* Lavar (*wPip*II) et Slab (*wPip*III) présentent un variant (*cidA*) dont la séquence d'acides aminés est identique. Puisque *cidA* est probablement impliqué dans la diversité des phénotypes *resc*, nous avons comparé les phénotypes *resc* de ces deux lignées (*i.e.* s'intéresser aux femelles). Les femelles de ces deux lignées présentent systématiquement des phénotypes *resc* opposés quand croisées avec la majorité des lignées du laboratoire (l'une est compatible et l'autre est incompatible), sauf avec les mâles de la lignée Tunis avec lesquels elles sont toutes les deux incompatibles et les mâles de la lignée Utique avec lesquels elles sont compatibles. Ainsi, ce variant *cidA* en commun pourrait permettre aux femelles de ces deux lignées de restaurer la fertilité quand croisées avec des mâles de la lignée Utique.

Quelle que soit la manière d'aborder cette question, il est nécessaire de se concentrer sur une matrice réduite de phénotypes d'IC et de construire une plus grande matrice pas à pas. En revanche, pour pouvoir associer spécifiquement un variant *cidA* à un phénotype *resc* et un variant *cidB* à un phénotype *mod*, il est nécessaire que nous soyons sûrs de la complétude de nos répertoires de variants *cidA* et *cidB* (*i.e.* avoir une connaissance exhaustive de tous les variants). Le clonage et le séquençage Sanger, de 24 clones minimum par gènes (48 clones pas souches de *Wolbachia*) ont été réalisés pour produire les répertoires présentés tout au long de cette thèse (Partie I et Partie II). Pour certaines souches, lorsque qu'un variant n'était observé que sur un seul clone, des clones supplémentaires (jusqu'à 72 clones par gène) ont été séquencés. Lorsqu'un variant apparaissait très majoritairement dans les clones, pouvant nous empêcher l'accès aux variants plus minoritaires, une PCR spécifique de ce variant a été développée

afin de trier et de ne séquencer que les clones ne le portant pas. Grâce à cette méthode, nous avons pu obtenir une meilleure représentation des répertoires *cidA/cidB* de nos souches. Il est clair que l'acquisition des répertoires par clonage et séquençage Sanger est laborieuse (96 séquences minimum nécessaires). De plus, on ne peut exclure d'avoir manqué la détection d'un variant rare. Pour optimiser la reconstitution des répertoires, nous sommes en train de mettre au point avec Erick Desmarais, de la Plateforme GenSeq, une méthode de séquençage NGS de ces deux gènes i) beaucoup plus rapide, ii) avec une profondeur de couverture suffisamment importante pour obtenir une liste plus exhaustive des variants constituant nos répertoires et iii) permettant potentiellement d'obtenir le répertoire de 96 souches en une seule fois. Cette méthode nous permettrait de confirmer nos répertoires déjà acquis et d'en acquérir de nouveaux afin de construire de nouvelles matrices variants *cidA-cidB* versus phénotypes d'IC.

c. Nos données soutiennent l'hypothèse d'un système toxine-antidote

L'opéron *cidA/cidB* a été proposé par Beckmann et al. (2013, 2017) comme un exemple de système toxine-antidote (TA) dans lequel *cidA* serait l'antidote et *cidB* la toxine. En effet, l'opéron *cidA/cidB* revêt plusieurs caractéristiques observées dans un système TA, plus particulièrement les systèmes TA de type II (Gerdes & Maisonneuve, 2012). Premièrement, un système TA est composé de deux gènes avec le gène de l'antidote disposé en amont de celui de la toxine, ce qui est le cas pour *cidA/cidB* où ces deux gènes sont toujours retrouvés en paire dans les génomes des souches de *Wolbachia* séquencés (Beckmann *et al.*, 2017; Lindsey *et al.*, 2018). Deuxièmement, les deux protéines produites par un tel système interagissent spécifiquement l'une avec l'autre, ce qui a été démontré pour *cidA* et *cidB* grâce à des expériences de Pull-Down (Beckmann *et al.*, 2017). Troisièmement, l'antidote est toujours plus fortement exprimé que la toxine, ce qui a été observé chez *D. melanogaster* (LePage *et al.*, 2017; Lindsey *et al.*, 2018). Enfin, Beckmann et al. (2017) n'ont pas réussi à obtenir de *D. melanogaster* contenant uniquement le gène *cidB*^{wPip} par transgénèse, ce qui suggère un effet toxique de cette protéine sur l'hôte (Beckmann et al. 2017). De la même manière, l'expression de *cidB* chez la levure est létale, ce qui conforte le rôle de toxine de *cidB*. Inversement, l'expression de *cidA* chez la levure prévient de l'activité toxique de *cidB*, ce qui supporte son rôle d'antidote (Beckmann *et al.*, 2017). Ces résultats, associés à l'incapacité de *cidB* à induire l'IC lorsque sa fonction catalytique DUB est inactivée, ont conduit Beckmann et al. (2017) à proposer le modèle simple et parcimonieux suivant : *cidB* serait la toxine impliquée dans la fonction *mod* et *cidA* serait à la fois l'antidote de *cidB* qui empêcherait *cidB* d'être toxique dans l'hôte mais aussi l'antidote de la fonction *resc* qui permettrait de restaurer la fertilité.

Plusieurs résultats obtenus au cours de ma thèse soutiennent l'hypothèse de John Beckmann, à savoir que les gènes *cidA* et *cidB* forment un opéron de type TA et que *cidA* est impliqué dans le *resc* et *cidB* dans le *mod*. Le nombre de copies de *cidA* par rapport à celui de *cidB* par génome de wPip donne toujours

une valeur proche de 1, ce qui suggère que ces gènes sont présents en même nombre de copies par génome et forment très probablement un opéron. De plus, *cidA* a toujours été trouvé plus fortement exprimé que *cidB* chez toutes les *wPips* analysées. En revanche, contrairement à *D. melanogaster* où *cidA* est environ 100 fois plus exprimé que *cidB*, chez les *wPips* *cidA* est exprimé environ 1,5-2 fois plus que *cidB*. Cette différence entre *C. pipiens* et *D. melanogaster* pourrait provenir du fait que l'expression de *cidA* et *cidB* a été réalisée dans des pools de testicules chez *D. melanogaster* alors que nous l'avons mesurée dans le corps entier chez *C. pipiens*. Les spermatozoïdes ne contiennent pas de *Wolbachia* pour produire de nouvelles protéines CidA. Les CidA qui seraient des protéines instables (forte dégradation) du fait de leur rôle probable d'antitoxine, seraient régulièrement dégradées dans le sperme. Ainsi, afin que la toxicité de CidB dans le sperme ne soit pas révélée et entraîne la mort des spermatozoïdes, l'introduction de beaucoup de protéines CidA dans ceux-ci apparaît comme essentielle. Une forte production de protéines CidA pourrait résulter d'une forte expression dans les testicules, expression qui serait par conséquent spécifiquement plus forte dans les testicules que dans le reste du corps. Nous avons tenté d'étudier l'expression de ces deux gènes dans des testicules de *C. pipiens*, mais sans réussir à obtenir suffisamment d'ARNm. Il faudra donc réitérer l'expérience en essayant d'obtenir plus d'ARNm, soit en partant de plus d'organes collectés, soit en améliorant la technique d'extraction.

Nos analyses de répertoires de variants de populations naturelles de *C. pipiens* infectés par des *wPipIV* soutiennent également le modèle simple proposé par John Beckmann, à savoir que *cidA* est impliqué dans le *resc* et *cidB* dans le *mod*. Ces résultats sont, par ailleurs, en discordance avec le modèle « deux-pour-un » proposé par Lepage et al. (2017) et Shropshire et al. (2018) chez *D. melanogaster*. En effet, dans ce modèle, *cidB* serait impliqué dans le *mod* quand *cidA* serait impliqué à la fois dans la fonction *mod* quand exprimé chez les mâles et dans la fonction *resc* quand exprimé chez les femelles. L'opéron *cidA/cidB* ne fonctionnerait donc pas exactement de la même manière chez ces deux espèces. Cette constatation est étonnante même s'il est vrai que l'IC chez *C. pipiens* et *D. melanogaster* diffère sur plusieurs aspects (diversité de phénotype, pénétrance ...). L'hypothèse de l'implication de *cidA* dans la fonction *mod* repose sur les résultats de deux expériences de transgénèse chez *D. melanogaster* (LePage et al., 2017). Premièrement, les auteurs montrent que *cidA* serait nécessaire à l'induction de l'IC car des mâles transgéniques portant *cidA^{wMel}* ou *cidB^{wMel}* ne sont pas capables d'induire l'IC alors que ceux portant *cidA^{wMel}* et *cidB^{wMel}* le sont. Deuxièmement, les auteurs montrent que *cidA* aurait un effet additif sur la pénétrance de l'IC car des mâles infectés par *Wolbachia* dans lesquels le gène *cidA^{wMel}* est exprimé par transgénèse, induisent une IC plus forte que les mâles uniquement infectés par *Wolbachia* sans le transgène. Ces résultats sont critiquables pour plusieurs raisons. Premièrement, comme nous le proposons dans un article de synthèse en cours de préparation (Beckmann, Bonneau et al., in prep Trends in Genetics), le fait que l'induction de l'IC nécessite la co-expression de *cidA* et de *cidB* ne démontre pas l'implication de *cidA* dans le *mod*. Lorsque *cidB* est exprimé seul, l'absence de phénotype d'IC pourrait très bien résulter de la mort de tous les spermatozoïdes contenant le transgène.

Ainsi, seuls les spermatozoïdes ne contenant pas la toxine survivraient et iraient féconder les ovocytes, résultant en un croisement tout à fait compatible. En revanche, si *cidA*^{wMel} et *cidB*^{wMel} sont tous les deux exprimés, leurs protéines sont toutes les deux présentes dans les spermatozoïdes. Ainsi, *CidA*^{wMel} pourrait prévenir de l'effet toxique de *CidB*^{wMel} sur les spermatozoïdes jusqu'à la fécondation qui induira l'IC. La capacité de *CidB*^{wMel} à tuer les spermatozoïdes est soutenue par l'effet très toxique de *CidB*^{wPip} lorsqu'exprimé tout seul en transgénése chez *D. melanogaster* (aucun individu viable n'a été obtenu, Beckmann *et al.*, 2017). L'effet toxique de *CidB*^{wPip} ne serait pas restreint à la lignée germinale contrairement à *CidB*^{wMel}, potentiellement à cause de différences physiologiques entre *D. melanogaster* et *C. pipiens* qui rendraient *CidB*^{wPip} plus toxique chez *D. melanogaster*.

III. Sur quel déterminisme génétique est basée la diversité de l'IC chez les drosophiles ?

Alors que *D. melanogaster* n'est infectée que par une seule souche de *Wolbachia* (*wMel*), *D. simulans*, peut être infectée par les cinq souches suivantes : *wRi*, *wHa*, *wNo*, *wAu* et *wMa* (voir Merçot & Charlat, 2004 pour revue). Chez *D. simulans*, certaines souches ne sont pas capables d'induire l'IC, ce qui n'a jamais été observé chez *C. pipiens*. Ainsi, la souche *wMa* n'induit pas d'IC même lorsque les mâles sont croisés avec des femelles non infectées (James & Ballard, 2000; Charlat *et al.*, 2003b). En revanche, il a été montré dans une population au pied du Kilimandjaro qu'elle restaurait l'IC de *wNo* dont elle est très proche génétiquement (Merçot & Poinot, 1998; Poinot & Merçot, 1999; Charlat *et al.*, 2002). Cela veut dire que si cette souche ne doit pas coder de facteur *mod* (ou toxine) dans son génome, elle doit en revanche coder un facteur *resc* (ou antidote). Cette souche présente un opéron *cinA/cinB* dont le *cinA* semble fonctionnel alors que le *cinB* comporte un codon stop (Beckmann, Bonneau *et al.*, in prep). La souche *wAu* n'est, elle non plus, pas capable d'induire l'IC même quand les mâles sont croisés avec des femelles non infectées et est aussi incapable de restaurer la fertilité (Hoffmann *et al.*, 1996; James & Ballard, 2000; Reynolds & Hoffmann, 2002; Charlat *et al.*, 2003b). Cette souche ne possède aucun homologue de *cidA* et *cidB* dans son génome (Lindsey *et al.*, 2018).

Les autres souches infectant *D. simulans*, *wRi*, *wHa*, *wNo*, sont bidirectionnellement incompatibles entre elles. Lorsque *D. simulans* est transinfectée par *wMel*, les croisements avec des individus infectés soit par *wHa* soit par *wNo* sont bidirectionnellement incompatibles. En revanche, les femelles infectées par *wRi* sont capables de restaurer la fertilité lorsque croisées avec des mâles infectés par *wMel* (IC unidirectionnelle, Poinot *et al.*, 1998). Une première hypothèse pour expliquer que ces souches soient bidirectionnellement et unidirectionnellement incompatibles serait que comme chez *C. pipiens*, chacune de ces souches possède des variants de *cidA* et/ou de *cidB* différents. L'analyse des génomes de ces

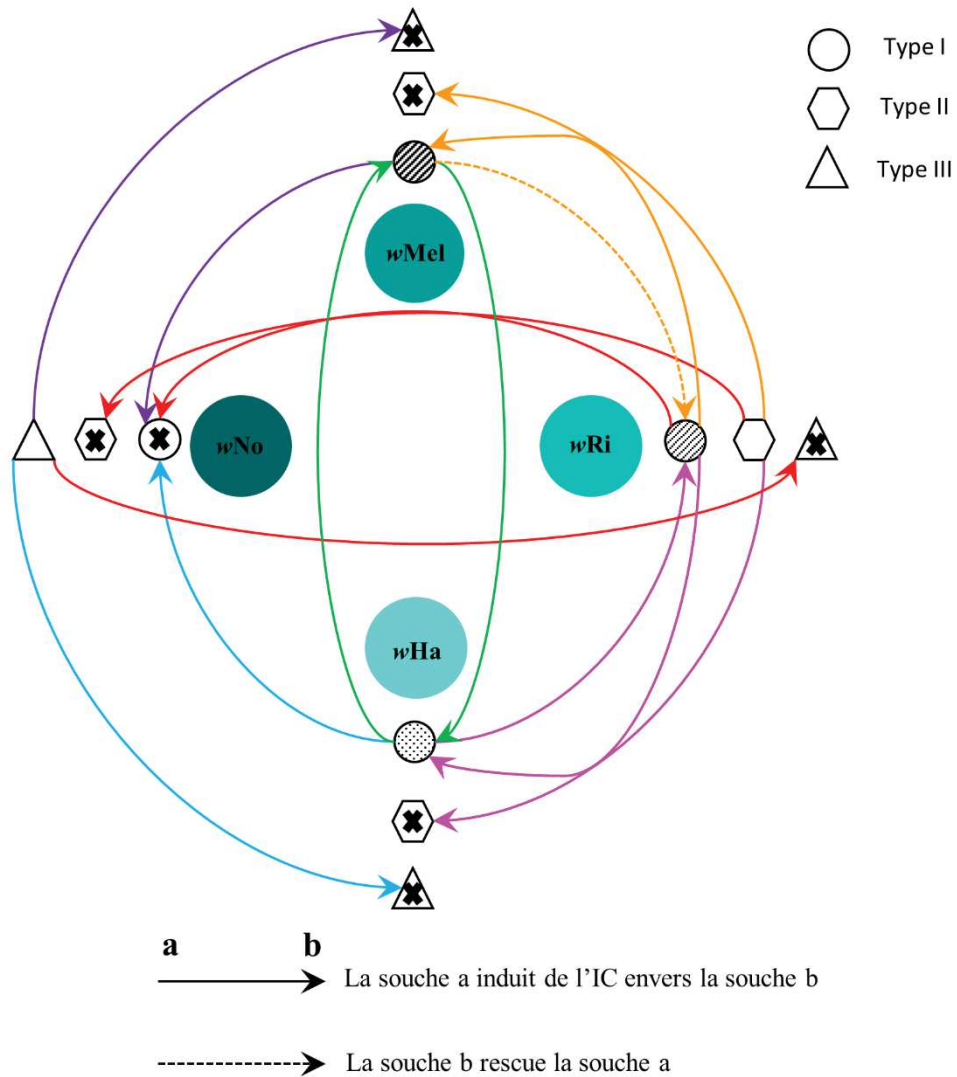


Figure 2 : Relations d'incompatibilité entre les souches de *Wolbachia* wMel, wNo, wHa et wRi en lien avec les types d'opérons qu'elles possèdent dans leur génome.

Type I: DUB; **Type II, III:** Nucléases (LePage *et al.*, 2017; Lindsey *et al.*, 2018).

Les croix dans les symboles indiquent l'absence de ce type d'opéron dans le génome de la souche de *Wolbachia*.

Les motifs à l'intérieur des cercles figurant les types I indiquent le degré d'identité des séquences d'acides aminés entre les types I des différentes souches. Plus le motif est proche, plus le pourcentage d'identité est élevé. Par exemple, les types I de wMel et wRi possèdent 99% d'identité alors que les types I de wMel et wHa possèdent seulement 65% en moyenne

souches a révélé que *wMel* et *wHa* possèdent bien une seule copie des gènes *cidA* et *cidB* dont le pourcentage d'identité entre les deux souches est d'environ de 65%. *wRi* possède quant à elle deux copies identiques de cet opéron dont les deux gènes possèdent 99% d'identité avec ceux de *wMel* et 68% (*cidA*) et 62% (*cidB*) d'identité avec ceux de *wHa*. Enfin, *wNo* ne possède pas de gènes *cidA* et *cidB* (LePage *et al.*, 2017). Cette faible diversité de séquence entre *wRi* et *wMel* et l'absence d'opéron *cidA/cidB* chez *wNo* nécessitent de faire appel à d'autres gènes, comme les paralogues *cinA/cinB*, afin d'expliquer la diversité des phénotypes d'IC observés chez les drosophiles (LePage *et al.*, 2017).

C'est pourquoi, dans Lepage *et al.* (2017) et Lindsey *et al.* (2018), les auteurs se sont lancés dans la caractérisation des différents paralogues de l'opéron *cidA/cidB* dans les génomes des *Wolbachia* infectant les drosophiles. Ils ont caractérisé en plus du type I (l'opéron possédant une DUB *cidA/cidB*), trois types d'opéron dont le deuxième gène possède non pas une DUB mais une nucléase : les types II, III et IV (*cinA^{wPip}/cinB^{wPip}* étant un type IV). Les auteurs ont proposé que ces types d'opérons en fonction de leur pourcentage d'identité en acides aminés et de leur présence/absence puissent permettre d'expliquer les phénotypes d'IC observés entre *wRi*, *wHa*, *wNo* et *wMel* (Figure 2, LePage *et al.*, 2017). Ainsi, *wRi* serait capable de « rescuer » *wMel* car elles possèdent toutes les deux un opéron *cidA/cidB* avec 99% d'identité (Figure 2 couleur orange). En revanche *wMel* ne pourrait pas « rescuer » *wRi* car *wRi* possède un opéron de type II absent du génome de *wMel* (Figure 2 couleur orange). *wNo* serait bidirectionnellement incompatible avec toutes les autres souches car elle est la seule à posséder un type III (Figure 2 couleurs bleue, rouge et violette). *wHa* et *wRi* seraient bidirectionnellement incompatibles à cause i) du type II présent chez *wRi* et absent de chez *wHa* et ii) du faible pourcentage d'identité entre leurs opérons de type I *cidA/cidB* (68% et 62% d'identité pour *cidA* et *cidB* respectivement, Figure 2 couleur rose). Enfin, *wMel* et *wHa* seraient bidirectionnellement incompatibles du fait que leurs opérons de type I ne partagent que 67% et 62% d'identité pour *cidA* et *cidB* respectivement (Figure 2 couleur verte). Ce pourcentage d'identité est en effet faible. A titre de comparaison, les différents variants *cidA* et *cidB* séquencés dans nos souches de *wPips* présentent au minimum 91% d'identité de séquence d'acide aminés et sont pourtant impliqués dans la grande diversité phénotypique d'IC chez *C. pipiens*.

Chez les Drosophiles, la diversité phénotypique d'IC reposerait sur la présence ou l'absence des deux types d'opérons (DUB vs Nucléase) et/ou résulterait d'une divergence importante entre les opérons de même type. Il serait pertinent, pour étayer cette hypothèse, de séquencer et de croiser d'autres souches de *Wolbachia* de Drosophiles. De plus, afin de vérifier que les opérons nucléases sont effectivement capables d'induire l'IC et d'étudier le rôle respectif des deux gènes dans les fonctions *mod* et *resc*, il faudrait mener des expériences de transgénèse de ces opérons chez *D. melanogaster*. De ce point de vue, le cas de *wNo* est particulièrement intéressant car cette souche induit l'IC sans opéron *cidA/cidB* dans son génome, potentiellement grâce à la seule présence de l'opéron *cinA/cinB*. Il est donc nécessaire d'étudier les mécanismes cytologiques sous-jacent à l'IC entre mâles infectés par *wNo* et femelles non infectées afin de déterminer si ce sont les mêmes que ceux caractérisés dans les organismes qui

possèdent l'opéron *cidA/cidB*. Si tel est le cas, il y aurait convergence des phénotypes cellulaires de l'IC induits par ces deux types d'opérons à partir de voies moléculaires différentes.

L'opéron *cinA/cinB* est également présent chez *C. pipiens*, mais n'est pas variable ce qui suggère qu'il ne joue pas de rôle dans la diversité de l'IC. Cependant, puisqu'il est probablement impliqué dans l'induction de l'IC chez *wNo*, il pourrait également jouer un rôle dans l'IC chez *C. pipiens*. Une hypothèse serait que la présence de l'opéron *cinA/cinB* en plus de *cidA/cidB* contribuerait à la très forte pénétrance de l'IC chez *C. pipiens* (Partie II: Bonneau *et al.*, 2018c). Un argument en faveur de l'implication de *cinA/cinB* dans la pénétrance de l'IC est que *wRi*, *wPip* et *wVitA*, qui possèdent toutes les trois les deux types d'opérons, induisent des pénétrances d'IC très élevées alors que *wMel*, *wNo* et *wHa* qui ne possèdent que l'un ou l'autre n'induisent que de l'IC de faible intensité. La pénétrance d'IC partielle de ces souches pourrait résulter, soit de la présence de spermatozoïdes ne contenant pas de toxines, soit du fait que certains spermatozoïdes contiennent une trop faible dose de la toxine.

Si l'on suppose que les nucléases auraient un effet additif à celui des DUB dans la modification de la chromatine paternelle, chez *C. pipiens*, cela pourrait augmenter la quantité de toxines et la proportion des spermatozoïdes impactés par la toxine. Cela conduirait, dans les deux cas, à une augmentation de la pénétrance de l'IC. Cette théorie est cependant problématique si l'on se place du côté de la « rescue ». En effet, si l'opéron *cinA/cinB* est identique chez les différentes *wPips*, alors toutes les *wPips* produiraient la protéine CinA capable d'empêcher la protéine CinB d'être toxique. Pour que l'hypothèse de synergie des voies moléculaires fonctionne, il faudrait faire, par exemple, appel à un modèle quantitatif où on pourrait imaginer que la quantité de protéines CinA et CinB varie selon les souches et quand CinA est produit en trop faible quantité ça ne suffit pas pour contrecarrer les effets toxiques de CinB. Pour tester ces hypothèses, il faudrait comparer la pénétrance de l'IC chez des mâles *D. melanogaster* dans lesquels les opérons *cidA/cidB* et *cinA/cinB* de *wRi* auraient été intégrés en transgénèse, ensemble ou séparément. Cette expérience ne serait pas pertinente en utilisant les opérons *cidA/cidB* et *cinA/cinB* de *wPip* car *cidA^{wPip}/cidB^{wPip}* seuls induisent déjà une pénétrance de l'IC totale en transgénèse chez des mâles *D. melanogaster* (Beckmann *et al.*, 2017). En revanche, la transgénèse de *cidA^{wMel}/cidB^{wMel}* n'induit qu'une IC partielle, ce qui en fait de meilleurs candidats (LePage *et al.*, 2017).

IV. Quid de l'IC induite par d'autres endosymbiotes : une toute autre histoire ?

L'IC, bien que pratiquée par une grande diversité de souches de *Wolbachia* chez un grand nombre d'espèces d'hôtes arthropodes, n'est induite que par deux autres endosymbiotes : une autre alpha-protéobactérie récemment décrite chez le bronstipe du cocotier *Brontispa longissima* (Takano *et al.*,

2017) et chez la Bacteroidetes *Cardinium hertigii* (Hunter *et al.*, 2003). Cette dernière est capable d'induire l'IC, seule, chez l'acarien *Eotetranychus suginamensis* (Gotoh *et al.*, 2007) et la guêpe parasitoïde *Encarsia suzannae* (Hunter *et al.*, 2003). Cette bactérie a été détectée en coïnfection avec *Wolbachia* chez le thrips *Pezothrips kellyanus* (Nguyen *et al.*, 2017) et chez la cicadelle à dos blanc *Sogatella furcifera* (Nakamura *et al.*, 2012) où elles induisent toutes les deux l'IC.

Les mécanismes cytologiques sous-jacents à l'IC que *Cardinium* induit chez *E. suzannae* sont extrêmement similaires à ceux induits par *Wolbachia* (*i.e.* problème de condensation et de ségrégation de la chromatine paternelle (Gebiola *et al.*, 2017). Pour le moment, les bases génétiques de l'IC n'ont pas été découvertes chez *Cardinium* mais une étude du transcriptome de la souche cEper1 qui infecte la guêpe parasitoïde *E. suzannae* a révélé la présence de gènes dont le niveau d'expression est particulièrement élevé. Ces gènes sont impliqués dans la manipulation des voies d'ubiquitination, de l'apoptose et des processus impactant l'ADN de l'hôte (Mann *et al.*, 2017). Cette souche de *Cardinium* ne présente pas d'homologues de *cidA* et *cidB*, ce qui suggère que les bases moléculaires de l'IC ne sont pas les mêmes chez *Cardinium* et *Wolbachia* (Penz *et al.*, 2012). En revanche, le génome de cette souche présente une protéine avec une potentielle fonction de ligase d'ubiquitine (CAHE_p0026) et une protéase spécifique d'ubiquitine (*i.e.* une DUB, CAHE_0028) qui sont de bons candidats pour être impliqués dans l'IC (Penz *et al.*, 2012; Mann *et al.*, 2017). Ce résultat, associé à la similarité des défauts cellulaires, suggère que les voies moléculaires différentes employées par ces bactéries pour induire l'IC aient convergé vers un phénotype de manipulation du sexe commun. Le séquençage du génome de l'alpha-protéobactérie induisant l'IC chez le bronstipe du cocotier ainsi que l'étude des mécanismes cellulaires à l'œuvre pourraient permettre de déterminer si les voies d'ubiquitination/déubiquitination sont aussi favorisées chez cette bactérie.

Récemment, il a été montré qu'une autre bactérie, *Spiroplasma poulsonii*, induit du male-killing chez *D. melanogaster* via une protéine nommé Spaid qui contient un domaine DUB (et des domaines ankyrines) (Harumoto & Lemaitre, 2018). L'expression de Spaid par transgénèse chez *D. melanogaster* produit une forte apoptose et une désorganisation neurale chez les embryons mâles similaire à ce qui est observé chez les embryons mâles infectés par *S. poulsonii*. Il est pertinent de souligner que dans les embryons mâles exprimant Spaid, de nombreux ponts chromatiniens ont été observés, ce qui rappelle fortement les mécanismes cytologiques de l'IC induit par *Wolbachia* et *Cardinium* (Gebiola *et al.*, 2017; Harumoto & Lemaitre, 2018, Partie II: Bonneau *et al.*, 2018c). Ainsi, les voies d'ubiquitination/déubiquitination utilisées par *S. poulsonii* dans le MK et *Wolbachia* voire *Cardinium* dans l'IC apparaissent comme des voies de prédilection dans l'altération du matériel génétique et la manipulation de la reproduction.

V. L'IC dans la lutte antivectorielle

Les maladies vectorielles transmises par les moustiques telles que la malaria, la dengue, le chikungunya, la fièvre de la vallée du Rift, Zika ou encore le virus West Nile sont parmi les maladies responsables des plus fortes mortalités humaines. Les premières stratégies mises en place pour contrôler les populations naturelles de moustiques utilisaient des insecticides mais leurs utilisations répétées et non optimales ont conduit à la sélection de nombreux individus résistants voire même multi-résistants. Même si l'utilisation d'insecticides continue à être l'une des stratégies de contrôle les plus employées, du fait de l'émergence de multiples résistances et de l'impact négatif de ces insecticides sur d'autres arthropodes non visés par les traitements, des solutions alternatives sont en cours de développement (voir Bourtzis *et al.*, 2016 et Flores & O'Neill, 2018 pour revue). Certaines de ces méthodes alternatives font appel à l'utilisation de *Wolbachia*. A l'heure actuelle, deux types de stratégie existent (Xi *et al.*, 2005; Brelsfoard *et al.*, 2008; Walker *et al.*, 2011; Bian *et al.*, 2013; Bourtzis *et al.*, 2014).

La première méthode utilise *Wolbachia* pour remplacer les individus d'une population de moustiques ciblée par des individus contenant une *Wolbachia* qui confère une protection des moustiques vis-à-vis des virus ou autres agents pathogènes humains (voir Bourtzis *et al.*, 2016 et Flores & O'Neill, 2018 pour revue). C'est par exemple le cas des souches *wMel*, *wMelPop-CLA*, *wAu*, *wMelCS* et *wAlbB* qui, lorsqu'elles sont transinfectées par microinjection dans *Aedes aegypti*, limitent la transmission du virus de la dengue, de Zika, ou du chikungunya (Moreira *et al.*, 2009; Bian *et al.*, 2010; Walker *et al.*, 2011; Aliota *et al.*, 2016b; a; Dutra *et al.*, 2016; Ant *et al.*, 2018). Ces mêmes effets « protecteurs » contre i) la dengue sont observés lorsque *Aedes albopictus* est transinfecté avec *wMel* (Blagrove *et al.*, 2012, 2013; Mousson *et al.*, 2012) et contre ii) la malaria lorsque *Anopheles gambiae* est transinfectée avec *wMelPop* (Kambris *et al.*, 2010; Hughes *et al.*, 2011). Afin de remplacer une population naturelle par une nouvelle population de moustiques contenant des *Wolbachia* « protectrices », des mâles et des femelles moustiques infectés par ces *Wolbachia*, sont libérés dans les populations naturelles. Il est nécessaire, au préalable, de tester l'incompatibilité cytoplasmique entre la *Wolbachia* protectrice et celle(s) déjà présente(s) dans les populations naturelles (si infectées) afin de remplacer efficacement les individus de la population ciblée. Une souche uni-directionnellement incompatible sera favorisée pour son pouvoir invasif. L'IC peut donc, si la souche est bien choisie, conférer un avantage reproductif aux femelles infectées par la *Wolbachia* « protectrice », ce qui permet théoriquement la propagation et l'établissement de celle-ci dans les populations. Une étude préliminaire a été menée en Australie pour tester l'efficacité du remplacement de populations naturelles d'*A. aegypti*, moustiques majoritairement non infectés par *Wolbachia*, par des populations infectées par *wMel* (Hoffmann *et al.*, 2011). Cette étude a montré l'efficacité de cette méthode car on observe dans les populations une prévalence de l'infection par *Wolbachia* supérieure à 90%, même plusieurs années après le lâché initial (Hoffmann *et al.*, 2014). A l'heure actuelle, plusieurs essais en populations naturelles au Brésil, en Colombie, en Indonésie et au Vietnam, où des virus et autres pathogènes humains sont présents, sont en cours dans le cadre du « World

Mosquito Program » pour tester l'efficacité de cette méthode contre la Dengue (Servick, 2016; Anders *et al.*, 2018).

La deuxième méthode de stratégie de lutte antivectorielle qui implique *Wolbachia* s'appelle la technique de l'insecte incompatible (TII). Cette technique repose une nouvelle fois sur la capacité de *Wolbachia* à induire l'IC chez ses hôtes. Dans cette méthode, des moustiques mâles infectés par une souche de *Wolbachia* (soit naturellement soit pas transinfection) sont lâchés dans la population ciblée. Soit les populations locales ne sont pas infectées et n'importe quelle souche de *Wolbachia* peut faire l'affaire comme c'est généralement le cas chez *A. aegypti*, soit les populations locales sont déjà infectées par *Wolbachia* comme c'est toujours le cas pour *A. albopictus* et dans ce cas-là il faut choisir une souche de *Wolbachia* incompatible avec la souche de *Wolbachia* locale. Dans l'idéal, la *Wolbachia* « relâchée » devra être bidirectionnellement incompatible avec la *Wolbachia* locale car, si jamais des femelles sont relâchées par mégarde (à cause de la difficulté de trier parfaitement les deux sexes), celles-ci seront incompatibles avec les mâles locaux. Ces mâles, infectés ou transinfectés incompatibles avec les femelles locales, produiront une descendance non viable, ce qui diminuera la taille de la population du moustique vecteur et potentiellement la transmission des virus et des parasites qu'il transmet. Des expériences en laboratoire (Atyame *et al.*, 2011c, 2015a; Walker *et al.*, 2011; Moretti & Calvitti, 2013) et des essais pilotes ont été réalisés sur le terrain et semi-terrain en Birmanie, en Polynésie française et dans le Kentucky avec des résultats encourageants (Laven, 1967a; Chambers *et al.*, 2011; O'Connor *et al.*, 2012; Mains *et al.*, 2016).

La découverte des gènes de *cidA* et *cidB* et leur association avec la diversité des phénotypes d'IC chez *C. pipiens* ouvrent la voie vers une utilisation optimale de l'IC comme stratégie de lutte antivectorielle. En effet, une fois qu'il sera établi plus précisément quels variants de *cidA* et de *cidB* induisent tel ou tel phénotype d'IC, il serait possible de créer des moustiques génétiquement modifiés « sur-mesure », adaptés pour remplacer ou supprimer les populations locales via par exemple l'utilisation de Crispr Cas9. Crispr Cas9 est un outil très prometteur. En effet, il a été montré que son utilisation, dans des conditions de laboratoire, permettait la propagation dans des populations d'*Anopheles stephensi*, d'un gène de résistance à *Plasmodium falciparum*, parasite responsable de la malaria (Gantz *et al.*, 2015). Cependant, l'utilisation de moustique transgénique fait débat à l'heure actuelle car elle soulève de nombreuses questions, tant éthiques (Resnik, 2014) que sur leur efficacité en population naturelles (Scott *et al.*, 2002). Par ailleurs, la diversité des phénotypes d'IC et de variants de *cidA* et *cidB* présente chez *C. pipiens* constitue une source quasiment intarissable de *Wolbachia* incompatibles à transinfecter chez d'autres insectes, ce qui permettrait une utilisation de la TII sur le long terme. Il semble ainsi primordial de continuer à étudier le fonctionnement de ces deux gènes chez *C. pipiens*, ainsi que de pouvoir prédire plus précisément quels *cidA* et quels *cidB* induisent quels phénotypes d'IC afin de construire une banque de variants/phénotypes qui constituerait un outil très précieux pour une utilisation durable et efficace de *Wolbachia* en technique de l'insecte incompatible.

Bibliographie

- Aanen, D.K., de Fine Licht, H.H., Debets, A.J.M., Kerstes, N.A.G., Hoekstra, R.F. & Boomsma, J.J. 2009. High Symbiont Relatedness Stabilizes Mutualistic Cooperation in Fungus-Growing Termites. *Science*. **326**: 1103–1106.
- Adachi-Hagimori, T., Miura, K. & Stouthamer, R. 2008. A new cytogenetic mechanism for bacterial endosymbiont-induced parthenogenesis in Hymenoptera. *Proc. R. Soc. B Biol. Sci.* **275**: 2667–2673.
- Ahmed, M.Z., Breinholt, J.W. & Kawahara, A.Y. 2016. Evidence for common horizontal transmission of *Wolbachia* among butterflies and moths. *BMC Evol. Biol.* **16**: 1–16.
- Akashi, H., Osada, N. & Ohta, T. 2012. Weak selection and protein evolution. *Genetics* **192**: 15–31.
- Akman Gündüz, E. & Douglas, A.E. 2012. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proc. R. Soc. B Biol. Sci.* **276**: 987–991.
- Aliota, M.T., Peinado, S.A., Velez, I.D. & Osorio, J.E. 2016a. The wMel strain of *Wolbachia* Reduces Transmission of Zika virus by *Aedes aegypti*. *Sci. Rep.* **6**: 1–7.
- Aliota, M.T., Walker, E.C., Uribe Yepes, A., Dario Velez, I., Christensen, B.M. & Osorio, J.E. 2016b. The wMel Strain of *Wolbachia* Reduces Transmission of Chikungunya Virus in *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **10**: e0004677.
- Alizon, S., Hurford, A., Mideo, N. & Van Baalen, M. 2009. Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *J. Evol. Biol.* **22**: 245–259.
- Alout, H., Labbé, P., Berthomieu, A., Pasteur, N. & Weill, M. 2009. Multiple duplications of the rare ace-1 mutation F290V in *Culex pipiens* natural populations. *Insect Biochem. Mol. Biol.* **39**: 884–891.
- Altinli, M., Gunay, F., Alten, B., Weill, M. & Sicard, M. 2018. *Wolbachia* diversity and cytoplasmic incompatibility patterns in *Culex pipiens* populations in Turkey. *Parasites and Vectors* **11**: 1–9.
- Anders, K.L., Indriani, C., Ahmad, R.A., Tantowijoyo, W., Arguni, E., Andari, B., *et al.* 2018. The AWED trial (Applying *Wolbachia* to Eliminate Dengue) to assess the efficacy of *Wolbachia*-infected mosquito deployments to reduce dengue incidence in Yogyakarta, Indonesia: Study protocol for a cluster randomised controlled trial. *Trials* **19**: 1–16.
- Ant, T.H., Herd, C.S., Geoghegan, V., Hoffmann, A.A. & Sinkins, S.P. 2018. The *Wolbachia* strain wAu provides highly efficient virus transmission blocking in *Aedes aegypti*. *PLoS Pathog.* **14**: 1–19.
- Arakaki, N., Miyoshi, T. & Noda, H. 2001. *Wolbachia*-mediated parthenogenesis in the predatory thrips *Franklinothrips vespiformis* (Thysanoptera: Insecta). *Proc. R. Soc. B Biol. Sci.* **268**: 1011–1016.
- Aras, R.A., Kang, J., Tschumi, A.I., Harasaki, Y. & Blaser, M.J. 2003. Extensive repetitive DNA facilitates prokaryotic genome plasticity. *Proc. Natl. Acad. Sci.* **100**: 13579–13584.
- Atanassov, B.S., Koutelou, E. & Dent, S.Y. 2011. The role of deubiquitinating enzymes in chromatin regulation. *FEBS Lett.* **585**: 2016–2023.
- Atyame, C.M., Cattel, J., Lebon, C., Flores, O., Dehecq, J.S., Weill, M., *et al.* 2015a. *Wolbachia*-based population control strategy targeting *Culex quinquefasciatus* mosquitoes proves efficient under semi-field conditions. *PLoS One* **10**.
- Atyame, C.M., Delsuc, F., Pasteur, N., Weill, M. & Duron, O. 2011a. Diversification of *Wolbachia* Endosymbiont in the *Culex pipiens* Mosquito. *Mol. Biol. Evol.* **28**: 2761–2772.

- Atyame, C.M., Duron, O., Tortosa, P., Pasteur, N., Fort, P. & Weill, M. 2011b. Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito populations. *Mol. Ecol.* **20**: 286–298.
- Atyame, C.M., Labbé, P., Dumas, E., Milesi, P., Charlat, S., Fort, P., *et al.* 2014. *Wolbachia* Divergence and the Evolution of Cytoplasmic Incompatibility in *Culex pipiens*. *PLoS One* **9**: e87336.
- Atyame, C.M., Labbé, P., Rousset, F., Beji, M., Makoundou, P., Duron, O., *et al.* 2015b. Stable coexistence of incompatible *Wolbachia* along a narrow contact zone in mosquito field populations. *Mol. Ecol.* **24**: 508–521.
- Atyame, C.M., Pasteur, N., Dumas, E., Tortosa, P., Tantely, M.L., Pocquet, N., *et al.* 2011c. Cytoplasmic Incompatibility as a Means of Controlling *Culex pipiens quinquefasciatus* Mosquito in the Islands of the South-Western Indian Ocean. *PLoS Negl. Trop. Dis.* **5**: e1440.
- Augustinos, A.A., Santos-Garcia, D., Dionyssopoulou, E., Moreira, M., Papapanagiotou, A., Scarvelakis, M., *et al.* 2011. Detection and characterization of *Wolbachia* infections in natural populations of aphids: is the hidden diversity fully unraveled? *PLoS One* **6**: e28695.
- Badawi, M., Moumen, B., Giraud, I., Grève, P. & Cordaux, R. 2018. Investigating the Molecular Genetic Basis of Cytoplasmic Sex Determination Caused by *Wolbachia* Endosymbionts in Terrestrial Isopods. *Genes (Basel)*. **9**: 290.
- Balcewicz-Sablinska, M.K., Keane, J., Kornfeld, H. & Remold, H.G. 1998. Pathogenic *Mycobacterium tuberculosis* evades apoptosis of host macrophages by release of TNF-R2, resulting in inactivation of TNF-alpha. *J. Immunol.* **161**: 2636–41.
- Baldo, L., Hotopp, J.C.D., Jolley, K. a., Bordenstein, S.R., Biber, S. a., Choudhury, R.R., *et al.* 2006. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol.* **72**: 7098–7110.
- Baldo, L., Prendini, L., Corthals, A. & Werren, J.H. 2007. *Wolbachia* are present in Southern African scorpions and cluster with supergroup F. *Curr. Microbiol.* **55**: 367–373.
- Bandi, C., Dunn, A.M., Hurst, G.D.. & Rigaud, T. 2001. Inherited microorganisms, sex-specific virulence and reproductive parasitism. *Trends Parasitol.* **17**: 88–94.
- Barr, A.R. 1957. The Distribution of *Culex P. Pipiens* and *C. P. Quinquefasciatus* in North America. *Am. J. Trop. Med. Hyg.* **6**: 153–165.
- Barr, R. 1980. Cytoplasmic incompatibility in natural populations of a mosquito, *Culex pipiens*. *Nature.* **283**:71-72
- Bateman, A.J. 1948. Intra-sexual selection in *Drosophila*. *Heredity.* **2**: 349–368.
- Beckmann, J.F. & Fallon, A.M. 2013. Detection of the *Wolbachia* protein WPIPO282 in mosquito spermathecae: Implications for cytoplasmic incompatibility. *Insect Biochem. Mol. Biol.* **43**: 867–878.
- Beckmann, J.F., Ronau, J.A. & Hochstrasser, M. 2017. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2**: 17007.
- Behar, S.M., Martin, C.J., Booty, M.G., Nishimura, T., Zhao, X., Gan, H.X., *et al.* 2011. Apoptosis is an innate defense function of macrophages against *Mycobacterium tuberculosis*. *Mucosal Immunol.* **4**: 279–287. Nature Publishing Group.
- Bell, G. 1986. The Evolution of Empty Flowers. *J. Theor. Biol.* **118**: 253–258.
- Beltran-Bech, S. & Richard, F.J. 2014. Impact of infection on mate choice. *Anim. Behav.* **90**: 159–170.

- Ben Cheikh, H., Ben Ali-Haouas, Z., Marquine, M. & Pasteur, N. 1998. Resistance to Organophosphorus and Pyrethroid Insecticides in *Culex pipiens* (Diptera: Culicidae) from Tunisia. *J. Med. Entomol.* **35**: 251–260.
- Bentley, J.K., Veneti, Z., Heraty, J. & Hurst, G.D.D. 2007. The pathology of embryo death caused by the male-killing *Spiroplasma* bacterium in *Drosophila nebulosa*. *BMC Biol.* **5**: 1–7.
- Bian, G., Joshi, D., Dong, Y., Lu, P., Zhou, G., Pan, X., *et al.* 2013. *Wolbachia* Invades *Anopheles stephensi*. *Science.* **340**: 748–751.
- Bian, G., Xu, Y., Lu, P., Xie, Y. & Xi, Z. 2010. The Endosymbiotic Bacterium *Wolbachia* Induces Resistance to Dengue Virus in *Aedes aegypti*. *PLoS Pathog.* **6**: e1000833.
- Birky, C.W. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl. Acad. Sci.* **92**: 11331–11338.
- Blagrove, M.S.C., Arias-Goeta, C., Di Genua, C., Failloux, A.B. & Sinkins, S.P. 2013. A *Wolbachia* wMel Transinfection in *Aedes albopictus* Is Not Detrimental to Host Fitness and Inhibits Chikungunya Virus. *PLoS Negl. Trop. Dis.* **7**: e2152.
- Blagrove, M.S.C., Arias-Goeta, C., Failloux, A.-B. & Sinkins, S.P. 2012. *Wolbachia* strain wMel induces cytoplasmic incompatibility and blocks dengue transmission in *Aedes albopictus*. *Proc. Natl. Acad. Sci.* **109**: 255–260.
- Bleidorn, C. & Gerth, M. 2018. A critical re-evaluation of multilocus sequence typing (MLST) efforts in *Wolbachia*. *FEMS Microbiol. Ecol.* **94**: 1997–1998.
- Bonneau, M., Atyame, C., Beji, M., Justy, F., Cohen-Gonsaud, M., Sicard, M., *et al.* 2018a. *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat. Commun.* **9**: 319.
- Bonneau, M., Caputo, B., Ligier, A., Caparros, R., Unal, S., Arnoldi, D., *et al.* 2018b. The cytoplasmic incompatibility mod diversity is governed by *Wolbachia cidB* gene and not by *cidA* gene in *Culex pipiens* mosquitoes. *Proc. R. Soc. B Biol. Sci.* Submitted.
- Bonneau, M., Landmann, F., Labbé, P., Justy, F., Weill, M. & Sicard, M. 2018c. The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity. *PLoS Pathog.* Minor Revision.
- Bordenstein, S. & Rosengaus, R.B. 2005. Discovery of a novel *Wolbachia* supergroup in isoptera. *Curr. Microbiol.* **51**: 393–398.
- Bordenstein, S.R. & Bordenstein, S.R. 2016. Eukaryotic association module in phage WO genomes from *Wolbachia*. *Nat. Commun.* **7**: 1–10.
- Bordenstein, S.R. & Bordenstein, S.R. 2011. Temperature affects the tripartite interactions between bacteriophage WO, *Wolbachia*, and cytoplasmic incompatibility. *PLoS One* **6**: e29106.
- Bordenstein, S.R., Paraskevopoulos, C., Dunning Hotopp, J.C., Sapountzis, P., Lo, N., Bandi, C., *et al.* 2009. Parasitism and mutualism in *Wolbachia*: What the phylogenomic trees can and cannot say. *Mol. Biol. Evol.* **26**: 231–241.
- Bordenstein, S.R. & Reznikoff, W.S. 2005. Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* **3**: 688–699.
- Bordenstein, S.R., Uy, J.J. & Werren, J.H. 2003. Host genotype determines cytoplasmic incompatibility type in the haplodiploid genus *Nasonia*. *Genetics* **164**: 223–33.
- Bordenstein, S.R. & Wernegreen, J.J. 2004. Bacteriophage flux in endosymbionts (*Wolbachia*): Infection frequency, lateral transfer, and recombination rates. *Mol. Biol. Evol.* **21**: 1981–1991.

- Bordenstein, S.R. & Werren, J.H. 2007. Bidirectional incompatibility among divergent *Wolbachia* and incompatibility level differences among closely related *Wolbachia* in *Nasonia*. *Heredity*. **99**: 278–287.
- Bordenstein, S.R. & Werren, J.H. 1998. Effects of A and B *Wolbachia* and host genotype on interspecies cytoplasmic incompatibility in *Nasonia*. *Genetics* **148**: 1833–1844.
- Bourtzis, K., Dobson, S.L., Xi, Z., Rasgon, J.L., Calvitti, M., Moreira, L.A., *et al.* 2014. Harnessing mosquito-*Wolbachia* symbiosis for vector and disease control. *Acta Trop.* **132**: S150–S163.
- Bourtzis, K., Lees, R.S., Hendrichs, J. & Vreysen, M.J.B. 2016. More than one rabbit out of the hat: Radiation, transgenic and symbiont-based approaches for sustainable management of mosquito and tsetse fly populations. *Acta Trop.* **157**: 115–130.
- Boyd, E.F. & Brüssow, H. 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol.* **10**: 521–529.
- Boyd, E.F., Davis, B.M. & Hochhut, B. 2001. Bacteriophage-bacteriophage interactions in the evolution of pathogenic bacteria. *Trends Microbiol.* **9**: 137–144.
- Boyle, L., O’Neill, S.L., Robertson, H.M. & Karr, T.L. 1993. Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. *Science*. **260**: 1796–1799.
- Braquart-Varnier, C., Altinli, M., Pigeault, R., Chevalier, F.D., Grève, P., Bouchon, D., *et al.* 2015. The mutualistic side of *Wolbachia*-isopod interactions: *Wolbachia* mediated protection against pathogenic intracellular bacteria. *Front. Microbiol.* **6**: 1–15.
- Braquart-Varnier, C., Grève, P., Félix, C. & Martin, G. 2005. Bacteriophage WO in *Wolbachia* infecting terrestrial isopods. *Biochem. Biophys. Res. Commun.* **337**: 580–585.
- Breeuwer, J.A.J. 1997. *Wolbachia* and cytoplasmic incompatibility in the spider mites *Tetranychus urticae* and *T. turkestanii*. *Heredity*. **79**: 41–47.
- Breeuwer, J.A.J. & Jacobs, G. 1996. *Wolbachia*: Intracellular manipulators of mite reproduction. *Exp. Appl. Acarol.* **20**: 421–434.
- Breeuwer, J.A.J. & Werren, J.H. 1990. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature* **346**: 558–560.
- Brelsfoard, C.L., Séchan, Y. & Dobson, S.L. 2008. Interspecific hybridization yields strategy for South Pacific filariasis vector elimination. *PLoS Negl. Trop. Dis.* **2**: 1–6.
- Brian, P.W. 1966. Obligate parasitism in fungi. *Proc. R. Soc. London. Ser. B, Biol. Sci.* **168**: 101–118.
- Bronstein, J.L. 1994. Conditional outcomes in mutualistic interactions. *Trends Ecol. Evol.* **9**: 214–217.
- Bronstein, J.L. 2001. The exploitation of mutualisms. *Ecol. Lett.* **4**: 277–287.
- Bronstein, J.L., Barker, J.L., Lichtenberg, E.M., Richardson, L.L. & Irwin, R.E. 2017. The behavioral ecology of nectar robbing: why be tactic constant? *Curr. Opin. Insect Sci.* **21**: 14–18
- Buchner, P. 1965. *Endosymbiosis of Animals with Plant Microorganisms*. Interscience Publishers, New York.
- Bull, J.J. & Rice, W.R. 1991. Distinguishing mechanisms for the evolution of co-operation. *J. Theor. Biol.* **149**: 63–74.
- Callaini, G., Dallai, R. & Riparbelli, M.G. 1997. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. *J. Cell Sci.* **110**: 271–80.

- Callaini, G., Riparbelli, M.G., Giordano, R. & Dallai, R. 1996. Mitotic Defects Associated with Cytoplasmic Incompatibility in *Drosophila simulans*. *J. Invertebr. Pathol.* **67**: 55–64.
- Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M.L. & Brüssow, H. 2003. Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.* **6**: 417–424.
- Casiraghi, M., Anderson, T.J.C., Bandi, C., Bazzochi, C. & Genchi, C. 2001. A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of *Wolbachia* endosymbionts. *Parasitology* **122**: 93–103.
- Casiraghi, M., Bordenstein, S.R., Baldo, L., Lo, N., Beninati, T., Wernegreen, J.J., *et al.* 2005. Phylogeny of *Wolbachia pipientis* based on *gltA*, *groEL* and *ftsZ* gene sequences: Clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. *Microbiology* **151**: 4015–4022.
- Caspari, E. & Watson, G.S. 1959. On the Evolutionary Importance of Cytoplasmic Sterility in Mosquitoes. *Evolution*. **13**: 568.
- Cerveau, N., Leclercq, S., Leroy, E., Bouchon, D. & Cordaux, R. 2011. Short- and long-term evolutionary dynamics of bacterial insertion sequences: Insights from *Wolbachia* endosymbionts. *Genome Biol. Evol.* **3**: 1175–1186.
- Chafee, M.E., Funk, D.J., Harrison, R.G. & Bordenstein, S.R. 2010. Lateral Phage Transfer in Obligate Intracellular Bacteria (*Wolbachia*): Verification from Natural Populations. *Mol. Biol. Evol.* **27**: 501–505.
- Chambers, E.W., Hapairai, L., Peel, B.A., Bossin, H. & Dobson, S.L. 2011. Male mating competitiveness of a *Wolbachia*-introgressed *Aedes polynesiensis* strain under semi-field conditions. *PLoS Negl. Trop. Dis.* **5**: 1–6.
- Charlat, S., Engelstädter, J., Dyson, E.A., Hornett, E.A., Duploux, A., Tortosa, P., *et al.* 2006. Competing Selfish Genetic Elements in the Butterfly *Hypolimnas bolina*. *Curr. Biol.* **16**: 2453–2458.
- Charlat, S., Hurst, G.D.D. & Merçot, H. 2003a. Evolutionary consequences of *Wolbachia* infections. *Trends Genet.* **19**: 217–223.
- Charlat, S., Le Chat, L. & Merçot, H. 2003b. Characterization of non-cytoplasmic incompatibility inducing *Wolbachia* in two continental African populations of *Drosophila simulans*. *Heredity (Edinb)*. **90**: 49–55.
- Charlat, Sy., Nlrgianaki, A. & Bourtzis, K. 2002. Evolution of *Wolbachia* -Induced Cytoplasmic Incompatibility in *Drosophila Simulans* and *D. Sechellia*. **56**: 1735–1742.
- Cheng, T.C. 1991. *Is parasitism symbiosis ? A definition of terms and the evolution of concepts*. Oxford University Press, Oxford.
- Chrostek, E., Marialva, M.S.P., Esteves, S.S., Weinert, L.A., Martinez, J., Jiggins, F.M., *et al.* 2013. *Wolbachia* Variants Induce Differential Protection to Viruses in *Drosophila melanogaster*: A Phenotypic and Phylogenomic Analysis. *PLoS Genet.* **9**: e1003896.
- Clancy, D. & Hoffmann, A.A. 1998. Environmental effects on cytoplasmic incompatibility and bacterial load in *Wolbachia*-infected *Drosophila simulans*. *Entomol. Exp. Appl.* **86**: 13–24.
- Clancy, D.J. & Hoffmann, a a. 1996. Cytoplasmic incompatibility in *Drosophila simulans*: evolving complexity. *Trends Ecol. Evol. (Personal Ed.* **11**: 145–146.
- Combes, C. 2001. *Parasitism: the ecology and evolution of intimate interactions. Interspecific interactions*. The University of Chicago, Chicago.

- Combes, C., Gavotte, L., Mouli a, C. & Sicard, M. 2018. *Parasitisme: Ecologie et  volution des interactions durables.*, Dunod. Paris.
- Cordaux, R., Bouchon, D. & Gr ve, P. 2011. The impact of endosymbionts on the evolution of host sex-determination mechanisms. *Trends Genet.* **27**: 332–341.
- Cordaux, R., Michel-Salzat, A. & Bouchon, D. 2001. *Wolbachia* infection in crustaceans: Novel hosts and potential routes for horizontal transmission. *J. Evol. Biol.* **14**: 237–243.
- Cordaux, R., Michel-Salzat, A., Frelon-Raimond, M., Rigaud, T. & Bouchon, D. 2004. Evidence for a new feminizing *Wolbachia* strain in the isopod *Armadillidium vulgare*: evolutionary implications. *Heredity.* **93**: 78–84.
- Cosmides, L.M. & Tooby, J. 1981. Cytoplasmic Inheritance and Intragenomic Conflict. *Phys. Basis Hered.* **89**: 83–129.
- Covacin, C. & Barker, S.C. 2007. Supergroup F *Wolbachia* bacteria parasitise lice (Insecta: Phthiraptera). *Parasitol. Res.* **100**: 479–485.
- Coyne, J.A. 1992. Genetics and speciation. *Nature* **355**: 511–515.
- Crainey, J.L., Hurst, J., Lamberton, P.H.L., Cheke, R.A., Griffin, C.E., Wilson, M.D., *et al.* 2017. The genomic architecture of novel *simulium damnosum wolbachia* prophage sequence elements and implications for onchocerciasis epidemiology. *Front. Microbiol.* **8**: 1–17.
- Cube nas-Potts, C. & Matunis, M.J. 2013. SUMO: A Multifaceted Modifier of Chromatin Structure and Function. *Dev. Cell* **24**: 1–12.
- Darby, A.C., Armstrong, S.D., Bah, G.S., Kaur, G., Hughes, M.A., Kay, S.M., *et al.* 2012. Analysis of gene expression from the *Wolbachia* genome of a filarial nematode supports both metabolic and defensive roles within the symbiosis. *Genome Res.* **22**: 2467–2477.
- Dawkins, R. 1982. *The Extended Phenotype: The Long Reach of the Gene.* Oxford University Press, Oxford.
- Dedeine, F., Boul treau, M. & Vavre, F. 2005. *Wolbachia* requirement for oogenesis: Occurrence within the genus *Asobara* (Hymenoptera, Braconidae) and evidence for intraspecific variation in *A. tabida*. *Heredity.* **95**: 394–400.
- Dedeine, F., Vavre, F., Fleury, F., Loppin, B., Hochberg, M.E. & Bouletreau, M. 2001. Removing symbiotic *Wolbachia* bacteria specifically inhibits oogenesis in a parasitic wasp. *Proc. Natl. Acad. Sci.* **98**: 6247–6252.
- Dedeine, F., Vavre, F., Shoemaker, D.D.W. & Boul treau, M. 2004. Intra-individual coexistence of a *Wolbachia* strain required for host oogenesis with two strains inducing cytoplasmic incompatibility in the wasp *Asobara tabida*. *Evolution.* **58**: 2167–2174.
- Deshaies, R.J. 1999. SCF and Cullin/RING H2-Based Ubiquitin Ligases. *Annu. Rev. Cell Dev. Biol.* **15**: 435–467.
- Dobson, S.L., Bourtzis, K., Braig, H.R., Jones, B.F., Zhou, W., Rousset, F., *et al.* 1999. *Wolbachia* infections are distributed throughout insect somatic and germ line tissues. *Insect Biochem. Mol. Biol.* **29**: 153–160.
- Dobson, S.L., Fox, C.W. & Jiggins, F.M. 2002a. The effect of *Wolbachia*-induced cytoplasmic incompatibility on host population size in natural and manipulated systems. *Proc. R. Soc. B Biol. Sci.* **269**: 437–445.
- Dobson, S.L., Marsland, E.J. & Rattanadechakul, W. 2002b. Mutualistic *Wolbachia* infection in *Aedes albopictus*: Accelerating cytoplasmic drive. *Genetics* **160**: 1087–1094.

- Dobson, S.L., Marsland, E.J. & Rattanadechakul, W. 2001. *Wolbachia*-induced cytoplasmic incompatibility in single- and superinfected *Aedes albopictus* (Diptera: Culicidae). *J. Med. Entomol.* **38**: 382–387.
- Douglas, A.E. 1998. Nutritional interactions in insect-microbial symbioses: Aphids and Their Symbiotic Bacteria *Buchnera*. *Annu. Rev. Entomol.* **43**: 17–37.
- Dumas, E., Atyame, C.M., Milesi, P., Fonseca, D.M., Shaikovich, E. V, Unal, S., *et al.* 2013. Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol. Biol.* **13**: 181.
- Dunn, A.M. & Smith, J.E. 2001. Microsporidian life cycles and diversity: the relationship between virulence and transmission. *Microbes Infect.* **3**: 381–388.
- Duron, O., Bernard, J., Atyame, C.M., Dumas, E. & Weill, M. 2012. Rapid evolution of *Wolbachia* incompatibility types. *Proc. R. Soc. B Biol. Sci.* **279**: 4473–4480.
- Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L., Engelstädter, J., *et al.* 2008. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol.* **6**: 27.
- Duron, O., Boureux, A., Echaubard, P., Berthomieu, A., Berticat, C., Fort, P., *et al.* 2007a. Variability and expression of ankyrin domain genes in *Wolbachia* variants infecting the mosquito *Culex pipiens*. *J. Bacteriol.* **189**: 4442–4448.
- Duron, O., Fort, P. & Weill, M. 2006a. Hypervariable prophage WO sequences describe an unexpected high number of *Wolbachia* variants in the mosquito *Culex pipiens*. *Proc. R. Soc. B-Biological Sci.* **273**: 495–502.
- Duron, O., Fort, P. & Weill, M. 2007b. Influence of aging on cytoplasmic incompatibility, sperm modification and *Wolbachia* density in *Culex pipiens* mosquitoes. *Heredity.* **98**: 368–374.
- Duron, O., Lagnel, J., Raymond, M., Bourtzis, K., Fort, P. & Weill, M. 2005. Transposable element polymorphism of *Wolbachia* in the mosquito *Culex pipiens*: Evidence of genetic diversity, superinfection and recombination. *Mol. Ecol.* **14**: 1561–1573.
- Duron, O., Raymond, M. & Weill, M. 2011. Many compatible *Wolbachia* strains coexist within natural populations of *Culex pipiens* mosquito. *Heredity.* **106**: 986–993
- Duron, O. & Weill, M. 2006. *Wolbachia* infection influences the development of *Culex pipiens* embryo in incompatible crosses. *Heredity.* **96**: 493–500.
- Duron, O., Bernard, C., Unal, S., Berthomieu, A., Berticat, C. & Weill, M. 2006b. Tracking factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. *Mol. Ecol.* **15**: 3061–3071.
- Dutra, H.L.C., Rocha, M.N., Dias, F.B.S., Mansur, S.B., Caragata, E.P. & Moreira, L.A. 2016. *Wolbachia* Blocks Currently Circulating Zika Virus Isolates in Brazilian *Aedes aegypti* Mosquitoes. *Cell Host Microbe* **19**: 771–774.
- Ebbert, M.A. 1991. The Interaction Phenotype in the *Drosophila willistoni*-*Spiroplasma* Symbiosis. *Evolution.* **45**: 971–988.
- Ebert, D. 2013. The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. *Annu. Rev. Ecol. Evol. Syst.* **44**: 623–643.
- Engelstädter, J. & Hurst, G.D.D. 2009. The Ecology and Evolution of Microbes that Manipulate Host Reproduction. *Annu. Rev. Ecol. Evol. Syst.* **40**: 127–149.
- Engelstädter, J. & Telschow, A. 2009. Cytoplasmic incompatibility and host population structure. *Heredity.* **103**: 196–207.

- Everett, H. & McFadden, G. 1999. Apoptosis: An innate immune response to virus infection. *Trends Microbiol.* **7**: 160–165.
- Ewald, P. & De Leo, G. 2002. Alternative transmission modes and the evolution of virulence. In: *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management* (M. S. and K. S. U. Dieckmann, JAJ Metz, ed), pp. 10–25. Cambridge University Press, Cambridge.
- Ferdy, J.B., Liu, N. & Sicard, M. 2016. Transmission modes and the evolution of feminizing symbionts. *J. Evol. Biol.* **29**: 2395–2409.
- Fine, P.E.M. 1978. On the dynamics of symbiote-dependent cytoplasmic incompatibility in culicine mosquitoes. *J. Invertebr. Pathol.* **31**: 10–18.
- Fine, P.E.M. 1975. Vectors and Vertical Transmission: an Epidemiologic Perspective. *Ann. N. Y. Acad. Sci.* **266**: 173–194.
- Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. The Clarendon Press, Oxford.
- Flor, M., Hammerstein, P. & Telschow, A. 2007. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and the stability of infection polymorphism in parapatric host populations. *J. Evol. Biol.* **20**: 696–706.
- Flores, H.A. & O’Neill, S.L. 2018. Controlling vector-borne diseases by releasing modified mosquitoes. *Nat. Rev. Microbiol.* **16**: 508–518.
- Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., *et al.* 2005. The *Wolbachia* Genome of *Brugia malayi*: Endosymbiont Evolution within a Human Pathogenic Nematode. *PLoS Biol.* **3**: e121.
- Frank, S.A. 1996. Host-symbiont conflict over the mixing of symbiotic lineages. *Proc. R. Soc. B Biol. Sci.* **263**: 339–344.
- Fry, A.J., Palmer, M.R. & Rand, D.M. 2004. Variable fitness effects of *Wolbachia* infection in *Drosophila melanogaster*. *Heredity.* **93**: 379–389.
- Fukatsu, T. & Hosokawa, T. 2002. Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, *Megacopta punctatissima*. *Appl. Environ. Microbiol.* **68**: 389–396.
- Gantz, V.M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V.M., Bier, E., *et al.* 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc. Natl. Acad. Sci.* **112**: E6736–E6743.
- Gao, L.Y. & Abu Kwaik, Y. 2000. Hijacking of apoptotic pathways by bacterial pathogens. *Microbes Infect.* **2**: 1705–1719.
- Gavotte, L., Henri, H., Stouthamer, R., Charif, D., Charlat, S., Boulétreau, M., *et al.* 2007. A survey of the bacteriophage WO in the endosymbiotic bacteria *Wolbachia*. *Mol. Biol. Evol.* **24**: 427–435.
- Gavotte, L., Vavre, F., Henri, H., Ravallec, M., Stouthamer, R. & Boulétreau, M. 2004. Diversity, distribution and specificity of WO phage infection in *Wolbachia* of four insect species. *Insect Mol. Biol.* **13**: 147–153.
- Gebiola, M., Giorgini, M., Kelly, S.E., Doremus, M.R., Ferree, P.M. & Hunter, M.S. 2017. Cytological analysis of cytoplasmic incompatibility induced by *Cardinium* suggests convergent evolution with its distant cousin *Wolbachia*. *Proc. R. Soc. B Biol. Sci.* **284**: 20171433.
- Georghiou, G.P., Metcalf, R.L. & Gidden, F.E. 1966. Carbamate-resistance in mosquitos. *Bull. World Health Organ.* **35**: 691–708.
- Gerdes, K. & Maisonneuve, E. 2012. Bacterial Persistence and Toxin-Antitoxin Loci. *Annu. Rev. Microbiol.* **66**: 103–123.

- Ghera, R.L., Werren, J.H., Weisburg, W., Cote, R., Woese, C.R., Mandelco, L., *et al.* 1991. *Arsenophonus nasoniae* gen. nov., sp. nov., the Causative Agent of the Son-Killer Trait in the Parasitic Wasp *Nasonia vitripennis*. *Int. J. Syst. Bacteriol.* **41**: 563–565.
- Gilbert, F.S., Haines, N. & Kickson, K. 1991. Empty flowers. *Funct. Ecol.* **5**: 29–39.
- Glickman, M.H. & Ciechanover, A. 2002. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol. Rev.* **82**: 373–428.
- Glowska, E., Dragun-Damian, A., Dabert, M. & Gerth, M. 2015. New *Wolbachia* supergroups detected in quill mites (Acari: Syringophilidae). *Infect. Genet. Evol.* **30**: 140–146.
- Gotoh, T., Noda, H. & Ito, S. 2007. *Cardinium* symbionts cause cytoplasmic incompatibility in spider mites. *Heredity.* **98**: 13–20.
- Gottlieb, Y., Zchori-Fein, E., Werren, J.H. & Karr, T.L. 2002. Diploidy restoration in *Wolbachia*-infected *Muscidifurax uniraptor* (Hymenoptera: Pteromalidae). *J. Invertebr. Pathol.* **81**: 166–174.
- Grenier, S., Pintureau, B., Heddi, A., Lassabliere, F., Jager, C., Louis, C., *et al.* 1998. Successful horizontal transfer of *Wolbachia* symbionts between *Trichogramma* wasps. *Proc. R. Soc. Lond. B. Biol. Sci.* **265**: 1441–1445.
- Guillemaud, T., Pasteur, N. & Rousset, F. 1997. Contrasting levels of variability between cytoplasmic genomes and incompatibility types in the mosquito *Culex pipiens*. *Proc. R. Soc. B Biol. Sci.* **264**: 245–251.
- Haegeman, A., Vanholme, B., Jacob, J., Vandekerckhove, T.T.M., Claeys, M., Borgonie, G., *et al.* 2009. An endosymbiotic bacterium in a plant-parasitic nematode: Member of a new *Wolbachia* supergroup. *Int. J. Parasitol.* **39**: 1045–1054.
- Hagimori, T., Abe, Y., Date, S. & Miura, K. 2006. The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. *Curr. Microbiol.* **52**: 97–101.
- Hamilton, W.D. 1967. Extraordinary sex ratios. *Science.* **156**: 477–488.
- Harumoto, T., Anbutsu, H. & Fukatsu, T. 2014. Male-Killing *Spiroplasma* Induces Sex-Specific Cell Death via Host Apoptotic Pathway. *PLoS Pathog.* **10**.
- Harumoto, T. & Lemaître, B. 2018. Male-killing toxin in a *Drosophila* bacterial symbiont. *Nature.* **557**:252-255.
- Heath, B.D., Butcher, R.D.J., Whitfield, W.G.F. & Hubbard, S.F. 1999. Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. *Curr. Biol.* **9**: 313–316.
- Hershberg, R. & Petrov, D.A. 2010. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet.* **6**: e100115.
- Hertig, M. 1936. The Rickettsia, *Wolbachia pipientis* (gen. et sp.n.) and Associated Inclusions of the Mosquito, *Culex pipiens*. *Parasitology* **28**: 453–486.
- Hertig, M. & Wolbach, S.B. 1924. Studies on Rickettsia-Like Micro-Organisms in Insects. *J. Med. Res.* **44**: 329–374.7.
- Hildebrand, F., Meyer, A. & Eyre-Walker, A. 2010. Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet.* **6**: e1001107.
- Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A. & Werren, J.H. 2008. How many species are infected with *Wolbachia*? a statistical analysis of current data. *FEMS Microbiol. Lett.* **281**: 215–220.

- Hochstrasser, M., Chen, H., Beckmann, J.F., Ronau, J.A. & Berk, J. 2018. Toxin-Antitoxin Modules Drive Cytoplasmic Incompatibility by Intracellular *Wolbachia* Bacteria. *Wolbachia Conf.*
- Hoffmann, A.A., Clancy, D. & Duncan, J. 1996. Naturally-occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. *Heredity*. **76**: 1–8.
- Hoffmann, A.A., Iturbe-Ormaetxe, I., Callahan, A.G., Phillips, B.L., Billington, K., Axford, J.K., *et al.* 2014. Stability of the wMel *Wolbachia* Infection following Invasion into *Aedes aegypti* Populations. *PLoS Negl. Trop. Dis.* **8**: e3115.
- Hoffmann, A.A., Montgomery, B.L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P.H., Muzzi, F., *et al.* 2011. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* **476**: 454–457
- Hoffmann, A.A., Turelli, M. & Harshman, L.G. 1990. Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. *Genetics* **126**: 933–948.
- Hoffmann, A.A., Turelli, M. & Simmons, G.M. 1986. Unidirectional Incompatibility between Populations of *Drosophila simulans*. *Evolution*. **40**: 692–701.
- Hornett, E.A., Charlat, S., Duploux, A.M.R., Davies, N., Roderick, G.K., Wedell, N., *et al.* 2006. Evolution of male-killer suppression in a natural population. *PLoS Biol.* **4**: 1643–1648.
- Hoshizaki, S. & Shimada, T. 1995. PCR- based detection of *Wolbachia*, cytoplasmic incompatibility microorganisms, infected in natural populations of *Laodelphax striatellus* (Homoptera: Delphacidae) in central Japan: has the distribution of *Wolbachia* spread recently? *Insect Mol. Biol.* **4**: 237–243.
- Hosokawa, T., Kikuchi, Y. & Fukatsu, T. 2007. How many symbionts are provided by mothers, acquired by offspring, and needed for successful vertical transmission in an obligate insect-bacterium mutualism? *Mol. Ecol.* **16**: 5316–5325.
- Hosokawa, T., Kikuchi, Y., Nikoh, N., Shimada, M. & Fukatsu, T. 2006. Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. *PLoS Biol.* **4**: 1841–1851.
- Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X.-Y. & Fukatsu, T. 2010. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc. Natl. Acad. Sci.* **107**: 769–774.
- Hughes, G.L., Koga, R., Xue, P., Fukatsu, T. & Rasgon, J.L. 2011. *Wolbachia* infections are virulent and inhibit the human malaria parasite *Plasmodium falciparum* in *Anopheles gambiae*. *PLoS Pathog.* **7**: 3–10.
- Huigens, M.E., De Almeida, R.P., Boons, P.A.H., Luck, R.F. & Stouthamer, R. 2004. Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing *Wolbachia* in *Trichogramma* wasps. *Proc. R. Soc. B Biol. Sci.* **271**: 509–515.
- Hunter, M.S., Perlman, S.J. & Kelly, S.E. 2003. A bacterial symbiont in the Bacteroidetes induces cytoplasmic incompatibility in the parasitoid wasp *Encarsia pergandiella*. *Proc. R. Soc. B Biol. Sci.* **270**: 2185–2190.
- Hurst, G.D.D., Bandi, C., Sacchi, L., Cochrane, A.G., Bertrand, D., Karaca, I., *et al.* 1999. *Adonia variegata* (Coleoptera: Coccinellidae) bears maternally inherited *Flavobacteria* that kill males only. *Parasitology* **118**: 125–134.
- Hurst, G.D.D. & Frost, C.L. 2015. Reproductive Parasitism: Maternally Inherited Symbionts in a Biparental World. *Cold Spring Harb. Perspect. Biol.* **7**: a017699.
- Hurst, G.D.D. & Majerus, M.E.N. 1993. Why do maternally inherited microorganisms kill males? *Heredity*. **71**: 81–95.

- Hurst, G.D.D., Majerus, M.E.N. & Walker, L.E. 1993. The importance of cytoplasmic male killing elements in natural populations of the two spot ladybird, *Adalia bipunctata* (Linnaeus) (Coleoptera: Coccinellidae). *Biol. J. Linn. Soc.* **49**: 195–202.
- Hurst, G.D.D. & Schilthuizen, M. 1998. Selfish genetic elements and speciation. *Heredity (Edinb)*. **80**: 2–8.
- Hurst, G.D.D. & Werren, J.H. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.* **2**: 597–606.
- Hurst, L.D. 1991a. The evolution of cytoplasmic incompatibility or when spite can be successful. *J. Theor. Biol.* **148**: 269–277.
- Hurst, L.D. 1991b. The Incidences and Evolution of Cytoplasmic Male Killers. *Proc. R. Soc. B Biol. Sci.* **244**: 91–99.
- Ikeda, T., Ishikawa, H. & Sasaki, T. 2003. Infection density of *Wolbachia* and level of cytoplasmic incompatibility in the Mediterranean flour moth, *Ephestia kuehniella*. *J. Invertebr. Pathol.* **84**: 1–5.
- Ironside, J.E., Smith, J.E., Hatcher, M.J., Sharpe, R.G., Rollinson, D. & Dunn, A.M. 2003. Two species of feminizing microsporidian parasite coexist in populations of *Gammarus duebeni*. *J. Evol. Biol.* **16**: 467–473.
- Irving-Bell, R. 1983. Cytoplasmic incompatibility within and between *Culex molestus* and *Culex quinquefasciatus* (Diptera: Culicidae). *J. Med. Entomol.* **20**: 44–48.
- Irwin, R.E. & Brody, A.K. 1999. Nectar-robbing bumble bees reduce the fitness of *Ipomopsis aggregata* (Polemoniaceae). *Ecology* **80**: 1703–1712.
- Iturbe-Ormaetxe, I., Burke, G.R., Riegler, M. & O'Neill, S.L. 2005. Distribution, expression, and motif variability of ankyrin domain genes in *Wolbachia pipientis*. *J. Bacteriol.* **187**: 5136–5145.
- Jaenike, J., Dyer, K.A., Cornish, C. & Minhas, M.S. 2006. Asymmetrical reinforcement and *Wolbachia* infection in *Drosophila*. *PLoS Biol.* **4**: 1852–1862.
- James, A.C. & Ballard, J.W.O. 2000. Expression of cytoplasmic incompatibility in *Drosophila simulans* and its impact on infection frequencies and distribution of *Wolbachia pipientis*. *Evolution.* **54**: 1661–1672.
- Jamnongluk, W., Kittayapong, P., Baisley, K.J. & O'Neill, S.L. 2000. *Wolbachia* infection and expression of cytoplasmic incompatibility in *Armigeres subalbatus* (Diptera: Culicidae). *J. Med. Entomol.* **37**: 53–57.
- Jiggins, F.M., Hurst, G.D.D. & Majerus, M.E.N. 2000. Sex-ratio-distorting *Wolbachia* causes sex-role reversal in its butterfly host. *Proc. R. Soc. B Biol. Sci.* **267**: 69–73.
- Kambhampati, S., Rai, K.S. & Verleye, D.M. 1992. Frequencies of mitochondrial DNA haplotypes in laboratory cage populations of the mosquito, *Aedes albopictus*. *Genetics* **132**: 205–209.
- Kambris, Z., Blagborough, A.M., Pinto, S.B., Blagrove, M.S.C., Godfray, H.C.J., Sinden, R.E., *et al.* 2010. *Wolbachia* stimulates immune gene expression and inhibits *Plasmodium* development in *Anopheles gambiae*. *PLoS Pathog.* **6**: e1001143.
- Katakura, Y. 1989. Endocrine and genetic control of sex differentiation in the Malacostracan Crustacea. *Invertebr. Reprod. Dev.* **16**: 177–181.
- Keeling, M.J., Jiggins, F.M. & Read, J.M. 2003. The invasion and coexistence of competing *Wolbachia* strains. *Heredity (Edinb)*. **91**: 382–388.

- Kent, B.N., Funkhouser, L.J., Setia, S. & Bordenstein, S.R. 2011a. Evolutionary genomics of a temperate bacteriophage in an obligate intracellular bacteria (*Wolbachia*). *PLoS One* **6**: e24984.
- Kent, B.N., Salichos, L., Gibbons, J.G., Rokas, A., Newton, I.L.G., Clark, M.E., *et al.* 2011b. Complete bacteriophage transfer in a bacterial endosymbiont (*Wolbachia*) determined by targeted genome capture. *Genome Biol. Evol.* **3**: 209–218.
- Kittayapong, P., Mongkalagoon, P., Baimai, V. & O'Neill, S.L. 2002. Host age effect and expression of cytoplasmic incompatibility in field populations of *Wolbachia*-superinfected *Aedes albopictus*. *Heredity*. **88**: 270–274.
- Klasson, L., Walker, T., Sebaihia, M., Sanders, M.J., Quail, M. a., Lord, A., *et al.* 2008. Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol. Biol. Evol.* **25**: 1877–1887.
- Klasson, L., Westberg, J., Sapountzis, P., Naslund, K., Lutnaes, Y., Darby, A.C., *et al.* 2009. The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proc. Natl. Acad. Sci.* **106**: 5725–5730.
- Kose, H. & Karr, T.L. 1995. Organization of *Wolbachia pipientis* in the *Drosophila fertilized* egg and embryo revealed by an anti-*Wolbachia* monoclonal antibody. *Mech. Dev.* **51**: 275–288.
- Kremer, N., Charif, D., Henri, H., Bataille, M., Prévost, G., Kraaijeveld, K., *et al.* 2009. A new case of *Wolbachia* dependence in the genus *Asobara*: Evidence for parthenogenesis induction in *Asobara japonica*. *Heredity*. **103**: 248–256.
- Landmann, F., Orsi, G.A., Loppin, B. & Sullivan, W. 2009. *Wolbachia*-Mediated Cytoplasmic Incompatibility Is Associated with Impaired Histone Deposition in the Male Pronucleus. *PLoS Pathog.* **5**: e1000343.
- Landmann, F., Voronin, D., Sullivan, W. & Taylor, M.J. 2011. Anti-filarial activity of antibiotic therapy is due to extensive apoptosis after *Wolbachia* depletion from filarial nematodes. *PLoS Pathog.* **7**: 1–12.
- Lassy, C.W. & Karr, T.L. 1996. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. *Mech. Dev.* **57**: 47–58.
- Laven, H. 1956. Cytoplasmic Inheritance in *Culex*. *Nature* **177**: 141–142.
- Laven, H. 1967a. Eradication of *Culex pipiens fatigans* through Cytoplasmic Incompatibility. *Nature* **216**: 383–384.
- Laven, H. 1967b. Speciation and evolution in *Culex pipiens*. In: *Genetics of Insect Vectors of Disease* (J. Wright & R. Pal, eds), pp. 251–275. Elsevier: Amsterdam.
- Le Clec'h, W., Chevalier, F.D., Genty, L., Bertaux, J., Bouchon, D. & Sicard, M. 2013a. Cannibalism and Predation as Paths for Horizontal Passage of *Wolbachia* between Terrestrial Isopods. *PLoS One* **8**: e60232.
- Le Clec'h, W., Raimond, M., Guillot, S., Bouchon, D. & Sicard, M. 2013b. Horizontal transfers of feminizing versus non-feminizing *Wolbachia* strains: From harmless passengers to pathogens. *Environ. Microbiol.* **15**: 2922–2936.
- Leclercq, S., Thézé, J., Chebbi, M.A., Giraud, I., Moumen, B., Ernenwein, L., *et al.* 2016. Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome. *Proc. Natl. Acad. Sci.* **113**: 15036–15041.
- Lee, M.T., Bakir, A.A., Nguyen, K.N. & Bachant, J. 2011. The SUMO isopeptidase Ulp2p is required to prevent recombination-induced chromosome segregation lethality following DNA replication stress. *PLoS Genet.* **7**: 20–22.

- Lefoulon, E., Bain, O., Makepeace, B.L., d'Haese, C., Uni, S., Martin, C., *et al.* 2016. Breakdown of coevolution between symbiotic bacteria *Wolbachia* and their filarial hosts. *PeerJ* **4**: e1840.
- Lefoulon, E., Gavotte, L., Junker, K., Barbuto, M., Uni, S., Landmann, F., *et al.* 2012. A new type F *Wolbachia* from Splendidofilariinae (Onchocercidae) supports the recent emergence of this supergroup. *Int. J. Parasitol.* **42**: 1025–1036.
- Legrand, J.J. & Juchault, P. 1984. Nouvelles données sur le déterminisme génétique et épigénétique de la monogénie chez le Crustacé Isopode terrestre *Armadillidium vulgare* Latr. *Genet. Sel. Evol.* **16**: 57–84.
- Leigh, B.A., Shropshire, J.D., van Opstal, E., Ngo, K. & S., B. 2018. What is the role of phage WO in cytoplasmic incompatibility? *Wolbachia Conf.* Salem.
- LePage, D. & Bordenstein, S.R. 2013. *Wolbachia*: Can we save lives with a great pandemic? *Trends Parasitol.* **29**: 385–393.
- LePage, D.P., Metcalf, J.A., Bordenstein, S.R., On, J., Perlmutter, J.I., Shropshire, J.D., *et al.* 2017. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **543**: 243–247.
- Lind, P.A. & Andersson, D.I. 2008. Whole-genome mutational biases in bacteria. *Proc. Natl. Acad. Sci.* **105**: 17878–17883.
- Lindsey, A.R.I., Rice, D.W., Bordenstein, S.R., Brooks, A.W., Bordenstein, S.R. & Newton, I.L.G. 2018. Evolutionary Genetics of Cytoplasmic Incompatibility Genes *cifA* and *cifB* in Prophage WO of *Wolbachia*. *Genome Biol. Evol.* **10**: 434–451.
- Lo, N., Casiraghi, M., Salati, E., Bazzocchi, C. & Bandi, C. 2002. How Many *Wolbachia* Supergroups Exist? *Mol. Biol. Evol.* **19**: 341–346.
- Lorenzen, M.D., Gnirke, A., Margolis, J., Garnes, J., Campbell, M., Stuart, J.J., *et al.* 2008. The maternal-effect, selfish genetic element Medea is associated with a composite Tc1 transposon. *Proc. Natl. Acad. Sci.* **105**: 10085–10089.
- Lu, M.H., Zhang, K.J. & Hong, X.Y. 2012. Tripartite associations among bacteriophage WO, *Wolbachia*, and host affected by temperature and age in *Tetranychus urticae*. *Exp. Appl. Acarol.* **58**: 207–220.
- Ma, Y., Chen, W.J., Li, Z.H., Zhang, F., Gao, Y. & Luan, Y.X. 2017. Revisiting the phylogeny of *Wolbachia* in Collembola. *Ecol. Evol.* **7**: 2009–2017.
- Mackinnon, M.J. & Read, A.F. 1999. Genetic Relationships between Parasite Virulence and Transmission in the Rodent Malaria *Plasmodium chabaudi*. *Evolution.* **53**: 689–703.
- Mackinnon, M.J. & Read, A.F. 2003. The effects of host immunity on virulence-transmissibility relationships in the rodent malaria parasite *Plasmodium chabaudi*. *Parasitology* **126**: 103–112.
- Magnin, M., Pasteur, N. & Raymond, M. 1987. Multiple incompatibilities within populations of *Culex pipiens* L. in southern France. *Genetica* **74**: 125–130.
- Mains, J.W., Brelsfoard, C.L., Rose, R.I. & Dobson, S.L. 2016. Female adult *aedes Albopictus* suppression by *Wolbachia*-infected male mosquitoes. *Sci. Rep.* **6**: 1–7.
- Mann, E., Stouthamer, C.M., Kelly, S.E., Dzieciol, M., Hunter, M.S. & Schmitz-Esser, S. 2017. Transcriptome Sequencing Reveals Novel Candidate Genes for *Cardinium hertigii* -Caused Cytoplasmic Incompatibility and Host-Cell Interaction. *mSystems* **2**: 1–16.
- Marshall, J.F. & Staley, J. 1937. Some notes regarding the morphological and biological differentiation of *Culex pipiens* Linnaeus and *Culex molestus* Forskal (Diptera, Culicidae). *Proc. R. Entomol. Soc. London. Ser. A, Gen. Entomol.* **12**: 17–26.

- Martin, G., Sorokine, O., Moniatte, M., Bulet, P., Hetru, C. & Van Dorsselaer, A. 1999. The structure of a glycosylated protein hormone responsible for sex determination in the isopod, *Armadillidium vulgare*. *Eur. J. Biochem.* **262**: 727–736.
- Martin, J., Chong, T. & Ferree, P.M. 2013. Male killing *Spiroplasma* preferentially disrupts neural development in the *Drosophila melanogaster* embryo. *PLoS One* **8**: e79368.
- Masui, S., Kamoda, S., Sasaki, T. & Ishikawa, H. 2000. Distribution and evolution of bacteriophage WO in *Wolbachia*, the endosymbiont causing sexual alterations in Arthropods. *J. Mol. Evol.* **51**: 491–497.
- Masui, S., Kuroiwa, H., Sasaki, T., Inui, M., Kuroiwa, T. & Ishikawa, H. 2001. Bacteriophage WO and virus-like particles in *Wolbachia*, an endosymbiont of arthropods. *Biochem. Biophys. Res. Commun.* **283**: 1099–1104.
- Maynard, S. 1978. *The Evolution of Sex*. Cambridge University Press, Cambridge.
- McAleenan, A., Cordon-Preciado, V., Clemente-Blanco, A., Liu, I.C., Sen, N., Leonard, J., *et al.* 2012. SUMOylation of the α -kleisin subunit of cohesin is required for DNA damage-induced cohesion. *Curr. Biol.* **22**: 1564–1575.
- McCutcheon, J.P., McDonald, B.R. & Moran, N.A. 2009. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc. Natl. Acad. Sci.* **106**: 15394–15399.
- McCutcheon, J.P. & Moran, N. a. 2012. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**: 13–26.
- McCutcheon, J.P. & Moran, N.A. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc. Natl. Acad. Sci.* **104**: 19392–19397.
- McGraw, E. a, Merritt, D.J., Droller, J.N. & O'Neill, S.L. 2001. *Wolbachia*-mediated sperm modification is dependent on the host genotype in *Drosophila*. *Proc. Biol. Sci.* **268**: 2565–2570.
- Merçot, H. & Charlat, S. 2004. *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. *Genetica.* **120**: 51–59.
- Merçot, H. & Poinso, D. 1998. Rescuing *Wolbachia* have been overlooked and discovered on Mount Kilimanjaro. *Nature* **391**: 853–853.
- Miao, E. a & Miller, S.I. 1999. Bacteriophages in the evolution of pathogen-host interactions. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 9452–9454.
- Miller, B.R., Crabtree, M.B. & Savage, H.M. 1996. Phylogeny of fourteen *Culex* mosquito species, including the *Culex pipiens* complex, inferred from the internal transcribed spacers of ribosomal DNA. *Insect Mol. Biol.* **5**: 93–107.
- Moran, N., McCutcheon, J.P. & Nakabachi, A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* **42**: 165–90.
- Moran, N.A. 1996. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci.* **93**: 2873–2878.
- Moran, N.A. & Dunbar, H.E. 2006. Sexual acquisition of beneficial symbionts in aphids. *Proc. Natl. Acad. Sci.* **103**: 12803–12806.
- Moreira, L.A., Iturbe-Ormaetxe, I., Jeffery, J.A., Lu, G., Pyke, A.T., Hedges, L.M., *et al.* 2009. A *Wolbachia* Symbiont in *Aedes aegypti* Limits Infection with Dengue, Chikungunya, and *Plasmodium*. *Cell* **139**: 1268–1278.

- Moretti, R. & Calvitti, M. 2013. Male mating performance and cytoplasmic incompatibility in a wPip *Wolbachia* trans-infected line of *Aedes albopictus* (*Stegomyia albopicta*). *Med. Vet. Entomol.* **27**: 377–386.
- Mousson, L., Zouache, K., Arias-Goeta, C., Raquin, V., Mavingui, P. & Failloux, A.B. 2012. The Native *Wolbachia* Symbionts Limit Transmission of Dengue Virus in *Aedes albopictus*. *PLoS Negl. Trop. Dis.* **6**: e1989.
- Nakamura, Y., Yukuhiro, F., Matsumura, M. & Noda, H. 2012. Cytoplasmic incompatibility involving *Cardinium* and *Wolbachia* in the white-backed planthopper *Sogatella furcifera* (Hemiptera: Delphacidae). *Appl. Entomol. Zool.* **47**: 273–283.
- Narita, S., Kageyama, D., Nomura, M. & Fukatsu, T. 2007. Unexpected mechanism of symbiont-induced reversal of insect sex: Feminizing *Wolbachia* continuously acts on the butterfly *Eurema hecabe* during larval development. *Appl. Environ. Microbiol.* **73**: 4332–4341.
- Newton, I.L., Clark, M.E., Kent, B.N., Bordenstein, S.R., Qu, J., Richards, S., *et al.* 2016. Comparative genomics of two closely related *Wolbachia* with different reproductive effects on hosts. *Genome Biol Evol.* **8**: 1526–1542.
- Nguyen, D.T., Morrow, J.L., Spooner-Hart, R.N. & Riegler, M. 2017. Independent cytoplasmic incompatibility induced by *Cardinium* and *Wolbachia* maintains endosymbiont coinfections in haplodiploid thrips populations. *Evolution.* **71**: 995–1008.
- Nirgianaki, A., Banks, G.K., Frohlich, D.R., Veneti, Z., Braig, H.R., Miller, T.A., *et al.* 2003. *Wolbachia* infections of the whitefly *Bemisia tabaci*. *Curr. Microbiol.* **47**: 93–101.
- Nishino, T. & Morikawa, K. 2002. Structure and function of nucleases in DNA repair: Shape, grip and blade of the DNA scissors. *Oncogene* **21**: 9022–9032.
- Noda, H., Koizumi, Y., Zhang, Q. & Deng, K. 2001. Infection density of *Wolbachia* and incompatibility level in two planthopper species, *Laodelphax striatellus* and *Sogatella furcifera*. *Insect Biochem. Mol. Biol.* **31**: 727–737.
- Nor, I., Engelstädter, J., Duron, O., Reuter, M., Sagot, M.-F. & Charlat, S. 2013. On the genetic architecture of cytoplasmic incompatibility: inference from phenotypic data. *Am. Nat.* **182**: E15–24.
- Nor, I., Hermelin, D., Charlat, S., Engelstadter, J., Reuter, M., Duron, O., *et al.* 2012. Mod/Resc parsimony inference: Theory and application. *Inf. Comput.* **213**: 23–32.
- Nyholm, S. V. & McFall-Ngai, M.J. 2004. The winnowing: Establishing the squid - Vibrios symbiosis. *Nat. Rev. Microbiol.* **2**: 632–642.
- O'Connor, L., Plichart, C., Sang, A.C., Brelsfoard, C.L., Bossin, H.C. & Dobson, S.L. 2012. Open Release of Male Mosquitoes Infected with a *Wolbachia* Biopesticide: Field Performance and Infection Containment. *PLoS Negl. Trop. Dis.* **6**: e1797.
- O'Neill, S.L. & Karr, T.L. 1990. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature* **348**: 178–180.
- O'Neill, S.L. & Paterson, H.E.H. 1992. Crossing type variability associated with cytoplasmic incompatibility in Australian populations of the mosquito *Culex quinquefasciatus* Say. *Med. Vet. Entomol.* **6**: 209–216.
- Ohta, T. 1987. Very slightly deleterious mutations and the molecular clock. *J. Mol. Evol.* **26**: 1–6.
- Okuno, A., Hasegawa, Y. & Nagasawa, H. 1997. Purification and Properties of Androgenic Gland Hormone from the Terrestrial Isopod *Armadillidium vulgare*. *Zoolog. Sci.* **14**: 837–842.

- Oliver, K.M., Campos, J., Moran, N.A. & Hunter, M.S. 2008. Population dynamics of defensive symbionts in aphids. *Proc. R. Soc. B Biol. Sci.* **275**: 293–299.
- Oliver, K.M., Degnan, P.H., Hunter, M.S. & Moran, N.A. 2009. Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science*. **325**: 992–994.
- Oliver, K.M., Russell, J.A., Moran, N.A. & Hunter, M.S. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl. Acad. Sci.* **100**: 1803–1807.
- Panaram, K. & Marshall, J.L. 2007. F supergroup *Wolbachia* in bush crickets: What do patterns of sequence variation reveal about this supergroup and horizontal transfer between nematodes and arthropods? *Genetica* **130**: 53–60.
- Pannebakker, B. a, Pijnacker, L.P., Zwaan, B.J. & Beukeboom, L.W. 2004. Cytology of *Wolbachia*-induced parthenogenesis in *Leptopilina clavipes* (Hymenoptera: Figitidae). *Genome* **47**: 299–303.
- Pannebakker, B.A., Loppin, B., Elemans, C.P.H., Humblot, L. & Vavre, F. 2007. Parasitic inhibition of cell death facilitates symbiosis. *Proc. Natl. Acad. Sci.* **104**: 213–215.
- Pannebakker, B.A., Schidlo, N.S., Boskamp, G.J.F., Dekker, L., Van Dooren, T.J.M., Beukeboom, L.W., *et al.* 2005. Sexual functionality of *Leptopilina clavipes* (Hymenoptera: Figitidae) after reversing *Wolbachia*-induced parthenogenesis. *J. Evol. Biol.* **18**: 1019–1028.
- Paraskevopoulos, C., Bordenstein, S.R., Wernegreen, J.J., Werren, J.H. & Bourtzis, K. 2006. Toward a *Wolbachia* multilocus sequence typing system: Discrimination of *Wolbachia* strains present in *Drosophila* species. *Curr. Microbiol.* **53**: 388–395.
- Penz, T., Schmitz-Esser, S., Kelly, S.E., Cass, B.N., Müller, A., Woyke, T., *et al.* 2012. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in *Cardinium hertigii*. *PLoS Genet.* **8**: e1003012.
- Pijls, J.W.A.M., Van Steenberghe, H.J. & Van Alphen, J.J.M. 1996. Asexuality cured: The relations and differences between sexual and asexual *Apoanagyrus diversicornis*. *Heredity*. **76**: 506–513.
- Poinsot, D. 1997. Infection par *Wolbachia* chez *Drosophila simulans* : étude des interactions hôte-symbiote. Pierre et Marie CURIE (Paris IV).
- Poinsot, D., Bourtzis, K., Markakis, G., Savakis, C. & Merçot, H. 1998. *Wolbachia* Transfer from *Drosophila melanogaster* into *D. simulans*: Host Effect and Cytoplasmic Incompatibility Relationships. *Genet. Soc. Am.* **150**: 227–237.
- Poinsot, D., Charlat, S. & Merçot, H. 2003. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: Confronting the models with the facts. *BioEssays* **25**: 259–265.
- Poinsot, D. & Merçot, H. 1999. *Wolbachia* can rescue from cytoplasmic incompatibility while being unable to induce it. In: *From symbiosis to eukaryotism: Endocytobiology VII*, pp. 221–334. University of Geneva, Geneva.
- Poirié, M. & Coustau, C. 2011. The evolutionary ecology of aphids' immunity. *Invertebr. Surviv. J.* **8**: 247–255.
- Provencher, L.M., Morse, G.E., Weeks, a R. & Normak, B.B. 2005. Parthenogenesis in the *Aspidiotus nerii* complex (Hemiptera: Diaspididae): a single origin of a worldwide, polyphagous lineage associated with *Cardinium* bacteria. *Ann. Entomol. Soc. Am.* **98**: 629–635.
- Rankin, D.J., Rocha, E.P.C. & Brown, S.P. 2011. What traits are carried on mobile genetic elements, and why? *Heredity*. **106**: 1–10.
- Rasgon, J.L. & Scott, T.W. 2003. *Wolbachia* and Cytoplasmic Incompatibility in the California *Culex pipiens* Mosquito Species Complex: Parameter Estimates and Infection Dynamics in Natural Populations. *Genetics* **165**: 2029–2038.

- Raychoudhury, R., Baldo, L., Oliveira, D.C.S.G. & Werren, J.H. 2009. Modes of acquisition of *Wolbachia*: Horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. *Evolution*. **63**: 165–183.
- Raymond, M., Magnin, M., Pasteur, N., Pasteur, G. & Sinegre, G. 1986. Cytoplasmic incompatibility in the mosquito *Culex-pipiens* L. from southern france - implications for the selection and dispersal of insecticide resistance genes in natural-populations. *Genetica* **70**: 113–118.
- Reed, K.M. & Werren, J.H. 1995. Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria (*Wolbachia*): A comparative study of early embryonic events. *Mol. Reprod. Dev.* **40**: 408–418.
- Regus, J.U., Gano, K.A., Hollowell, A.C. & Sachs, J.L. 2014. Efficiency of partner choice and sanctions in *Lotus* is not altered by nitrogen fertilization. *Proc. R. Soc. B Biol. Sci.* **281** : 20132587
- Resnik, D.B. 2014. Ethical issues in field trials of genetically modified disease-resistant mosquitoes. *Dev. World Bioeth.* **14**: 37–46.
- Reyes-Turcu, F.E., Ventii, K.H. & Wilkinson, K.D. 2009. Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. *Annu. Rev. Biochem.* **78**: 363–397.
- Reynolds, K.T. & Hoffmann, A.. 2002. Male age, host effects and the weak expression or non-expression of cytoplasmic incompatibility in *Drosophila* strains infected by maternally transmitted *Wolbachia*. *Genet. Res.* **80**: 79–87.
- Reynolds, K.T., Thomson, L.J. & Hoffmann, A.A. 2003. The effects of host age, host nuclear background and temperature on phenotypic effects of the virulent *Wolbachia* strain popcorn in *Drosophila melanogaster*. *Genetics* **164**: 1027–1034.
- Riegler, M., Charlat, S., Stauffer, C. & Merçot, H. 2004. *Wolbachia* Transfer from *Rhagoletis cerasi* to *Drosophila simulans*: Investigating the Outcomes of Host-Symbiont Coevolution. *Appl. Environ. Microbiol.* **70**: 273–279.
- Rigaud, T. 1997. Inherited microorganisms and sex determination of arthropod host. In: *Influential Passenger: Inherited Microorganisms and Arthropod Reproduction* (S. L. O'Neill, J. H. Werren, & A. A. Hoffmann, eds), pp. 91–101. Oxford University Press, Oxford.
- Rigaud, T., Antoine, D., Marcade, I. & Juchault, P. 1997a. The effect of temperature on sex ratio in the isopod *Porcellionides pruinosus*: Environmental sex determination or a by-product of cytoplasmic sex determination? *Evol. Ecol.* **11**: 205–215.
- Rigaud, T. & Juchault, P. 1992. Genetic control of the vertical transmission of a cytoplasmic sex factor in *Armadillidium vulgare* Latr. (crustacea, oniscidea). *Heredity*. **68**: 47–52.
- Rigaud, T. & Juchault, P. 1995. Success and failure of horizontal transfers of feminizing *Wolbachia* endosymbionts in woodlice. *J. Evol. Biol.* **8**: 249–255.
- Rigaud, T., Juchault, P. & Mocquard, J.P. 1997b. The evolution of sex determination in isopod crustaceans. *BioEssays* **19**: 409–416.
- Rigaud, T., Moreau, J. & Juchault, P. 1999. *Wolbachia* infection in the terrestrial isopod *Oniscus asellus*: Sex ratio distortion and effect on fecundity. *Heredity*. **83**: 469–475.
- Rigaud, T., Pennings, P.S. & Juchault, P. 2001. *Wolbachia* bacteria effects after experimental interspecific transfers in terrestrial isopods. *J. Invertebr. Pathol.* **77**: 251–257.
- Rigaud, T., Souty-Grosset, C., Raimond, R., Mocquard, J. & Juchault, P. 1991. Feminizing endocytobiosis in the terrestrial crustacean *Armadillidium vulgare* Latr. (Isopoda): Recent acquisitions. *Endocyt Cell Res* **7**: 259–273.

- Rocha, E.P.C. & Danchin, A. 2002. Base composition bias might result from competition for metabolic resources. *Trends Genet.* **18**: 291–294.
- Rodgers-Gray, T.P., Smith, J.E., Ashcroft, A.E., Isaac, R.E. & Dunn, A.M. 2004. Mechanisms of parasite-induced sex reversal in *Gammarus duebeni*. *Int. J. Parasitol.* **34**: 747–753.
- Ros, V.I.D., Fleming, V.M., Feil, E.J. & Breeuwer, J.A.J. 2009. How diverse is the genus *Wolbachia*? Multiple-gene sequencing reveals a putatively new *Wolbachia* supergroup recovered from spider mites (Acari: Tetranychidae). *Appl. Environ. Microbiol.* **75**: 1036–1043.
- Rousset, F. & Raymond, M. 1991. Cytoplasmic incompatibility in insects: Why sterilize females? *Trends Ecol. Evol.* **6**: 54–57.
- Rousset, F., Raymond, M. & Kjellberg, F. 1991. Cytoplasmic incompatibilities in the mosquito *Culex pipiens*: How to explain a cytotype polymorphism? *J. Evol. Biol.* **4**: 69–81.
- Rousset, F. & Solignac, M. 1995. Evolution of single and double *Wolbachia* symbioses during speciation in the *Drosophila simulans* complex. *Proc. Natl. Acad. Sci. U. S. A.* **92**: 6389–6393.
- Rowley, S.M., Raven, R.J. & McGraw, E.A. 2004. *Wolbachia pipientis* in Australian spiders. *Curr. Microbiol.* **49**: 208–214.
- Ruang-Areerate, T., Kittayapong, P., Baimai, V. & O'Neill, S.L. 2003. Molecular phylogeny of *Wolbachia* endosymbionts in Southeast Asian mosquitoes (Diptera: Culicidae) based on wsp gene sequences. *J. Med. Entomol.* **40**: 1–5.
- Ryan, S.L. & Saul, G.B. 1968. Post-fertilization effect of incompatibility factors in *Mormoniella*. *MGG Mol. Gen. Genet.* **103**: 29–36.
- Salzberg, S.L., Puiu, D., Sommer, D.D., Nene, V. & Lee, N.H. 2009. Genome sequence of the *Wolbachia* endosymbiont of *Culex quinquefasciatus*JHB. *J. Bacteriol.* **191**: 1725.
- Sanogo, Y.O. & Dobson, S.L. 2006. WO bacteriophage transcription in *Wolbachia*-infected *Culex pipiens*. *Insect Biochem. Mol. Biol.* **36**: 80–85.
- Scarborough, C.L., Ferrari, J. & Godfray, H.C.J. 2005. Aphid Protected from Pathogen by Endosymbiont. *Science*. **310**: 1781–1781.
- Schmitz, A., Anselme, C., Ravallec, M., Rebuf, C., Simon, J.C., Gatti, J.L., *et al.* 2012. The cellular immune response of the pea aphid to foreign intrusion and symbiotic challenge. *PLoS One* **7**: e42114.
- Schnell, J.D. & Hicke, L. 2003. Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *J. Biol. Chem.* **278**: 35857–35860.
- Scott, T.W., Takken, W., Knols, B.G.J. & Boëte, C. 2002. The ecology of genetically modified mosquitoes. *Science*. **298**: 117–119.
- Servick, K. 2016. Winged warriors. *Science*. **354**: 164–167.
- Shaheen, M., Shanmugam, I. & Hromas, R. 2010. The role of PCNA posttranslational modifications in translesion synthesis. *J. Nucleic Acids* **2010**: 1-8.
- Shropshire, J.D., On, J., Layton, E.M., Zhou, H. & Bordenstein, S.R. 2018. One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **115**: 4987–4991.
- Sicard, M., Bouchon, D., Ceyrac, L., Raimond, R., Thierry, M., Le Clec'h, W., *et al.* 2014a. Bidirectional cytoplasmic incompatibility caused by *Wolbachia* in the terrestrial isopod *Porcellio dilatatus*. *J. Invertebr. Pathol.* **121**: 28–36.

- Sicard, M., Brugirard-Ricaud, K., Pagès, S., Lanois, A., Boemare, N.E., Brehélin, M., *et al.* 2004. Stages of infection during the tripartite interaction between *Xenorhabdus nematophila*, its nematode vector, and insect hosts. *Appl. Environ. Microbiol.* **70**: 6473–6480.
- Sicard, M., Dittmer, J., Grève, P., Bouchon, D. & Braquart-Varnier, C. 2014b. A host as an ecosystem: *Wolbachia* coping with environmental constraints. *Environ. Microbiol.* **16**: 3583–3607.
- Sigaud, S., Becquet, V., Frenco, P., Puppo, A. & Hérouart, D. 1999. Differential regulation of two divergent *Sinorhizobium meliloti* genes for HP11-like catalases during free-living growth and protective role of both catalases during symbiosis. *J. Bacteriol.* **181**: 2634–2639.
- Simms, E.L., Taylor, D.L., Povich, J., Shefferson, R.P., Sachs, J.L., Urbina, M., *et al.* 2006. An empirical test of partner choice mechanisms in a wild legume-*rhizobium* interaction. *Proc. R. Soc. B Biol. Sci.* **273**: 77–81.
- Sinkins, S., Braig, H. & O’neill, S. 1995. *Wolbachia* superinfections and the expression of cytoplasmic incompatibility. *Proc. Biol. Sci.* **261**: 325–30.
- Sinkins, S.P., Walker, T., Lynd, A.R., Steven, A.R., Makepeace, B.L., Godfray, H.C.J., *et al.* 2005. *Wolbachia* variability and host effects on crossing type in *Culex* mosquitoes. *Nature* **436**: 257–260.
- Sintupachee, S., Milne, J.R., Poonchaisri, S., Baimai, V. & Kittayapong, P. 2006. Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. *Microb. Ecol.* **51**: 294–301.
- Smith, D.C. & Douglas, A.E. 1987. *The biology of symbiosis*. Edward Arnold, London.
- Stouthamer, R., Luck, R.F. & Hamilton, W.D. 1990. Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex. *Proc. Natl. Acad. Sci. U. S. A.* **87**: 2424–7.
- Stouthamer, R., Russell, J.E., Vavre, F. & Nunney, L. 2010. Intragenomic conflict in populations infected by Parthenogenesis Inducing *Wolbachia* ends with irreversible loss of sexual reproduction. *BMC Evol. Biol.* **10**:229.
- Stouthamer, R., Van Tilborg, M., De Jong, J.H., Nunney, L. & Luck, R.F. 2001. Selfish element maintains sex in natural populations of a parasitoid wasp. *Proc. R. Soc. B Biol. Sci.* **268**: 617–622.
- Sullivan, J. & Jaenike, J. 2006. Male-killing *Wolbachia* and male mate choice: a test with *Drosophila innubila*. *Evol Ecol Res* **8**: 91–102.
- Sutton, E.R., Harris, S.R., Parkhill, J. & Sinkins, S.P. 2014. Comparative genome analysis of *Wolbachia* strain wAu. *BMC Genomics* **15**: 928.
- Takano, S., Tuda, M., Takasu, K., Furuya, N., Imamura, Y., Kim, S., *et al.* 2017. Unique clade of alphaproteobacterial endosymbionts induces complete cytoplasmic incompatibility in the coconut beetle. *Proc. Natl. Acad. Sci.* **114**: 6110–6115.
- Taylor, M.J., Bandi, C. & Hoerauf, A. 2005. *Wolbachia* bacterial endosymbionts of filarial nematodes. In: *Advances in Parasitology*. **60**: 245–284.
- Taylor, M.J., Voronin, D., Johnston, K.L. & Ford, L. 2013. *Wolbachia* filarial interactions. *Cell. Microbiol.* **15**: 520–526.
- Teixeira, L., Ferreira, Á. & Ashburner, M. 2008. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol.* **6**: 2753–2763.
- Telschow, A., Flor, M., Kobayashi, Y., Hammerstein, P. & Werren, J.H. 2007. *Wolbachia*-induced unidirectional cytoplasmic incompatibility and speciation: Mainland-island model. *PLoS One* **2**: e701.

- Telschow, A., Yamamura, N. & Werren, J.H. 2005. Bidirectional cytoplasmic incompatibility and the stable coexistence of two *Wolbachia* strains in parapatric host populations. *J. Theor. Biol.* **235**: 265–274.
- Terry, R.S., Smith, J.E. & Dunn, A.M. 1998. Impact of a novel, feminising *microsporidium* on its crustacean host. *J. Eukaryot. Microbiol.* **45**: 497–501.
- Teyssier, C., Marchandin, H. & Jumas-Bilak, E. 2004. Le génome des alpha-protéobactéries : complexité, réduction, diversité et fluidité. *Can. J. Microbiol.* **50**: 383–396.
- Thomas, F., Lefevre, T. & Raymond, M. 2010. *Biologie évolutive*, I. De Boeck, Bruxelles.
- Timmermans, M.J.T.N., Mariën, J., Roelofs, D., Van Straalen, N.M. & Ellers, J. 2004. Evidence for multiple origins of *Wolbachia* infection in springtails. *Pedobiologia (Jena)*. **48**: 469–475.
- Tinsley, M.C. & Majerus, M.E.N. 2006. A new male-killing parasitism: *Spiroplasma* bacteria infect the ladybird beetle *Anisosticta novemdecimpunctata* (Coleoptera: Coccinellidae). *Parasitology* **132**: 757–765.
- Touchon, M. & Rocha, E.P.C. 2007. Causes of insertion sequences abundance in prokaryotic genomes. *Mol. Biol. Evol.* **24**: 969–981.
- Tram, U., Ferree, P.M. & Sullivan, W. 2003. Identification of *Wolbachia*–host interacting factors through cytological analysis. *Microbes Infect.* **5**: 999–1011.
- Tram, U., Fredrick, K., Werren, J.H. & Sullivan, W. 2006. Paternal chromosome segregation during the first mitotic division determines *Wolbachia*-induced cytoplasmic incompatibility phenotype. *J. Cell Sci.* **119**: 3655–3663.
- Tram, U. & Sullivan, W. 2002. Role of Delayed Nuclear Envelope Breakdown and Mitosis in *Wolbachia*-Induced Cytoplasmic Incompatibility. *Science*. **296**: 1124–1126.
- Turelli, M. 1994. Evolution of Incompatibility-Inducing Microbes and Their Hosts. *Evolution*. **48**: 1500–1513.
- Turelli, M. & Hoffmann, A.A. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**: 440–442.
- Turelli, M., Hoffmann, A.A. & McKechnie, S.W. 1992. Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. *Genetics* **132**: 713–723.
- Vala, F., Egas, M., Breeuwer, J.A.J. & Sabelis, M.W. 2004. *Wolbachia* affects oviposition and mating behaviour of its spider mite host. *J. Evol. Biol.* **17**: 692–700.
- Valbuena, G. & Walker, D.H. 2009. Infection of the endothelium by members of the order Rickettsiales. *Thromb. Haemost.* **102**: 1071–1079.
- van Vugt, J.F., Salverda, M., de Jong, J.H. & Stouthamer, R. 2003. The paternal sex ratio chromosome in the parasitic wasp *Trichogramma kaykai* condenses the paternal chromosomes into a dense chromatin mass. *Genome* **46**: 580–587.
- Vandekerckhove, T.T.M., Watteyne, S., Willems, A., Swings, J.G., Mertens, J. & Gillis, M. 1999. Phylogenetic analysis of the 16S rDNA of the cytoplasmic bacterium *Wolbachia* from the novel host *Folsomia candida* (Hexapoda, Collembola) and its implications for wolbachial taxonomy. *FEMS Microbiol. Lett.* **180**: 279–286.
- Vavre, F., Dedeine, F., Quillon, M., Fouillet, P., Fleury, F. & Boulétreau, M. 2001. Within-species diversity of *Wolbachia*-induced cytoplasmic incompatibility in haplodiploid insects. *Evolution*. **55**: 1710–1714.

- Vavre, F., Fleury, F., Lepetit, D., Fouillet, P. & Boulétreau, M. 1999. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Mol. Biol. Evol.* **16**: 1711–1723.
- Vavre, F., Fleury, F., Varaldi, J., Fouillet, P. & Boulétreau, M. 2002. Infection polymorphism and cytoplasmic incompatibility in Hymenoptera-*Wolbachia* associations. *Heredity*. **88**: 361–365.
- Vinogradova E. B. 2000. *Culex pipiens pipiens Mosquitoes: Taxonomy, Distribution, Ecology, Physiology, Genetics, Applied Importance and Control*. Pensoft, Sofia.
- Waldor, M.K. & Mekalanos, J.J. 1996. Lysogenic Conversion by a Filamentous Phage Encoding Cholera Toxin. *Science*. **272**: 1910–1914.
- Walker, T., Johnson, P.H., Moreira, L.A., Iturbe-Ormaetxe, I., Frentiu, F.D., McMeniman, C.J., *et al.* 2011. The wMel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* **476**: 450–455.
- Walker, T., Klasson, L., Sebahia, M., Sanders, M.J., Thomson, N.R., Parkhill, J., *et al.* 2007. Ankyrin repeat domain-encoding genes in the wPip strain of *Wolbachia* from the *Culex pipiens* group. *BMC Biol.* **5**: 1–9.
- Wang, G.H., Sun, B.F., Xiong, T.L., Wang, Y.K., Murfin, K.E., Xiao, J.H., *et al.* 2016a. Bacteriophage WO can mediate horizontal gene transfer in endosymbiotic *Wolbachia* genomes. *Front. Microbiol.* **7**: 1–16.
- Wang, N., Jia, S., Xu, H., Liu, Y. & Huang, D. 2016b. Multiple horizontal transfers of bacteriophage WO and host *Wolbachia* in fig wasps in a closed community. *Front. Microbiol.* **7**: 1–10.
- Wang, Z., Su, X.M., Wen, J., Jiang, L.Y. & Qiao, G.X. 2014. Widespread infection and diverse infection patterns of *Wolbachia* in Chinese aphids. *Insect Sci.* **21**: 313–325.
- Watanabe, M., Tagami, Y., Miura, K., Kageyama, D. & Stouthamer, R. 2012. Distribution Patterns of *Wolbachia* Endosymbionts in the Closely Related Flower Bugs of the Genus Orius: Implications for Coevolution and Horizontal Transfer. *Microb. Ecol.* **64**: 537–545.
- Weeks, A.R. & Breeuwer, J.A.J. 2001. *Wolbachia*-induced parthenogenesis in a genus of phytophagous mites. *Proc. R. Soc. B Biol. Sci.* **268**: 2245–2251.
- Werren, J.H. 1997. Biology of *Wolbachia*. *Annu. Rev. Entomol.* **42**: 587–609.
- Werren, J.H., Baldo, L. & Clark, M.E. 2008. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**: 741–751.
- Werren, J.H. & Beukeboom, L.W. 1998. Sex Determination , Sex Ratios , and genetic conflict. *Annu. Rev. Ecol. Syst.* **29**: 233-61
- Williamson, D.L., Sakaguchi, B., Hackett, K.J., Whitcomb, R.F., Tully, J.G., Carle, P., *et al.* 1999. *Spiroplasma poulsonii* sp. nov., a new species associated with male-lethality in *Drosophila willistoni*, a neotropical species of fruit fly. *Int. J. Syst. Bacteriol.* **49**: 611–618.
- Wright, J.D., Sjostrand, F.S., Portaro, J.K. & Barr, A.R. 1978. The ultrastructure of the rickettsia-like microorganism *Wolbachia pipientis* and Associated virus-like bodies in the mosquito *Culex pipiens*. *J. Ultrastructure Res.* **63**: 79–85.
- Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress on Genetics.* **1**: 356–366.
- Wu, M., Sun, L. V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., *et al.* 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: A streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2**: 327–341.

- Xi, Z., Khoo, C.C.H. & Dobson, S.L. 2005. *Wolbachia* establishment and invasion in an *Aedes aegypti* laboratory population. *Science*. **310**: 326–328.
- Yamaguchi, Y., Park, J.-H. & Inouye, M. 2011. Toxin-Antitoxin Systems in Bacteria and Archaea. *Annu. Rev. Genet.* **45**: 61–79.
- Yen, J.H. & Barr, A.R. 1971. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature* **232**: 657–658.
- Zeh, D.W., Zeh, J.A. & Bonilla, M.M. 2005. *Wolbachia*, sex ratio bias and apparent male killing in the harlequin beetle riding pseudoscorpion. *Heredity*. **95**: 41–49.
- Zug, R. & Hammerstein, P. 2012. Still a Host of Hosts for *Wolbachia*: Analysis of Recent Data Suggests That 40% of Terrestrial Arthropod Species Are Infected. *PLoS One* **7**: e38544.

Annexes partie I

Chapitre I

ARTICLE

DOI: 10.1038/s41467-017-02749-w

OPEN

Culex pipiens crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*

Manon Bonneau¹, Celestine Atyame^{1,2}, Marwa Beji³, Fabienne Justy¹, Martin Cohen-Gonsaud⁴,
Mathieu Sicard¹ & Mylène Weill¹

Culex pipiens mosquitoes are infected with *Wolbachia* (wPip) that cause an important diversity of cytoplasmic incompatibilities (CIs). Functional transgenic studies have implicated the *cidA-cidB* operon from wPip and its homolog in wMel in CI between infected *Drosophila* males and uninfected females. However, the genetic basis of the CI diversity induced by different *Wolbachia* strains was unknown. We show here that the remarkable diversity of CI in the *C. pipiens* complex is due to the presence, in all tested wPip genomes, of several copies of the *cidA-cidB* operon, which undergoes diversification through recombination events. In 183 isofemale lines of *C. pipiens* collected worldwide, specific variations of the *cidA-cidB* gene repertoires are found to match crossing types. The diversification of *cidA-cidB* is consistent with the hypothesis of a toxin-antitoxin system in which the gene *cidB* co-diversifies with the gene *cidA*, particularly in putative domains of reciprocal interactions.

¹Institut des Sciences de l'Evolution de Montpellier (ISEM), UMR CNRS-IRD-EPHE-Université de Montpellier, Place Eugène Bataillon, 34095 Montpellier, France. ²Processus Infectieux en Milieu Insulaire Tropical (PIMIT), UMR CNRS-INSERM-IRD-Université de La Réunion, Sainte-Clotilde, Ile de La Réunion, 97490, France. ³Institut Pasteur Tunis, Laboratory of Epidemiology and Veterinary Microbiology, University of Tunis El Manar, 1068 Tunis, Tunisia. ⁴Centre de Biochimie Structurale (CBS), UMR CNRS-INSERM-Université de Montpellier, 29 rue de Navacelles, 34090 Montpellier, France. Mathieu Sicard and Mylène Weill contributed equally to this work. Correspondence and requests for materials should be addressed to M.S. (email: mathieu.sicard@umontpellier.fr) or to M.W. (email: mylene.weill@umontpellier.fr)

The most common way by which *Wolbachia* bacteria spread within insect populations is cytoplasmic incompatibility (CI), which results in the early death of embryos¹ when males infected with a given *Wolbachia* strain mate with uninfected females or females infected with an incompatible strain. The most popular model to date for conceptualizing CI involves the secretion, by *Wolbachia*, of a “modification” factor (mod, or toxin) in the sperm that impairs early embryogenesis, and a *Wolbachia* “rescue” factor (resc, or antitoxin) produced in the oocyte, which allow for a viable, diploid zygote to develop if the cross is compatible^{2–4}. Previous studies searching for *Wolbachia* effectors involved in CI by genomic comparison between mutualistic and manipulative *Wolbachia* revealed a higher number of genes with ankyrin repeats in the latter^{5, 6}. Ankyrin repeat genes were thus considered as potential mod and resc candidates due to their role in protein–protein interactions and cellular cycle regulation⁷. However, to date no correlation between the distributions of any ankyrin repeat genes and CI patterns was demonstrated⁸. The first convincing candidate gene for involvement in CI encodes a protein called CidA (WP0282), secreted by *Wolbachia* into *Culex pipiens* sperm⁹. The *cidA* gene is part of an operon also containing a second gene, *cidB* (WP0283). CidA and CidB proteins directly interact with each other as demonstrated by pull-down experiments¹⁰. The *cidA* and *cidB* genes (named *cifA* and *cifB* in *wMel*¹¹) are associated with the prophage WO modules¹¹, and paralogs of these genes have been found in all published genomes from *Wolbachia* strains known to induce CI in insects^{9, 11, 12}. The co-expression of *cidA/cidB* (or *cifA/cifB*) transgenes in *Drosophila melanogaster* reproduces the disturbance of the first embryonic division hallmark of CI, leading to embryo death^{10, 11}. In this context, disruption of the deubiquitylating (DUB) domain of CidB restores normal embryogenesis, demonstrating the contribution of this domain to the “mod” factor function¹⁰. The role of CidA in the mod-resc CI system is much more debated. Data obtained on yeast suggest that CidA may rescue CidB¹⁰, but CifA has also been described as a CifB elicitor in *D. melanogaster*¹¹.

cidA-cidB operon is involved in the CI induced by infected males in the progenies of uninfected females^{10, 11}, but the genetic basis of CI diversity (i.e., compatibility, uni or bidirectional incompatibility) between hosts infected with different *Wolbachia* strains has yet to be investigated. We addressed this question in *C. pipiens*, a powerful model in which hundreds of crosses between lines sampled worldwide have revealed unprecedented CI diversity^{13–15}. This unique CI diversity is solely governed by *wPip* diversification in this species complex since no other manipulative endosymbionts are present and no host genetic background influence has been yet demonstrated^{13, 15}. All *C. pipiens* individuals are infected with *Wolbachia*, which recently diverged into five distinct phylogenetic groups (*wPip*I to *wPip*V)¹⁶. The *wPip* strains of the same group generally ensure compatibility between their hosts, whereas intergroup crosses are more likely to be incompatible¹³. An analysis of multiple crosses concluded that each *Wolbachia* (*wPip*) genome must contain several mod and resc factors to account for the high diversity of CI in *C. pipiens*^{15, 17}. These multiple mod and resc factors could theoretically be encoded by different copies (i.e., variants) of same mod/resc genes or different mod and resc genes within the same *wPip* genome^{15, 17}. It therefore seems likely that (i) different copies (i.e., variants) of *cidA-cidB*, (ii) paralogs of these genes resulting from ancient duplication events^{10, 11, 17}, or (iii) totally unrelated genes¹³, could be involved in CI diversity in *C. pipiens*.

We thus investigated the contribution of polymorphism of the *cidA-cidB* operon and its *cinA-cinB* paralog with predicted nuclease activity¹⁰ to CI diversity in *C. pipiens* with the major hypothesis that if these genes encode for CI diversity determinants, they should be different between *C. pipiens* lines with different crossing types. We show that all tested *wPip* genomes contain several variants of the *cidA-cidB* operon while it is not the case for *cinA-cinB* operon. The *cidA-cidB* repertoires of variants differ between *wPip* strains from different groups exhibiting different crossing types. In 180 isofemale lines of *C. pipiens* exclusively infected with *wPip*IV strains, specific variations of the *cidA-cidB* gene repertoires are found to match crossing type

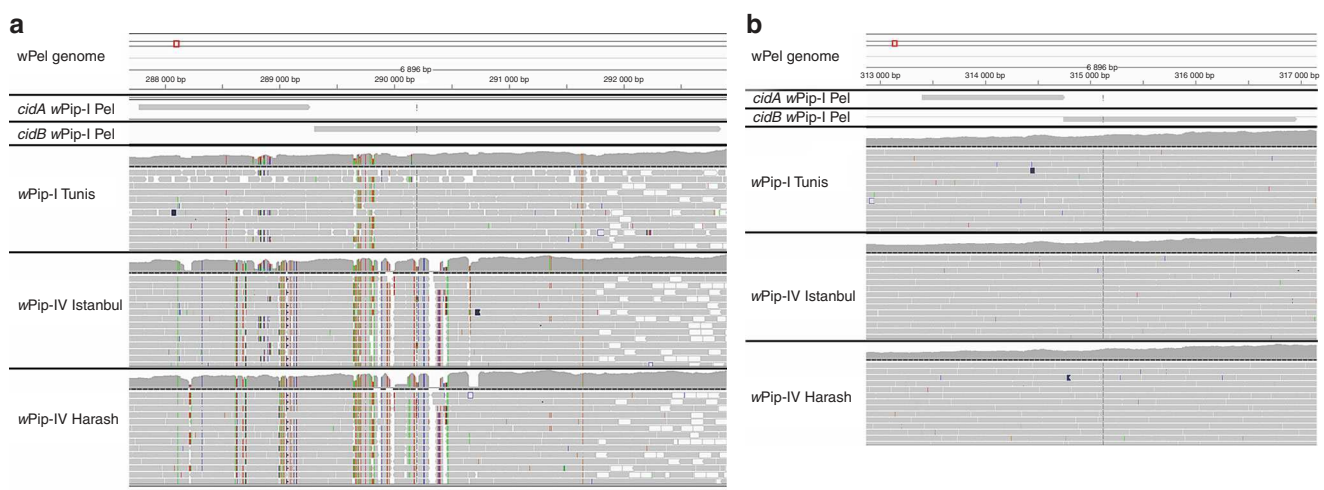


Fig. 1 The *cidA/cidB* operon displays polymorphism within and between different *wPip* strains, whereas no polymorphism is observed for the *cinA/cinB* operon. Mapping onto the reference genome *wPip_Pel* of the Illumina reads from the Tunis line infected with *wPip*I, and from the Harash and Istanbul lines infected with *wPip*IV. Colored residues are different from those in the reference *wPip_Pel* sequence. The pattern above the dotted lines represents the number of Illumina reads that have mapped for each position. The red box on the *wPip_Pel* reference genome allows the IGV user to locate the zone that is visualized with more details on the bottom panels. **a** Mapping of the *cidA/cidB* operon reads onto the *wPip_Pel* reference genome. Polymorphism between *wPip* groups was detected (differences between Harash and Istanbul infected with *wPip*IV, Tunis, and *wPip_Pel* infected with *wPip*I). Polymorphism was also detected within *Wolbachia* groups (variations observed between Tunis and *wPip_Pel* or between Harash and Istanbul). The operon was also found to be polymorphic within the same isofemale line of *C. pipiens*. **b** Mapping of the *cinA/cinB* operon reads onto the *wPip_Pel* reference genome. No polymorphism was detected between reads from the three isofemale lines and the reference line

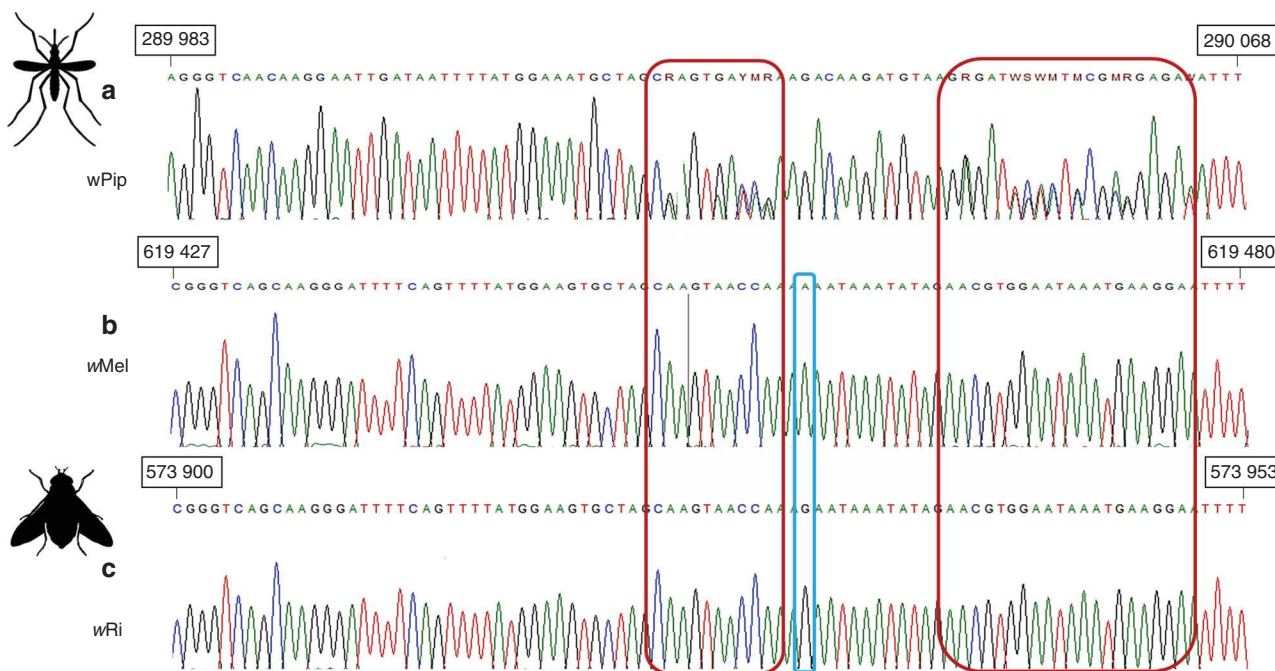


Fig. 2 Sanger sequencing electrophoregrams of homologous regions of **a** *cidB* in *wPip*, **b** *cifB* in *wMel*, **c** and *cifB* in *wRi*. The numbers in the black boxes indicate the positions of these regions in the *wMel*, *wRi*, and *wPip_Pel* genomes. The blue rectangle highlights the only polymorphism between *wRi* and *wMel*. Unlike the *Wolbachia* strains of *Drosophila*, *wPip* strains consistently give mixed signals (showed in red rectangles), suggesting the presence of at least two different sequences of this gene in DNA samples

Table 1 Crossing types of the four mosquito lines each infected with *Wolbachia wPip* from different groups (I, II, III, and IV)

♀ \ ♂	Tunis	Lavar	Maclo	Istanbul	Crossing type
Tunis (<i>wPip</i> I)	+	-	+	-	A
Lavar (<i>wPip</i> II)	+	+	+	-	B
Maclo (<i>wPip</i> III)	+	+	+	+	C
Istanbul (<i>wPip</i> IV)	-	-	-	+	D

+ means that the cross is compatible (many larvae hatch from the egg rafts) and - indicates an incompatible cross (no larvae hatch from the egg rafts). Each isofemale line has a different crossing type, as deduced from reciprocal crosses between the four lines.

variations. These variations occurred only in specific domains for both proteins, which consist exclusively of protein-protein interaction motifs and thus might be involved in the physical interaction between CidA and CidB.

Results

Illumina sequencing of several *wPip* strains. We first used Illumina technology to sequence three *wPip* genomes from strains inducing different crossing types¹³: two strains from group IV,

originating from Algeria and Turkey; and one strain from group I, originating from Tunisia. The mapping of the reads onto the reference genome *wPip_Pel*¹⁸ revealed an abnormally high coverage of the *WO*Pip1 region, which includes the *cidA* and *cidB* genes. The coverage for this region was, on average, three times greater than that for 14 single-copy housekeeping genes^{19, 20} (Supplementary Fig. 1). Moreover, the *cidA-cidB* operon was highly variable between and within the three isofemale lines sequenced (Fig. 1a). Contrary to the assembly of *wPip_Pel*

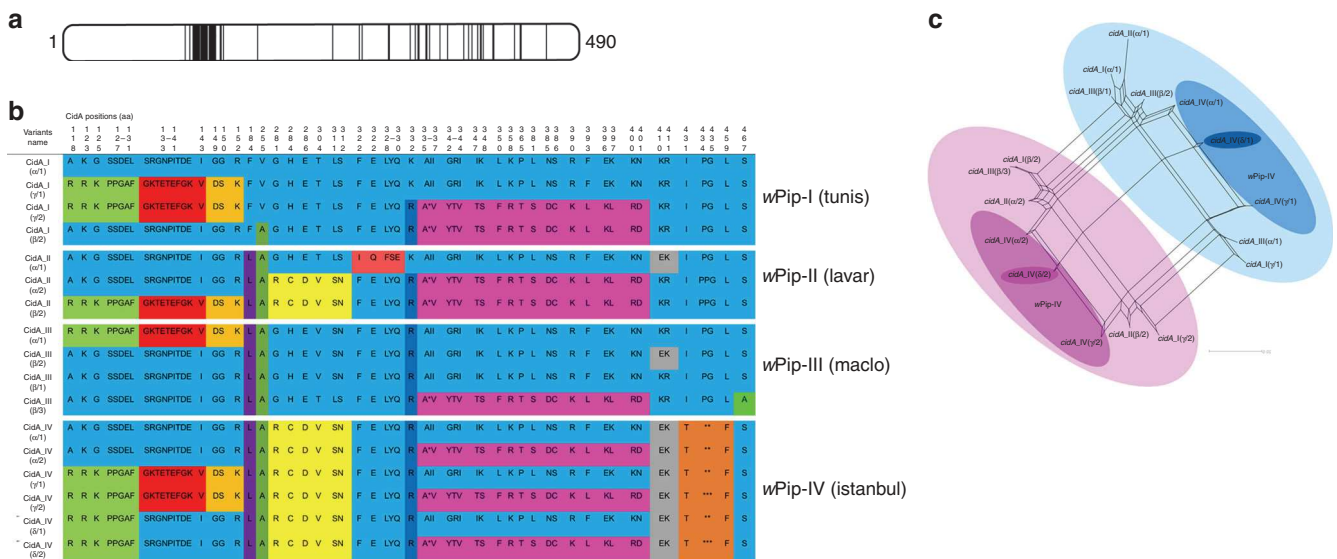


Fig. 3 The repertoire of CidA variants in the different wPip groups. **a** Localization of polymorphic zones within CidA proteins. Schematic representation of the CidA protein with polymorphic zones between CidA variants of the four wPip groups highlighted in black. **b** The polymorphism between CidA variants is distributed as blocks of variable amino acids. Protein sequences alignment of the CidA variants found in the four *Wolbachia* strains Tunis, Lavar, Maclo, and Istanbul. The first sequence is used as a reference to determine the polymorphic region. For more clarity, only polymorphic positions of the alignment are represented and amino-acid positions are not contiguous. When more than two contiguous amino acids were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic wPip group (I-IV). However, no variant (i.e., complete CidA sequence) is common to wPip strains from different groups. **c** *cidA* variants result from block recombination. Each edge (or set of parallel edges) corresponds to a split in the data set and has a length equal to the weight of the split. Incompatible splits produced by recombination are represented by boxes in the network. Most of the *cidA* variants are connected by multiple pathways resulting from block recombination between them. The largest circles highlight the two groups of *cidA* variants sharing the same amino-acid sequence from position 336 to 401 (**b**). Intermediate darker circles highlight *cidA* variants observed in the wPipIV group. The smallest bold circles highlight the two versions (1 and 2) of the *cidA_IV*(δ) variant matching the “incompatible” crossing type of mosquito lines infected with wPipIV (see the text)

genome, which exhibits only one copy of *cidA-cidB* operon¹⁸, our findings suggest the occurrence of several copies of the *cidA-cidB* operon within each wPip genome from Algeria, Tunisia, and Turkey. Consistent with this observation, the Sanger sequences of *cidA* and *cidB* obtained from single individuals displayed multiple overlaps (Fig. 2a). No such diversity was found in wMel and wRi from *D. melanogaster* and *D. simulans* (Figs. 2b, c). The *cinA* and *cinB* paralogs, which are also present in all sequenced wPip genomes and have been identified as possibly involved in incompatibility between *C. pipiens* lines¹⁰, displayed no polymorphism within and between wPip strains (Fig. 1b), ruling out their role in CI diversity¹⁰. These findings suggest a previously unsuspected diversity of *cidA-cidB* copies within wPip genomes, resulting from gene amplifications and divergence that may account for CI diversification in *C. pipiens*.

High *cidA-cidB* diversity between wPip groups. We explored the association between *cidA-cidB* genetic variants and CI diversity in *C. pipiens* further, by investigating the variability of these genes in four isofemale lines differing by their geographical origin, each infected with a wPip strain from a different group (I-IV) and exhibiting a different crossing type (Table 1). Each of these isofemale lines was founded with one initial egg-raft from a single female. Cloning and sequencing revealed a large number of *cidA-cidB* variants in each of these wPip strains. The variable domains that differ between variants were not contiguous but restricted to some specific zones of the proteins (Figs. 3a and 4a). Within a single line there were up to six different variants of *cidA* (Fig. 3b), and up to four different variants of *cidB* (Fig. 4b). Some specificity in the *cidA-cidB* repertoire was also detected, with none of the variants shared between strains from different wPip

groups (Figs. 3b and 4b). Network analyses of *cidA* and *cidB* suggested that the differences between copies stemmed mostly from block rearrangements within and/or between wPip genomes (Figs. 3c and 4c). These rearrangements were never observed in the region corresponding to the CidB catalytic domain, but in other regions of the CidA and CidB proteins (Figs. 3 and 4; Supplementary Fig. 2 and 3). The stability of the *cidA-cidB* repertoire over a decade (between 2006 and 2017) in addition with the stability in the *cidA* and *cidB* copy numbers relative to *wsp* gene (known to be present as a single copy in all *Wolbachia* genomes) inferred by quantitative PCR (q-PCR) in the Istanbul *C. pipiens* line are convincing elements that the *cidA-cidB* genes are present in multiple copies in wPip genome. The copy number ratio between *cidA-cidB* genes and *wsp* was always between 4 and 6 (Table 2). We also found that all copies of these genes were transcribed. Indeed, in the wPip strain from Istanbul the transcripts of the six polymorphic copies (i.e., variants) of *cidA* and the four polymorphic copies of *cidB* were detected in RNA extract after reverse transcription-PCR, cloning, and sequencing.

***cidA-cidB* repertoire variations are linked to crossing type.** We then focused on a simpler situation, by analyzing CI diversity within wPip group IV. We used six isofemale lines with two different crossing types. These lines had the same rescue properties but two different mods¹³. Males from “incompatible” lines sterilize females whenever they host a wPip group I, II, or III strain, whereas males from “compatible” lines always produce progenies with these females (Fig. 5). We compared the *cidA-cidB* repertoires of “compatible” and “incompatible” lines. We found that one CidA variant [*cidA_IV*(α)] was present in all wPipIV lines, regardless of their crossing type or geographic origin

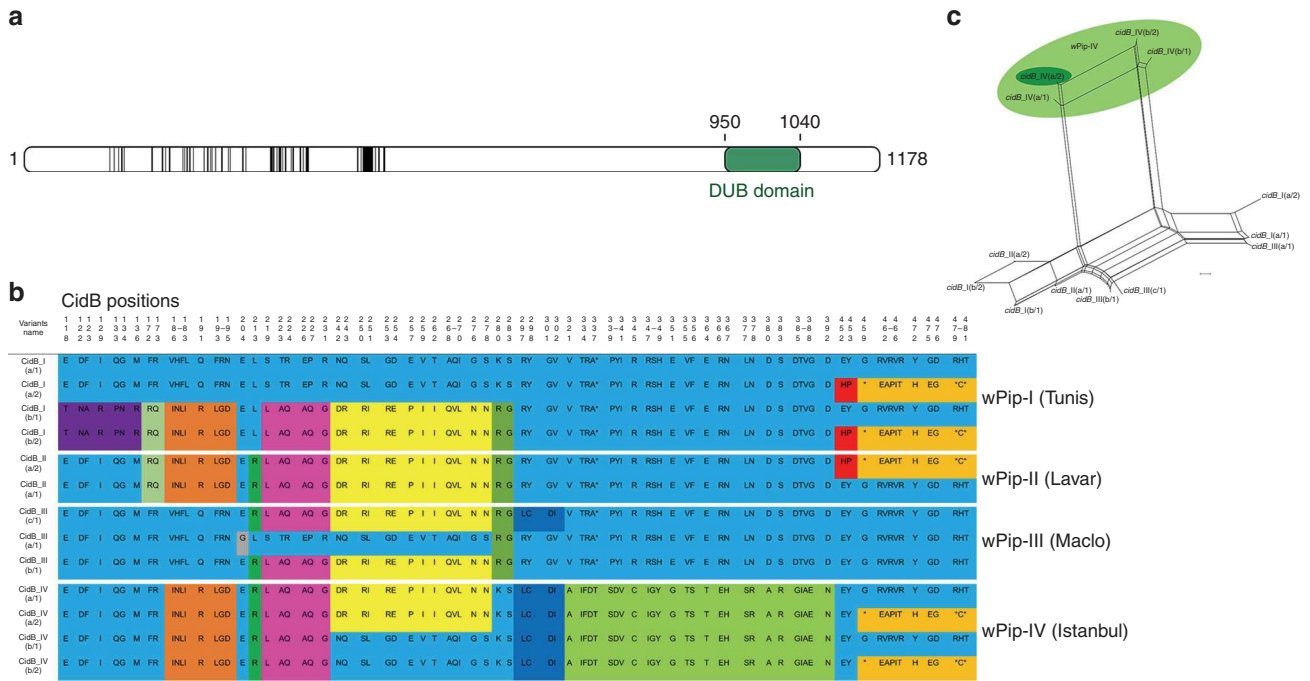


Fig. 4 The repertoire of CidB variants in the different wPip groups. **a** Localization of polymorphic zones within CidB proteins. Schematic representation of the CidB protein with polymorphic zones between CidB variants of the four wPip groups highlighted in black. The deubiquitylating protease domain (DUB) present in CidB, represented in green, displayed no polymorphism between wPip strains. **b** The polymorphism between CidB variants is distributed as blocks of variable amino acids. Protein sequences alignment of the CidB variants found in the four *Wolbachia* strains Tunis, Lavar, Maclo, and Istanbul. The first sequence is used as reference to determine the polymorphic region. For more clarity, only polymorphic positions of the alignment are represented and amino-acid positions are not contiguous. When more than two contiguous amino acids were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic wPip group (I–IV). However, no full variant (i.e., complete CidB sequence) is common to wPip strains from different groups. **c** *cidB* variants result from block recombination. Each edge (or set of parallel edges) corresponds to a split in the data set and has length equal to the weight of the split. Incompatible splits produced by recombination are represented by boxes in the network. Most of the *cidB* variants are connected by multiple pathways resulting from block recombination between them. The light green circle highlights *cidB*_{IV} variants, which cluster together due to their similar amino-acid sequences from position 321 to position 392 (**b**). The *cidB*_{IV(a/2)} variant matching the “incompatible” crossing type of mosquito lines infected with wPipIV (see the text) is highlighted by the darker green circle

Table 2 Quantification of *cidA-cidB* operon in different wPip genomes

Line	Tunis	Lavar	Maclo	Istanbul
	wPip-I	wPip-II	wPip-III	wPip-IV
<i>cidA/wsp</i>	5,99 ± 0,77	5,40 ± 0,84	5,1 ± 0,41	5,62 ± 0,72
<i>cidB/wsp</i>	6,21 ± 1,40	5,06 ± 1,46	4,17 ± 0,64	6,13 ± 1,26
<i>cidA/cidB</i>	0,99 ± 0,21	1,10 ± 0,19	1,24 ± 0,21	0,95 ± 0,19

The number of copies of the *cidA* and *cidB* genes was estimated by real-time q-PCR, relative to the reference gene *wsp* known to be present as a single copy per genome. The number of copies was determined by calculating the ratio between *cidA* (or *cidB*) signals and *wsp* signals on four individuals per line. ± means s.d. The *cidA* to *cidB* ratio was always close to one. The numbers of *cidA* and *cidB* copies present per *Wolbachia* genome were therefore similar, as expected for an operon.

(Fig. 5). Such a ubiquitous distribution is expected for the resc factor mediating compatibility between all isofemale lines hosting wPip strains from the same group. By contrast, no CidB variant was shared between compatible and incompatible lines (Fig. 5). Strikingly, *cidA* and *cidB* repertoire variations were clearly associated with crossing type variations, as *cidA*_{IV(δ)} and *cidB*_{IV}

(a/2) variants were found uniquely in “incompatible” lines (Fig. 5). This may suggest that both *cidA*_{IV(δ)} and *cidB*_{IV(a/2)} are required for the incompatibility phenotype (Fig. 5).

We further investigated the distribution of these two specific *cidA/cidB* variants [*cidA*_{IV(δ)} and *cidB*_{IV(a/2)}] at a larger scale, by performing diagnostic restricted fragment length

a

Line name	Crossing type	Polymorphism location	<i>cidA_IV</i>				
			Upstream	α	β	γ	δ
			Downstream	1/2	1/2	1/2	1/2
Istanbul	Incompatible		P	A	P	P	
Hang Zhou	Incompatible		P	A	A/P	P	
Ichkeul 09	Incompatible		P	A	A	P	
Ichkeul 21	Incompatible		P	A	A	P	
Harash	Compatible		P	P	A	A	
Ichkeul 13	Compatible		P	P/A	A	A	

b

Line name	Crossing type	<i>cidB_IV</i>					
		a1	a2	a3	b1	b2	b3
Istanbul	Incompatible	P	P	A	P	P	A
Hang Zhou	Incompatible	P	P	P	P	A	P
Ichkeul 09	Incompatible	P	P	A	A	A	A
Ichkeul 21	Incompatible	P	P	P	P	A	P
Harash	Compatible	P	A	P	P	A	P
Ichkeul 13	Compatible	A	A	A	P	A	P

Fig. 5 *cidA_IV*(δ) and *cidB_IV*(a/2) are present specifically in lines infected with wPipIV strains displaying the “incompatible” crossing type. The “incompatible” or “compatible” crossing type of the *C. pipiens* lines infected with wPipIV strains were determined by crossing males with females infected with wPipI, wPipII, or wPipIII. A “compatible” type corresponds to a line in which all males are compatible when crossed with all tested females, and an “incompatible” type corresponds to a line in which males are incompatible when crossed with all tested females. The column of the two variants *cidA_IV*(δ) and *cidB_IV*(a/2) associated with the “incompatible” crossing type were shaded in green. **a** Distribution of *cidA* variants (P for present, A for absent) in six wPipIV strains. Each column corresponds to the specific upstream polymorphic region named α , β , γ , or δ . The downstream polymorphism is represented by either the sequence 1 or the sequence 2 (1/2). For a given upstream sequence P/A means that the strain exhibits the sequence 1 but not the sequence 2. On the opposite A/P means that the strain exhibits the sequence 2 but not the sequence 1. **b** Distribution of *cidB* variants (P for present, A for absent) in six wPipIV strains

polymorphism (RFLP) tests on 180 isofemale lines from 15 populations [13 in Algeria and Tunisia; 1 each in China and Turkey], for which we determined crossing types. A very strong association was established between the presence of both *cidA_IV*(δ) and *cidB_IV*(a/2) in “incompatible” isofemale lines and their absence from “compatible” isofemale lines: all 17 “incompatible” lines had both these variants, whereas 147 of the 163 “compatible” lines had neither of them ($\chi^2 = 78$, $df = 1$, $p < 2.2e-16$). Among the 16 “compatible” isofemale that appeared discordant: (i) 8 were truly discordant as they display both *cidA_IV*(δ) and *cidB_IV*(a/2) while being compatible but (ii) 8 were only discordant for *cidA* as they displayed *cidA_IV*(δ) but lacked *cidB_IV*(a/2) (as confirmed by cloning and sequencing; Supplementary Data 1). To summarize the data: the absence of these variants is always associated with a “compatible” crossing type while its presence is mostly associated with “incompatible” crossing type with few discrepancies.

Putative ankyrin interaction domains between CidA and CidB.

The analysis of CidA and CidB polymorphic regions revealed that the sequence variations associated with crossing type variations are restricted to about 30 amino-acid positions, for both proteins (Fig. 6a). Extensive fold-recognition analysis of the CidA and CidB proteins were performed using the @tome2 in-house server²¹. Except for the DUB protease domain, no tridimensional model could be firmly obtained. Nevertheless, both CidA and CidB exhibit sequence identity homology (under 20%) with protein–protein interaction domains with repeated structure. Careful examination of hydrophobic patterns suggests internal helical structure repetitions as found in ankyrin or HEAT repeats²². The size of ~33 amino acids is the universal expected

size for tandem ankyrin repeat domains⁶. However, due to the lack of available homolog sequences in databases, it is difficult to determine the helical structure repetition boundaries in both N- and C-terminal part of the protein domains. As the protein variations matching with crossing type variations occur in putative ankyrin domains, we hypothesize that CidA and CidB co-diversify in these domains because they are involved in their reciprocal interactions (Fig. 6).

Discussion

In the *C. pipiens* mosquitoes, the genetic architecture of the *Wolbachia cidA-cidB* operon is characterized by variable multicopies associated to the unique CI diversification observed in this species. These different copies of *cidA* and *cidB* likely result from genic amplification followed by diversification since we revealed a high diversity within each wPip strain. As these genes are associated with the prophage WO modules¹¹, the number of *cidA-cidB* copies assessed by q-PCR may not reflect the exact number of *cidA-cidB* copies in a single wPip genome since viral particle multiplication resulting from lytic phage activity can influence the accuracy of this quantification²³. However, the diversity of *cidA-cidB* sequences in a single mosquito, the geographical and temporal stability of sequences found in wPipIV group clearly point out that several copies of the operon do co-exist within the same wPip genome. As *cidA* and *cidB* are amplified in all wPip genomes from each *C. pipiens* individual we sequenced so far, the presence of only one copy of the *cidA/cidB* operon in wPip_Pel seems to result from an error during assembly likely caused by too many repeats. The reason(s) why such amplification and diversification of *cidA/cidB* has evolved in wPip is still an open

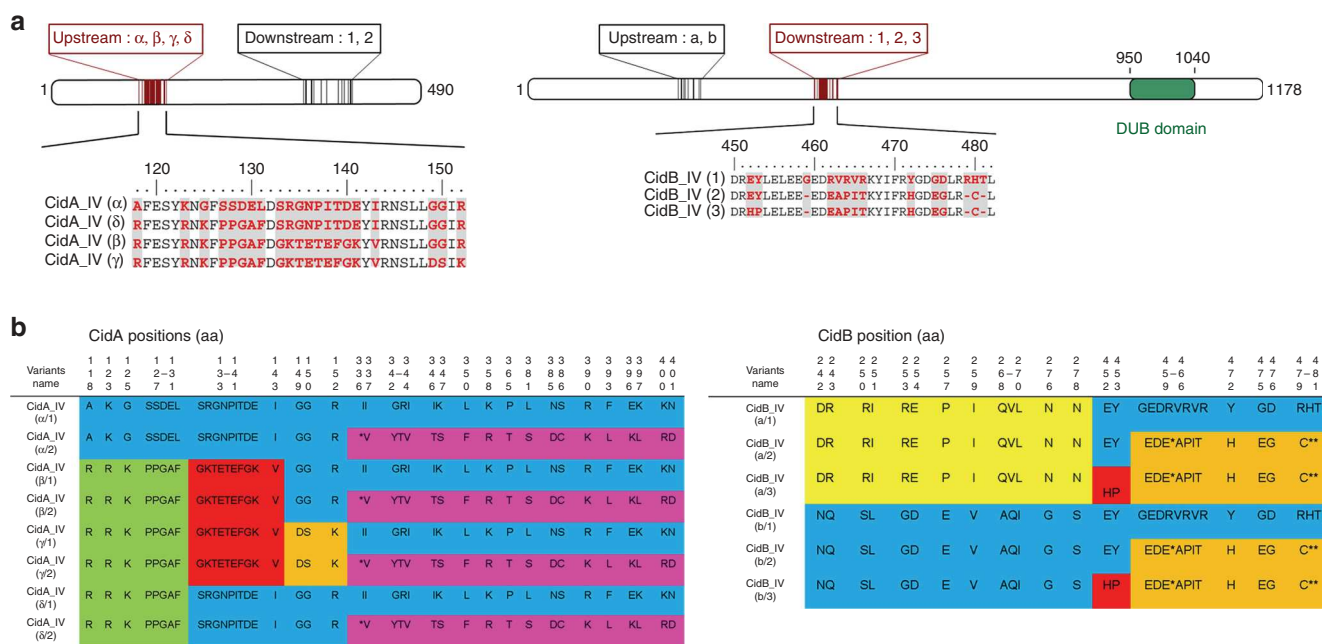


Fig. 6 The putative domains of interaction between the CidA_IV and CidB_IV proteins. **a** Localization of polymorphic zones within CidA_IV and CidB_IV proteins. Secondary structure prediction and fold-recognition analysis predicted that the CidA protein could almost exclusively display protein/protein interaction repeated motifs as ankyrin- (or HEAT-like repeats). Similar domains of ~30 amino acids (expected size for ankyrin domains) are also present in CidB. The deubiquitylating protease domain (DUB) present in CidB, represented in green, displayed no polymorphism between wPipIV strains as already observed between wPip strains from different groups. The polymorphism of the wPipIV CidA and CidB variants associated with crossing type (“compatible” vs. “incompatible”, see Fig. 5 for definition) is highlighted in red. The corresponding variations in protein sequences are reported below the schematic representation of the full-length proteins, with the variable residues highlighted in red. Other sequence variations not predicted to be involved in CI variations are shown in black. **b** CidA_IV and CidB_IV variants sequences detected in all sequenced wPip strains from group IV. CidA_IV variants display two regions of polymorphism resulting from recombination revealed by block colors: the upstream one from 118 to 152 aa and the downstream one from 336 to 401 aa. Four possible sequences (noted α, β, γ , or δ) were found in the upstream polymorphic region followed by one of the two sequences in the downstream region (noted 1 or 2). Only the upstream polymorphic region of CidA is associated with crossing type variations. CidB_IV variants display two regions of polymorphism resulting from recombination: the upstream one from 242 to 278 aa and the second one from 450 to 481 aa. Two possible sequences (noted a and b) were found in the upstream polymorphic region followed by one of the three possible sequences in the downstream polymorphic region (noted 1, 2, and 3). Only the downstream region is associated with crossing type variations

question. However, this diversity likely reflects high levels of intra- or intergenome recombination, which may be favored by multiple infections with different wPip and/or promoted by lytic phage activity in *C. pipiens*²⁴.

In a more global context, our results provide insights into the respective roles of CidA and CidB in the mod/resc model. Indeed, the association between *cidA_IV*(δ)-*cidB_IV*(a/2) presence/absence and incompatible/compatible crossing types, together with the results of functional studies^{10, 11}, demonstrate that CidB is involved in the mod function responsible for the variation in crossing types. However, some discrepancies were found in few isofemale lines, which harbor *cidA_IV*(δ) and/or *cidB_IV*(a/2) but are “compatible”. These discrepancies suggest that either (i) our strategy, involving larval field sampling, the establishment of isofemale lines and crossing experiments, generated a few errors or (ii) other genetic or epigenetic mechanisms may disrupt the relationship between genomic potential and the resulting crossing type.

The role of CidA is still unclear. The association of *cidA_IV*(δ) with variation in crossing types does not necessarily implicate CidA in the mod function. Indeed, as CidA and CidB interact with each other¹⁰, their co-diversification could result in specific CidA variations whenever involved in the mod (i.e., toxin) or in resc (i.e., antitoxin) function. Several lines of evidence suggest that CidA could be the antitoxin of CidB and that the *cidA-cidB* operon encodes a classic toxin–antitoxin system: (i) the simple two-gene structure of the operon, with the putative antitoxin gene

located upstream from the putative toxin gene⁹; (ii) the over-expression of the antitoxin gene relative to the toxin gene¹¹; (iii) the ability of the two proteins to form a complex¹⁰; and (iv) the required co-expression of *cidA* and *cidB* for the production of live transgenic *D. melanogaster* and *Saccharomyces cerevisiae*¹⁰. Additional argument of CidA being an antitoxin is provided by the presence of *cidA_IV*(α), a ubiquitous CidA variant in wPip strains within the same group wPipIV, potentially accounting for their hosts’ compatibility. Moreover, if CidA is indeed the antitoxin, then a new toxin (i.e., CidB) produced by block recombination between variants would be positively selected only if a compatible CidA, capable of interaction with this new toxin, was already present in the same wPip genome to prevent its producer from being killed^{2, 25}. This is coherent with our results since we found more *cidA* variants than *cidB* variants in the same isofemale line (except for Tunis, which has as many *cidA* as *cidB*).

Most toxin–antitoxin systems require an interaction between both proteins²⁵. Pull-down experiments have demonstrated that CidA can interact with CidB¹⁰. Our analysis of CidA–CidB repertoires in wPipIV group showed that genetic variations linked to crossing type variations occurred only in specific domains of about ~30 amino acids in length, for both proteins (Fig. 6a). These domains consist exclusively of protein–protein interaction motifs (ankyrin or HEAT repeats) and thus might be involved in the reciprocal interaction between CidA and CidB. The variability of these domains that co-diversify may account for specific interaction between one variant of CidA and one variant of CidB.

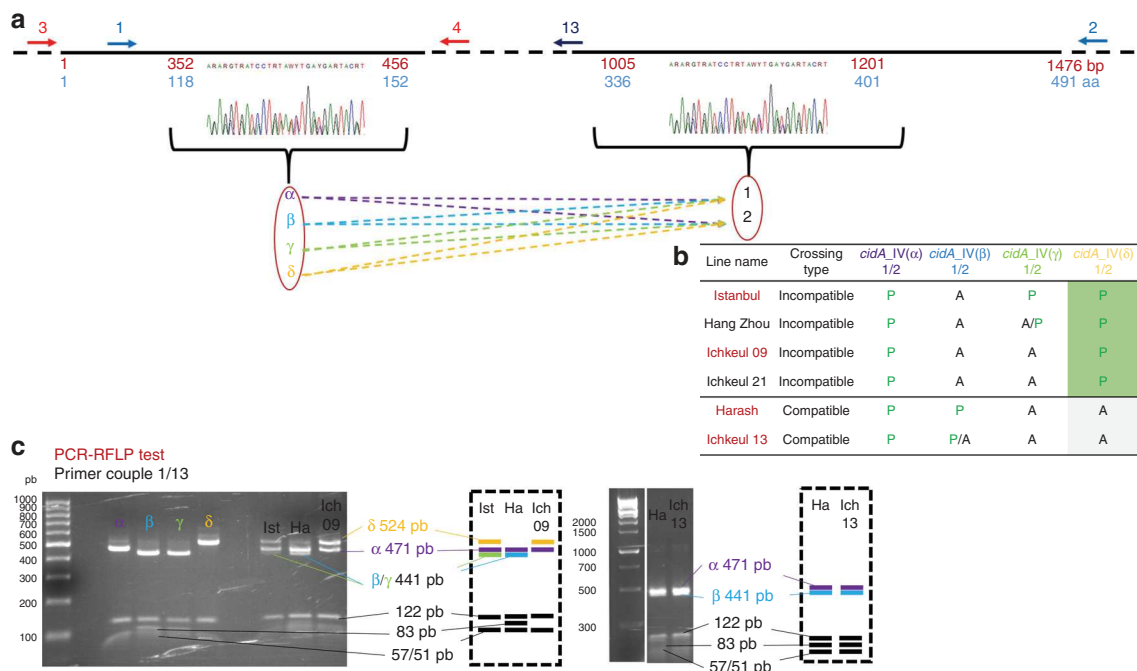


Fig. 7 *cidA_IV* gene polymorphic regions and PCR-RFLP tests for specific variants. **a** Schematic representation of the architecture of *cidA_IV* polymorphism. The black line represents the *cidA_IV* sequence with a focus on the two polymorphic regions. Non-polymorphic regions were shortened and represented as a dashed line. Numbers in red under the line represent nucleotide positions, and the numbers in blue indicate amino-acid positions. The overlapping oligonucleotides used for PCR amplification are represented by arrows numbered as in Supplementary Table 2. Both polymorphic regions were studied, but only the upstream matched the compatibility profile; its four different sequences (α , β , γ , and δ) were followed by one of the two possible sequences (1 or 2) in the downstream polymorphic region. A different color code was used for α (purple), β (light blue), γ (light green), and δ (yellow) sequences in the upstream part of the *cidA* gene. **b** The repertoire of *cidA_IV* variants is different in “compatible” and “incompatible” lines. *cidA_IV*(δ) is present only in lines with “incompatible” crossing type. The names of the *C. pipiens* lines used to set-up the PCR-RFLP (**c**) test are highlighted in red. **c** PCR-RFLP tests for distinguishing between *cidA_IV* variants on the basis of the upstream polymorphic region. A 778 bp fragment was amplified with primers 1/13. Double digestion with *ApoI* and *Hpy188I* distinguished between *cidA_IV*(α) (six fragments: 471; 122; 57; 53; 51; and 24 bp), *cidA_IV*(β) and *cidA_IV*(γ) (six fragments: 441; 122; 83; 57; 51; and 24 bp), and *cidA_IV*(δ) (five fragments: 524; 122; 57; 51; and 24 bp). Panel to the left of the electrophoresis gel: PCR-RFLP on DNA from clones; right panel: PCR-RFLP on DNA from the Istanbul, Harash, Ichkeul 13, and Ichkeul 09 lines. On the right of the gel, a schematic representation of the PCR-RFLP profiles of these lines. Digestion bands that are specific of the variants are represented with the color code established in **a**. Bands that are not used to discriminate between the variants are represented in black on the bottom of the schematic gel

Such variability in interactions between CidA and CidB could drive the unrivaled diversity of crossing types in *C. pipiens*.

Methods

Isofemale line construction. *C. pipiens* larvae and pupae were collected in the field and reared to adulthood in the laboratory. Females were then fed on blood to lay eggs that served to establish isofemale lines. Each egg-raft (containing 100–300 eggs) was individually isolated from hatching, and the *Wolbachia* group present was determined by performing *pk1* PCR-RFLP tests²⁶ on two first-instar larvae (L1). Isofemale lines were created by rearing the offspring resulting from a single egg-raft (thus from a single female). We established 162 isofemale lines for this study. We also used 21 isofemale lines from laboratory stocks of various geographic origins (Supplementary Data 1 and Supplementary Table 1). Isofemale lines were reared in 65 dm³ screened cages kept in a single room at 22–25 °C, under a 12 h light/12 h dark cycle. Larvae were fed with a mixture of shrimp powder and rabbit pellets, and adults were fed on honey solution.

Crossing type determination. Crossing types were characterized by crossing males (25–50 virgin males) from each of the studied isofemale lines with females (25–50 virgin females) from four reference laboratory isofemale lines hosting different *Wolbachia* strains [Tunis (*wPipI*), Lavar (*wPipII*), Macló (*wPipIII*), and Istanbul (*wPipIV*)] or with females from Tunis only (Supplementary Data 1 and Supplementary Table 1).

Illumina sequencing. The genomes of three *Wolbachia* strains from three *C. pipiens* isofemale lines were fully sequenced: Tunis from group *wPipI*, and Harash and Istanbul from group *wPipIV*. *Wolbachia* is an intracellular uncultivable bacterium. To obtain satisfactory amounts of *Wolbachia* DNA, we therefore developed a protocol based on that described by Ellegaard et al.²⁷. This protocol involved *Wolbachia* enrichment on egg rafts immediately after oviposition, to maximize the

ratio of *Wolbachia* to *Culex* genomes. The freshly laid eggs-rafts were washed in bleach to crack their chorion. They were then rinsed in water and homogenized in phosphate-buffered saline (1× PBS). *Wolbachia* cells were isolated from each egg-raft separately and concentrated. Each egg-raft was crushed with a sterile pestle and the resulting suspension was centrifuged at 400 × *g* for 5 min at 4 °C to remove the cell debris. The supernatant was then centrifuged again at 6000 × *g* for 5 min at 4 °C, to obtain a pellet of *Wolbachia* cells. These cells were resuspended in 1× PBS and passed through a filter with 5 μm pores (GVS Filter Technology), and a filter with 2.7 μm pores (Whatman), to remove any remaining particles bigger than *Wolbachia*. The resulting filtrate was centrifuged at 6900 × *g* for 15 min at 4 °C to obtain a pellet of *Wolbachia* cells. Multiple-displacement amplification was carried out directly on the *Wolbachia* pellet, with the Repli-g- mini kit (Qiagen). For each strain, five amplifications were performed independently, to randomize potential amplification errors, and the amplicons were then pooled for sequencing. Illumina sequencing was performed by the 3 kb Long Jumping Distance Mate-Pair method, on a HiSeq2000 instrument (Eurofins MWG Operon platform in Ebersberg, Germany), generating 2 × 100 bp sequences for each fragment.

Cloning and Sanger sequencing of the *cidA* and *cidB* genes. For all experiments, total DNA was extracted with the Dneasy Blood & Tissue Spin-Column kit (Qiagen; bench protocol: animal tissues). For amplification of the full-length *cidA* gene, we used two pairs of overlapping primers, 1/2 and 3/4 (Fig. 7, Supplementary Table 2). For amplification of the full-length *cidB* gene, we used four pairs of overlapping primers: 5/6; 7/8; 9/10; and 11/12 (Fig. 8, Supplementary Table 2). We used *cidA* fragments amplified with primers 1/2 and *cidB* fragments amplified with primers 7/8 for cloning experiments, because these two regions contain all the polymorphism observed in *cidA* and *cidB* variants (Figs. 3 and 4). PCR products were extracted from the electrophoresis gel with the W-tracta disposable gel extraction tool (Star Lab) and purified with the QIAquick Gel Extraction Kit (Qiagen). The TOPO TA cloning Kit pCR 2.1-TOPO Vector (Invitrogen) was used for ligation. Before electroporation, the DNA was dialyzed with VSWP membrane

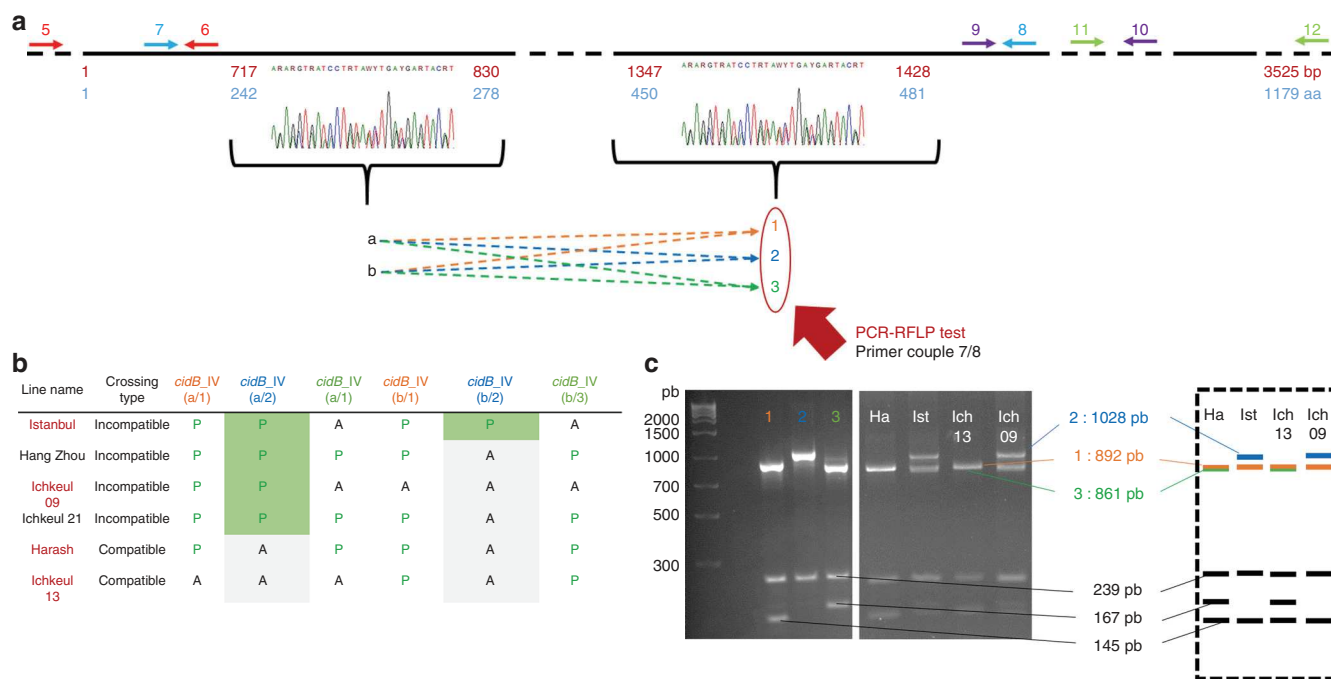


Fig. 8 *cidB_IV* gene polymorphic regions and RFLP tests for specific variants. **a** Schematic representation of the polymorphism architecture of *cidB_IV*. The black line represents the *cidB_IV* sequence with a focus on the polymorphic region. Non-polymorphic regions were shortened and represented as a dashed line. Numbers in red represent nucleotide positions and those in blue indicate amino-acid positions. The overlapping oligonucleotides used for amplification are represented by arrows numbered as in Supplementary Table 2. Both polymorphic regions are indicated but only the downstream (on the left, base pairs 1347–1428, amino acids 450–481) matches the compatibility profiles. The two sequences of the upstream polymorphic zone (a and b) are followed by one of the three possible sequences (1, 2, and 3) in the downstream polymorphic region. A different color code was used for 1 (orange), 2 (blue), and 3 (green) different sequences in the downstream part of the *cidB* gene. **b** The *cidB_IV* variant repertoire differs between “compatible” and “incompatible” lines. *cidB_IV*(a/2) is present only in lines with “incompatible” crossing type. The names of the *C. pipiens* lines used to set up the PCR-RFLP (**c**) test are highlighted in red. **c** PCR-RFLP tests for distinguishing between *cidB_IV* variants on the basis of the second polymorphic region. A 1267–1276 bp fragment was amplified with primers 7/8. Double digestion with *BanI* and *TaqI* distinguished between *cidB_IV*(1) (three fragments: 892; 239; and 145 bp), *cidB_IV*(2) (two fragments: 1028 and 239 bp), and *cidB_IV*(3) (three fragments: 861; 239; and 167 bp). Panel to the left of the electrophoresis gel: PCR-RFLP on DNA from clones; right panel: PCR-RFLP on DNA from lines. The panel on the right shows schematic representations of the PCR-RFLP profiles for these lines. Digestion bands that are specific of the variants are represented with the color code established in the **a**. Bands that are not used in to discriminate between the variants are represented in black on the bottom of the schematic gel

filters with 0.025 μm pores (Merck Millipore) to remove excess salt from the ligation. Electroporation was performed with One Shot TOP10 Electrocomp *Escherichia coli* (Invitrogen) and a Bio-Rad Micropulser with 2 mm-path electroporation cuvettes (Eurogentec) according to the manufacturer’s instructions. For repertoire acquisition of *cidA* and *cidB*, at least 48 clones for each individual were amplified and sequenced with M13 primers. Once the PCR-RFLP tests (see below) had been developed, more sequences were obtained for the variants initially underrepresented. For all the *C. pipiens* isofemale lines analyzed by cloning-sequencing, direct Sanger sequencing results for *cidA* and *cidB* were compared with the repertoire obtained by cloning-sequencing, to ensure that no variants were missing. For Sanger sequencing, the PCR products were purified with the Agencourt Ampure PCR purification kit (Agencourt) and directly sequenced with an ABI Prism 3130 sequencer using the BigDye Terminator Kit (Applied Biosystems).

Sanger sequencing data analyses. Sequence variants of *cidA* and *cidB* were aligned, with Muscle implemented in Seaview 6.4.1 software²⁸. Alignments were then analyzed within a phylogenetic network framework, to account for potentially conflicting signals due to recombination. A phylogenetic network was constructed from uncorrected *P* distances by the neighbor-net method²⁹ implemented in SplitsTree³⁰.

PCR-RFLP tests. We designed PCR-RFLP tests to detect the presence of the *cidA_IV*(δ) [*cidA_IV*(δ/1) and *cidA_IV*(δ/2)] and *cidB_IV*(2) [*cidB_IV*(a/2) and *cidB_IV*(b/2)] variants in isofemale lines. For *cidA_IV* variants, only the first polymorphic region was correlated with crossing types. We therefore designed the PCR-RFLP test to distinguish between the *cidA_IV*(α), *cidA_IV*(β), *cidA_IV*(γ), and *cidA_IV*(δ) sequences (Fig. 7). A 778 bp fragment was amplified with the 1/13 primer pair (Supplementary Table 2). Double digestion of the PCR products with *ApoI* and *Hpy188I* (New England Biolabs) identified three groups of sequencing: *cidA_IV*(α) (six fragments: 471; 122; 57; 53; 51; and 24 bp); *cidA_IV*(β) and

cidA_IV(γ) (six fragments: 441; 123; 83; 57; 51; and 24 bp); and *cidA_IV*(δ) (five fragments: 524; 122; 57; 51; and 24 bp).

For *cidB_IV* variants, only the second polymorphic region was found to be correlated with crossing types, so the PCR-RFLP test was designed to discriminate between the *cidB_IV*(1), *cidB_IV*(2), and *cidB_IV*(3) sequences (Fig. 8). A 1267–1276 bp fragment was amplified with the 7/8 primer pair (Supplementary Table 2). Double digestion of the PCR products with *BanI* and *TaqI* (New England Biolabs) distinguished three different groups of sequences: *cidB_IV*(1) (three fragments: 892, 239, 145 bp); *cidB_IV*(2) (two fragments: 1028 and 239 bp); and *cidB_IV*(3) (three fragments: 861; 239; and 167 bp).

Gene copy number quantification. The copy numbers of the *cidA* and *cidB* genes were estimated per *Wolbachia* genome, for four laboratory lines, on male and female adults and a pool of larvae at the earliest stage, by real-time q-PCR with a Roche Light Cycler. Three PCRs were performed, in triplicate, on each sample, one for the *Wolbachia wsp* locus present as single copy per genome (primer pair 18/19), one for *cidA* (primer pair 14/15), and one for *cidB* (primer pair 16/17). These primer pairs specifically bind to a region conserved in all wPip groups for each gene (Supplementary Table 2). The ratio between *cidA* or *cidB* and *wsp* signals was used to estimate the relative number of copies of *cidA* or *cidB* per *Wolbachia* genome.

***cidA-cidB* expressed repertoire.** Fresh mosquito samples from adult females and males and first (L1) and last (L4) instar larvae were used for RNA extraction with Trizol (Life Technologies). The RNA obtained was treated with DNase with the TURBO DNA-free Kit (Life Technologies), in accordance with the manufacturer’s instructions. We reverse-transcribed 2–5 μg of each total RNA sample with the SuperScript III Reverse Transcriptase Kit and 30 ng of random oligomer primers ((RP)10; Invitrogen, Life Technologies).

The polymorphic zones of the *cidA* and *cidB* cDNAs were amplified, cloned and 48 clones were sequenced, as described above.

Statistical analyses. All statistical analyses were performed with R version 3.0.2 software. We compared the proportions of “compatible” and “incompatible” isofemales harboring *cidA_IV*(δ) and *cidB_IV*(a/2) in a two-sample test for equality of proportions with continuity correction based on χ^2 .

Mapping of illumina reads on the reference genome. Illumina Long Jumping Distance sequencing data were trimmed with Trimmomatic version 0.36.

Paired reads and single reads for each line were compiled into Fastq files and mapped onto the reference genome wPip_Pel, with bwa mem and default parameters. Mappings were visualized with IGV version 2.3.68.

For each *Wolbachia* strain, the coverages of WP0282 (*cidA*), WP0283 (*cidB*), WP0294 (*cinA*), WP0295 (*cinB*), and MLST genes^{19, 31} were calculated with genomecov from bedtools, with the -d option, to calculate “per-base” genome coverage. Several adjustments were made: (i) coverage was normalized by dividing by the mean coverage on the whole reference genome of the strain concerned, to eliminate bias due to the higher coverage for strains for which there were larger numbers of reads in the Illumina data set; (ii) polymorphic regions of *cidA* and *cidB* were removed, due to the considerable variability of coverage in these regions; and (iii) the region of *cidB* from position 2461 to the end of the gene was removed because this region is identically duplicated in the wPip_Pel genome⁵, artificially increasing its coverage.

Data availability. The Illumina raw reads data of the three wPip genomes have been deposited in the Sequence Read Archive with Bioproject PRJNA393398 and BioSample SAMN07315580 (wPipI Tunis), SAMN07327402 (wPipIV Harash), and SAMN07327406 (wPipIV Istanbul). Nucleotide and proteic sequences of *cidA-cidB* variants have been deposited in GenBank and accession numbers are available in the Supplementary Table 3. The authors declare that all other data supporting the findings of this study are available within the article and its [Supplementary Information](#) files, or are available from the authors upon request.

Received: 17 August 2017 Accepted: 21 December 2017

Published online: 22 January 2018

References

- Yen, J. H. & Barr, A. R. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature* **232**, 657–658 (1971).
- Poinsot, D., Charlat, S. & Mercot, H. On the mechanism of *Wolbachia*-induced cytoplasmic incompatibility: confronting the models with the facts. *BioEssays* **25**, 259–265 (2003).
- Hurst, L. D. The evolution of cytoplasmic incompatibility or when spite can be successful. *J. Theor. Biol.* **148**, 269–277 (1991).
- Werren, J. H. Biology of *Wolbachia*. *Annu. Rev. Entomol.* **42**, 587–609 (1997).
- Foster, J. et al. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* **3**, 0599–0614 (2005).
- Siozios, S. et al. The diversity and evolution of *Wolbachia* ankyrin repeat domain genes. *PLoS ONE* **8**, e55390 (2013).
- Sedgwick, S. G. & Smerdon, S. J. The ankyrin repeat: a diversity of interaction on a common structural framework. *Trends Biochem. Sci.* **24**, 311–316 (1999).
- Duron, O. et al. Variability and expression of ankyrin domain genes in *Wolbachia* variants infecting the mosquito *Culex pipiens*. *J. Bacteriol.* **189**, 4442–4448 (2007).
- Beckmann, J. F. & Fallon, A. M. Detection of the *Wolbachia* protein WPiP0282 in mosquito spermathecae: Implications for cytoplasmic incompatibility. *Insect Biochem. Mol. Biol.* **43**, 867–878 (2013).
- Beckmann, J. F., Ronau, J. A., Hochstrasser, M., Tillett, D. & Ginalski, K. A. *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* **2**, 17007 (2017).
- Lepage, D. P. et al. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* **543**, 243–247 (2017).
- Sutton, E. R., Harris, S. R., Parkhill, J. & Sinkins, S. P. Comparative genome analysis of *Wolbachia* strain wAu. *BMC Genomics* **15**, 928 (2014).
- Atyame, C. M. et al. *Wolbachia* divergence and the evolution of cytoplasmic incompatibility in *Culex pipiens*. *PLoS ONE* **9**, 21–26 (2014).
- Laven, H., Wright, J. W. & Pal, R. In *Genetics of Insect Vectors of Disease* (eds Wright, J. W. & Pal, R.) 251–275 (Elsevier, Amsterdam, 1967).
- Atyame, C. M. et al. Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito populations. *Mol. Ecol.* **20**, 286–298 (2011).
- Atyame, C. M., Delsuc, F., Pasteur, N., Weill, M. & Duron, O. Diversification of *Wolbachia* endosymbiont in the *Culex pipiens* mosquito. *Mol. Biol. Evol.* **28**, 2761–2772 (2011).
- Nor, I. et al. On the genetic architecture of cytoplasmic incompatibility: inference from phenotypic data. *Am. Nat.* **182**, E15–E24 (2013).
- Klasson, L. et al. Genome evolution of *Wolbachia* strain wPip from the *Culex pipiens* group. *Mol. Biol. Evol.* **25**, 1877–1887 (2008).
- Baldo, L. et al. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl. Environ. Microbiol.* **72**, 7098–7110 (2006).
- Lo, N. et al. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. *Int. J. Syst. Evol. Microbiol.* **57**, 654–657 (2007).
- Pons, J. L. & Labesse, G. @TOME-2: A new pipeline for comparative modeling of protein-ligand complexes. *Nucleic Acids Res.* **37**, 485–491 (2009).
- Andrade, M. A., Perez-Iratxeta, C. & Ponting, C. P. Protein repeats: structures, functions, and evolution. *J. Struct. Biol.* **134**, 117–131 (2001).
- Bordenstein, S. R., Marshall, M. L., Fry, A. J., Kim, U. & Wernegreen, J. J. The Tripartite associations between bacteriophage, *Wolbachia*, and arthropods. *PLoS Pathog.* **2**, e43 (2006).
- Wright, J. D., Sjostrand, F. S., Portaro, J. K. & Barr, A. R. The ultrastructure of the rickettsia-like microorganism *Wolbachia pipientis* and associated virus-like bodies in the mosquito *Culex pipiens*. *J. Ultra Res.* **63**, 79–85 (1978).
- Aakre, C. D. et al. Evolving new protein-protein interaction specificity through promiscuous intermediates. *Cell* **163**, 594–606 (2015).
- Dumas, E. et al. Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol. Biol.* **13**, 181 (2013).
- Ellegaard, K. M., Klasson, L., Näslund, K., Bourtzis, K. & Andersson, S. G. E. Comparative genomics of *Wolbachia* and the bacterial species concept. *PLoS Genet* **9**, e1003381 (2013).
- Gouy, M., Guindon, S. & Gascuel, O. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).
- Bryant, D. & Moulton, V. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**, 255–265 (2004).
- Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).
- Paraskevopoulos, C., Bordenstein, S. R., Wernegreen, J. J., Werren, J. H. & Bourtzis, K. Toward a *Wolbachia* multilocus sequence typing system: Discrimination of *Wolbachia* strains present in *Drosophila* species. *Curr. Microbiol.* **53**, 388–395 (2006).

Acknowledgements

We would like to thank Sylvain Charlat, Pierrick Labbé, Benjamin Lopin, Nicole Pasteur, and François Rousset for helpful comments on the manuscript. We also thank Khalid Belkhir for bioinformatic analyses performed on the Montpellier Bioinformatics Biodiversity cluster computing platform. Frédéric Delsuc for phylogenetic analyses, and Arnaud Berthomieu, Ali Bouattour, Emilie Dumas, Patrick Makoundou, and Sandra Unal for technical support. Sequencing data were produced through the technical facilities of the LabEX “Centre Méditerranéen de l’Environnement et de la Biodiversité” in the GENSEQ platform. This work was funded by the French ANR (project “CIAWOL” ANR-16-CE02-0006-01 and ANR-10-BINF-0003).

Author contributions

M.B., M.S., and M.W. conceptualized and designed the study; M.B., C.A., MaBe, and F.J. performed the experiments; M.B., M.C.-G., M.S., and M.W. analyzed and interpreted the data; M.B., M.S., and M.W. wrote the manuscript.


Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-017-02749-w>.

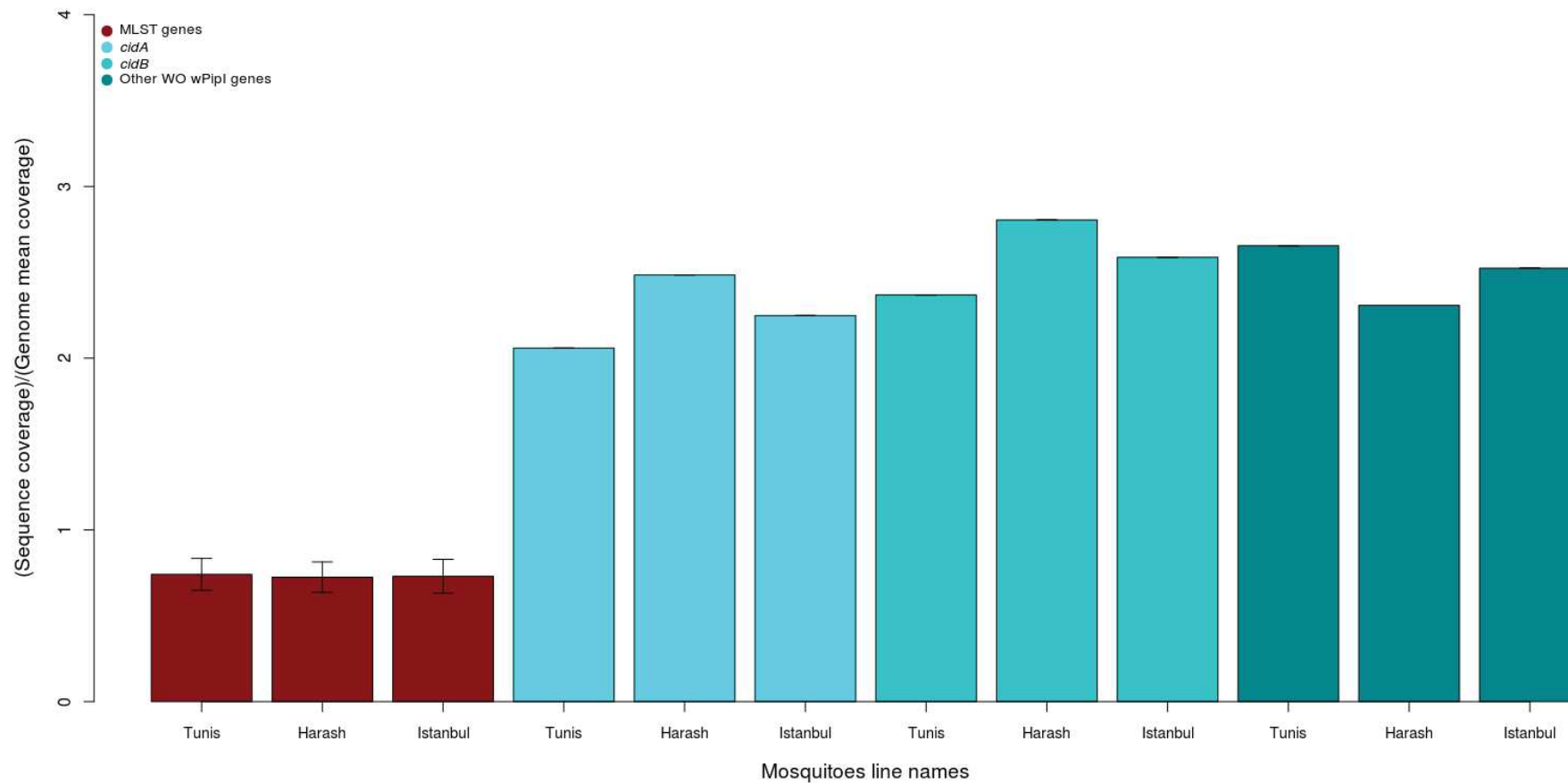
Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018



Supplementary Figure 1 | Mapping depth coverages of *cidA* and *cidB* suggest an amplification of these genes in *wPip* genomes. Comparison of the coverage depth obtained from the mapping of the Illumina reads of *wPip* hosted by Tunis, Istanbul and Harash *C. pipiens* isofemale lines on the reference genome *wPip_Pel* between *cidA*, *cidB*, WO *wPip* prophage I genes (without *cidA* and *cidB*) and 14 housekeeping genes (*gatB*, *coxA*, *hcpA*, *fbpA*, *ftsZ*, *aspC*; *atpD*, *sucB*, *pdhB*, *16s*, *wsp*, *groEL*, *dnaA*, *gltA*) that were predicted to be in single copy in the genome of all *Wolbachia*. Coverage for each line has been normalized by the mean coverage of this line on the reference genome. Error bars represent the standard error on the mean of the average coverage of the 14 housekeeping genes.

CidA_I ($\alpha/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_I ($\gamma/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_I ($\gamma/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_I ($\beta/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_II ($\alpha/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_II ($\alpha/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_II ($\beta/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_III ($\alpha/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_III ($\beta/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_III ($\beta/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_III ($\beta/3$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\alpha/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\alpha/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\gamma/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\gamma/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\delta/1$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL
 CidA_IV ($\delta/2$) MPTQKELRDTMSKKLQEAIKHPDPAVVAGRKSAIKRWVGVLDNFMHEHIKYFKGDKLKFL

CidA_I ($\alpha/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_I ($\gamma/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_I ($\gamma/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_I ($\beta/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_II ($\alpha/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_II ($\alpha/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_II ($\beta/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_III ($\alpha/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_III ($\beta/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_III ($\beta/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_III ($\beta/3$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\alpha/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\alpha/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\gamma/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\gamma/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\delta/1$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE
 CidA_IV ($\delta/2$) HNVFQDEGCWSGVRLDNAALGQRFTEEKIGGIDNPLRKYEMACSYCVVDKIHPLFQKAFE

CidA_I ($\alpha/1$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_I ($\gamma/1$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_I ($\gamma/2$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_I ($\beta/2$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_II ($\alpha/1$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_II ($\alpha/2$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_II ($\beta/2$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_III ($\alpha/1$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_III ($\beta/2$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_III ($\beta/1$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_III ($\beta/3$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\alpha/1$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\alpha/2$) SYKNGFSSDELD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\gamma/1$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\gamma/2$) SYRNKFPPGAFD GKTETEFGKY IIRNSLLDSIKRKGPFDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\delta/1$) SYRNKFPPGAFD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS
 CidA_IV ($\delta/2$) SYRNKFPPGAFD SRGNPITDEY IIRNSLLGGIIRKGPVDFWIDRESGELKKYDAVEGFDS

CidA_I ($\alpha/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_I ($\gamma/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_I ($\gamma/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_I ($\beta/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_II ($\alpha/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_II ($\alpha/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_II ($\beta/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_III ($\alpha/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_III ($\beta/2$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_III ($\beta/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_III ($\beta/3$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_IV ($\alpha/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_IV ($\alpha/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_IV ($\gamma/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_IV ($\gamma/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA
 CidA_IV ($\delta/1$) DIYK PDEEKEIEKKEILDMM LFAKNSFPERFEFFKEKIIKNLRLCGREGKRVNVDYGLFA
 CidA_IV ($\delta/2$) DIYK TDEEKEIEKKEILDMM SFAKDCFPFKFELFKKLIIRD LRLCGREGKRVNVDYGLFA

CidA_I ($\alpha/1$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_I ($\gamma/1$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_I ($\gamma/2$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_I ($\beta/2$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_II ($\alpha/1$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_II ($\alpha/2$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_II ($\beta/2$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_III ($\alpha/1$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_III ($\beta/2$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_III ($\beta/1$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_III ($\beta/3$) EELFSELEKT I L P P G P V G D G P C S N L R S R S K A H G S K K T T L P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\alpha/1$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\alpha/2$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\gamma/1$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\gamma/2$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\delta/1$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y
 CidA_IV ($\delta/2$) EELFSELEKT T L P --- P V G D G P C S N L R S R S K A H G S K K T T F P V D D S P Q S E L G T P S V S G V S S Y

CidA_I ($\alpha/1$) KKKSVFTLSGNK
 CidA_I ($\gamma/1$) KKKSVFTLSGNK
 CidA_I ($\gamma/2$) KKKSVFTLSGNK
 CidA_I ($\beta/2$) KKKSVFTLSGNK
 CidA_II ($\alpha/1$) KKKSVFTLSGNK
 CidA_II ($\alpha/2$) KKKSVFTLSGNK
 CidA_II ($\beta/2$) KKKSVFTLSGNK
 CidA_III ($\alpha/1$) KKKSVFTLSGNK
 CidA_III ($\beta/2$) KKKSVFTLSGNK
 CidA_III ($\beta/1$) KKKSVFTLSGNK
 CidA_III ($\beta/3$) KKKSVFTLSGNK
 CidA_IV ($\alpha/1$) KKKSVFTLSGNK
 CidA_IV ($\alpha/2$) KKKSVFTLSGNK
 CidA_IV ($\gamma/1$) KKKSVFTLSGNK
 CidA_IV ($\gamma/2$) KKKSVFTLSGNK
 CidA_IV ($\delta/1$) KKKSVFTLSGNK
 CidA_IV ($\delta/2$) KKKSVFTLSGNK

Supplementary Figure 2 | Complete sequences of the CidA variants from the different ω Pip groups (I to IV).

Polymorphic regions represented by colored blocks in Figure 3 are reported on the alignment.

CidB_I (a/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_I (a/2) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_I (b/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_I (b/2) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_II (a/2) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_II (a/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_III (c/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_III (a/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_III (b/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_IV (a/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_IV (a/2) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_IV (b/1) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL
 CidB_IV (b/2) MSNGDGLIRSLVDGDLEGFRQGFESFLDQCPSFLYHVSAGRFLPVFFFMSMFSTAHDANIL

CidB_I (a/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_I (a/2) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_I (b/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_I (b/2) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_II (a/2) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_II (a/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_III (c/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_III (a/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_III (b/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_IV (a/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_IV (a/2) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_IV (b/1) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE
 CidB_IV (b/2) NANERVYFRFDNHGVNPRNGENRNTANLKVAVYRDGQQVVCYSISDRPNSDGLRFSERE

CidB_I (a/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_I (a/2) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_I (b/1) RNALVQEIIRQNPNLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDRGRDKFAK
 CidB_I (b/2) RNALVQEIIRQNPNLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDRGRDKFAK
 CidB_II (a/2) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDRGRDKFAK
 CidB_II (a/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDRGRDKFAK
 CidB_III (c/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_III (a/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_III (b/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_IV (a/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_IV (a/2) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_IV (b/1) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK
 CidB_IV (b/2) RDFLVQEIIRQNQGLMEEDLNFEQYKVCMHGKGSQGEAIATVFEVIREKDFRGRDKFAK

CidB_I (a/1) YSASEVHFIRQLFRNHRLTIKEIEGRQLNQNLRLGRSVNFTRVEPGQQRIDNFMEMLA
 CidB_I (a/2) YSASEVHFIRQLFRNHRLTIKEIEGRQLNQNLRLGRSVNFTRVEPGQQRIDNFMEMLA
 CidB_I (b/1) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_I (b/2) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_II (a/2) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_II (a/1) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_III (c/1) YSASEVHFIRQLFRNHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_III (a/1) YSASEVHFIRQLFRNHRLTIKEIEGRQLNQNLRLGRSVNFTRVEPGQQRIDNFMEMLA
 CidB_III (b/1) YSASEVHFIRQLFRNHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_IV (a/1) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_IV (a/2) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_IV (b/1) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA
 CidB_IV (b/2) YSASEINLIRRLGGDHRLTIKEIEGRQLNQNLRLGRSVNFAQVAGQQGIDNFMEMLA

CidB_I (a/1) SNQRQDVRDSLIRGDIILEYVTDIYNNYRAQIENNIEGRSQKFE SHGFLGFLANF SHRYTI
 CidB_I (a/2) SNQRQDVRDSLIRGDIILEYVTDIYNNYRAQIENNIEGRSQKFE SHGFLGFLANF SHRYTI
 CidB_I (b/1) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHRYTI
 CidB_I (b/2) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHRYTI
 CidB_II (a/2) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHRYTI
 CidB_II (a/1) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHRYTI
 CidB_III (c/1) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHLCTI
 CidB_III (a/1) SNQRQDVRDSLIRGDIILEYVTDIYNNYRAQIENNIEGRSQKFE SHGFLGFLANF SHRYTI
 CidB_III (b/1) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHRYTI
 CidB_IV (a/1) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHLCTI
 CidB_IV (a/2) SDRRQDVRDRIRREILPYITDIYNNYRQVLENNIENRNQRFEGHGFLGFLANF SHLCTI
 CidB_IV (b/1) SNQRQDVRDSLIRGDIILEYVTDIYNNYRAQIENNIEGRSQKFE SHGFLGFLANF SHLCTI
 CidB_IV (b/2) SNQRQDVRDSLIRGDIILEYVTDIYNNYRAQIENNIEGRSQKFE SHGFLGFLANF SHLCTI

CidB_I (a/1) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_I (a/2) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_I (b/1) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_I (b/2) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_II (a/2) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_II (a/1) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_III (c/1) DIDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_III (a/1) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_III (b/1) GVDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_IV (a/1) DIDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_IV (a/2) DIDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_IV (b/1) DIDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP
 CidB_IV (b/2) DIDLDSLPRNSHVAFVLRHQVERENIPIVINLATRA-PPYIALNRRARSHAERLHVFSFIP

CidB_I (a/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_I (a/2) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_I (b/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_I (b/2) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_II (a/2) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_II (a/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_III (c/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_III (a/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_III (b/1) IHTESRNTVCVGLNFNLNLDPFVSDTVGLQQDRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_IV (a/1) IHTTSEHTVCVGLNFNSRLAPFRVGVIAELQQNRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_IV (a/2) IHTTSEHTVCVGLNFNSRLAPFRVGVIAELQQNRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_IV (b/1) IHTTSEHTVCVGLNFNSRLAPFRVGVIAELQQNRFPLVQRLFECLENEGIRENIRDFLLHH
 CidB_IV (b/2) IHTTSEHTVCVGLNFNSRLAPFRVGVIAELQQNRFPLVQRLFECLENEGIRENIRDFLLHH

CidB_I (a/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_I (a/2) LPAEIPRNAENYDRIFDCITGFAFGNSAFDRHELELEE-ED EAPITKYIFRHGDEGLR-C
 CidB_I (b/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_I (b/2) LPAEIPRNAENYDRIFDCITGFAFGNSAFDRHELELEE-ED EAPITKYIFRHGDEGLR-C
 CidB_II (a/2) LPAEIPRNAENYDRIFDCITGFAFGNSAFDRHELELEE-ED EAPITKYIFRHGDEGLR-C
 CidB_II (a/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_III (c/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_III (a/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_III (b/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_IV (a/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_IV (a/2) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEE-ED EAPITKYIFRHGDEGLR-C
 CidB_IV (b/1) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEEGEDRVRVRKYIFRYGDGDLRRH
 CidB_IV (b/2) LPAEIPRNAENYDRIFDCITGFAFGNSAFDREYLELEE-ED EAPITKYIFRHGDEGLR-C

CidB_I (a/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_I (a/2) LTMVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_I (b/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_I (b/2) LTMVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_II (a/2) LTMVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_II (a/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_III (c/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_III (a/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_III (b/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_IV (a/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_IV (a/2) LTMVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_IV (b/1) TLTMTVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL
CidB_IV (b/2) LTMVFHAEGSDIVILHIRAHDAQQQGAINLQTLNVNGNDVHVWEVSVCTLNNQLELDIDL

CidB_I (a/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_I (a/2) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_I (b/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_I (b/2) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_II (a/2) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_II (a/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_III (c/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_III (a/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_III (b/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_IV (a/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_IV (a/2) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_IV (b/1) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA
CidB_IV (b/2) PNDLGLYHDYQNNNANNFLAGDLVQVPNTENVHNTLNQVVNDGWKNIAQHRGLFQEISGA

CidB_I (a/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_I (a/2) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_I (b/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_I (b/2) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_II (a/2) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_II (a/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_III (c/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_III (a/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_III (b/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_IV (a/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_IV (a/2) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_IV (b/1) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR
CidB_IV (b/2) LMPLVDTINVNSEDKFRSILHGTFYASDNPKVFLAMYKVGQTYSLKRGQEEEGERVILTR

CidB_I (a/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_I (a/2) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_I (b/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_I (b/2) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_II (a/2) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_II (a/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_III (c/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_III (a/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_III (b/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_IV (a/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_IV (a/2) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_IV (b/1) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP
CidB_IV (b/2) ITEQRDLDDLLLRQPRENDLDTHPIGYVLRANNAEEVGGQQNDARQIEGRLLKKQHRGFIP

CidB_I (a/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_I (a/2) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_I (b/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_I (b/2) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_II (a/2) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_II (a/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_III (c/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_III (a/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_III (b/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_IV (a/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_IV (a/2) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_IV (b/1) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID
CidB_IV (b/2) ITSGNEVVLFPIVFNRDAHEAGNLILFPEGIGREEHVHRLDRHVRSSRPGLVGPESVID

CidB_I (a/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_I (a/2) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_I (b/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_I (b/2) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_II (a/2) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_II (a/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_III (c/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_III (a/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_III (b/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_IV (a/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_IV (a/2) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_IV (b/1) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR
CidB_IV (b/2) ENPPEGLLSDQTRENFRRFYEEKAPGQNSIFLLDIGDNLHVPFSYLQGTTRAQVIETLKSR

CidB_I (a/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_I (a/2) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_I (b/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_I (b/2) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_II (a/2) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_II (a/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_III (c/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_III (a/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_III (b/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_IV (a/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_IV (a/2) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_IV (b/1) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR
CidB_IV (b/2) IRGGGTPTAQGILQQINAILRRNNAREIEDVHDLLALDFATDNQNYRYWLQTHDMFFAAR

CidB_I (a/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_I (a/2) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_I (b/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_I (b/2) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_II (a/2) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_II (a/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_III (c/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_III (a/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_III (b/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_IV (a/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_IV (a/2) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_IV (b/1) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI
CidB_IV (b/2) QYTFLDNQSHSTNDHYGFEITSVGVDGNQNDPTGRGLLSSHITNFKQKVDSGEKDRLIAI

CidB_I (a/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_I (a/2)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_I (b/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_I (b/2)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_II (a/2)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_II (a/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_III (c/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_III (a/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_III (b/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_IV (a/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_IV (a/2)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_IV (b/1)	INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ
CidB_IV (b/2)	<u>INVGNRHWVTLVIVHQNGNYYGYADSLGPDSGIDNNIRGALRECDINDDNVHNI SVHQQ</u>

CidB_I (a/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_I (a/2)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_I (b/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_I (b/2)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_II (a/2)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_II (a/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_III (c/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_III (a/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_III (b/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_IV (a/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_IV (a/2)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_IV (b/1)	TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN
CidB_IV (b/2)	<u>TDGHNCGIWVYENARDINQAIDQALQGNNNFGEKGEIGIIGYIRGLLSAGIGNDTRQPRRN</u>

CidB_I (a/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_I (a/2)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_I (b/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_I (b/2)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_II (a/2)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_II (a/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_III (c/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_III (a/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_III (b/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_IV (a/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_IV (a/2)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_IV (b/1)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL
CidB_IV (b/2)	EQYFEDRRRDISQLLQNDPNLPSRRSDLIQAHPGIQHEIDPLLLQFLGLQYPQRGGGGAL

CidB_I (a/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_I (a/2)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_I (b/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_I (b/2)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_II (a/2)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_II (a/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_III (c/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_III (a/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_III (b/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_IV (a/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_IV (a/2)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_IV (b/1)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*
CidB_IV (b/2)	QLGGERVISIDFGNPQSALDKIDGVSrvYNHsNSRGSr*

Supplementary Figure 3 | Complete sequences of the CidB variants detected in the different wPip group of *Wolbachia* strains studied.

Polymorphic regions represented by colored blocks in Figure 4 are reported on the alignment. The deubiquitylating (DUB) domain is underlined in red.

Supplementary Table 1 | Information on the four reference lines used for crossing type determination

Line Name	<i>Wolbachia</i> group	<i>Culex pipiens</i> subspecies	Country	Years of collection	Reference
Tunis	wPip-I	<i>pipiens</i>	Tunisia	1992	Duron <i>et al.</i> 2005
Lavar	wPip-II	<i>pipiens</i>	France	2003	Duron <i>et al.</i> 2005
Maclo	wPip-III	<i>quinquefasciatus</i>	USA	1984	Duron <i>et al.</i> 2005
Istanbul	wPip-IV	<i>pipiens</i>	Turkey	2003	Duron <i>et al.</i> 2005

Supplementary Table 2 | List of the primers used in this study.

Primers' number	Primers' name	Primers' sequences	PCR primers couple	Fragment size (bp)	Tm (°C)
1	wpip_0282_287981_dir	TGGTCAGGTGTAAGGTTGGA	1/2	1300	57
2	wpip_0282_289280_rev	CGACCAGAAACACCAAGAGT			
3	wpip0282_287607_Dir	ACTCAAACACTGGTCTTTCTGT	3/4	81	59
4	wpip0282_288423_Rev	TGAACGCGAGAAAGAGCAAG			
5	wpip0282_289157_Dir	GATAGTCCGCAGTCTGAGCT	5/6	486	57
6	wpip0283_289642_Rev	ACCCATCACTATTAGGACGATCA			
7	wpip_283_289555_Dir	ACGGCAAACCTAAAAGTTGCT	7/8	1264	59
8	wpip_283_290818_Rev	TGATGGCTCCTTGTTGTTGC			
9	wpip0283_290614_Dir	AACTGGTTTTGCTTTTGGGAA	9/10	1280	57
10	wpip0283_291893_Rev	ATCTCACGAGCGTTGTTTCT			
11	wpip0283_291699_Dir	AGCACCAGGACAAAATTCGA	11/12	1211	59
12	wpip0283_292909_Rev	GCAGAAACAACCTAGAAGACCG			
13	wpip_0282_288758_rev	ACAAGCTACCACATTCCCAA	1/13	778	57
14	wpip_0282_QPCR_2_Dir	AGGTCCTGTATTTGATTTCTGGA	14/15	189	58
15	wpip_0282_QPCR_2_Rev	TGAACGCGAGAAAGAGCAAG			
16	wpip_0283_QPCR_1_Dir	TGAGTGTTTGGAGAATGAAGGA	16/17	135	58
17	wpip_0283_QPCR_1_Rev	TTCCCAAAGCAAACCAAGTT			
18	wolpipdir	AGAATTGACGGCATTGAATA	18/19	151	58
19	wolpiprev	CGTCGTTTTTGTTTAGTTGTG			

Supplementary Table 3 | Accession numbers.

Accession numbers for *cidA cidB* variants analyzed in Figure 2,3,4,5, in Extended Data Figures 2, 3 and in Supplementary Data Figure 2,3,4,5 and Table 2.

Gene name	Variant sequence submission name	Accession number	Gene name	Variant sequence submission name	Accession number
cidA	cidA_I(alpha/1)	MF444963	cidB	cidB_I(a/1)	MF444982
	cidA_I(gamma/1)	MF444964		cidB_I(a/2)	MF444983
	cidA_I(gamma/2)	MF444965		cidB_I(b/1)	MF444984
	cidA_I(beta/2)	MF444966		cidB_I(b/2)	MF444985
	cidA_II(alpha/1)	MF444967		cidB_II(a/2)	MF444986
	cidA_II(alpha/2)	MF444968		cidB_II(a/1)	MF444987
	cidA_II(beta/2)	MF444969		cidB_III(c/1)	MF444988
	cidA_III(alpha/1)	MF444970		cidB_III(a/1)	MF444989
	cidA_III(beta/2)	MF444971		cidB_III(b/1)	MF444990
	cidA_III(beta/1)	MF444972		cidB_IV(a/1)	MF444991
	cidA_III(beta/3)	MF444973		cidB_IV(a/2)	MF444992
	cidA_IV(alpha/1)	MF444974		cidB_IV(b/1)	MF444993
	cidA_IV(alpha/2)	MF444975		cidB_IV(b/2)	MF444994
	cidA_IV(gamma/1)	MF444976		cidB_IV(a/3)	MF444995
	cidA_IV(gamma/2)	MF444977		cidB_IV(b/3)	MF444996
	cidA_IV(delta/1)	MF444978			
	cidA_IV(delta/2)	MF444979			
	cidA_IV(beta/1)	MF444980			
	cidA_IV(beta/2)	MF444981			


Supplementary Data 1 : 180 Isofemale lines used in this study and the distribution of *cidA_IV*(δ) and *cidB_IV*(a/2) variants

Line Name	Isofemale Number	Country	Year of collection	Reference	Crossing Type	Presence of <i>cidA_IV</i>(δ) variant in the PCR-RFLP Test	Presence of <i>cidB_IV</i>(a/2) variant in the PCR-RFLP Test
Istanbul	Ist-1*	Turkey	2003	Duron et al. 2005	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Hang Zhou	Hang-1*	China	2003	Atyame et al. 2014	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Harash	Ha-1*	Algeria	2006	Atyame et al. 2015	Compatible	-	-
Souk ahras	Souk-1*	Algeria	2008	Atyame et al. 2014	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Souk ahras	Souk-2*	Algeria	2008	Atyame et al. 2014	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Guelma	Gue-1*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Guelma	Gue-2*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Douas	Dou-1*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Douas	Dou-2*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Kala	Kala-1*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Lac	Lac-1*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Lac	Lac-2*	Algeria	2008	Atyame et al. 2014	Compatible	-	-
Boussaleme	Bouss-1*	Tunisia	2008	Atyame et al. 2014	Compatible	-	-
Boussaleme	Bouss2*	Tunisia	2008	Atyame et al. 2014	Compatible	-	-
Kef	Kef-1*	Tunisia	2008	Atyame et al. 2014	Compatible	-	-
Kef	Kef-2*	Tunisia	2008	Atyame et al. 2014	Compatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Tabarka	Tab-1*	Tunisia	2008	Atyame et al. 2014	Compatible	-	<i>cidB_IV</i> (a/2)
Tabarka	Tab-2*	Tunisia	2008	Atyame et al. 2014	Compatible	-	-
Ichkeul	Ich-01	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-02	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-03	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-04	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-05	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-06	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-07	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-08	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-09*	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-10	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-11	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-12	Tunisia	2013	This study	Compatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-13*	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-14	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-15	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-16	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-17	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-18	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-19	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-20	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-21*	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-22	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-23	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-24	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-25	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-26	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-27	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-28	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-29	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-30	Tunisia	2013	This study	Compatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-31	Tunisia	2013	This study	Incompatible	<i>cidA_IV</i> (δ)	<i>cidB_IV</i> (a/2)
Ichkeul	Ich-32	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-33	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-34	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-35	Tunisia	2013	This study	Compatible	-	-
Ichkeul	Ich-36	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-1	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-2	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-5	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-6	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-7	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-9	Tunisia	2013	This study	Compatible	-	-
Rasjbal	RJ-10	Tunisia	2013	This study	Compatible	-	-


Annexes partie I

Chapitre II

Variants Name	CidA positions (aa)																			
	1	1	1	1 1	1 1	1	1 1	1	3 3	3 3	3 3	3	3	3	3	3 3	3	3	3 3	4 4
	1	2	2	2-3	3-4	4	4 5	5	3 3	4-4	4 4	5	5	6	8	8 8	9	9	9 9	0 0
	8	3	5	7 1	3 1	3	9 0	2	6 7	2 4	6 7	0	8	5	1	5 6	0	3	6 7	0 1
CidA_IV (α/1)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN
CidA_IV (α/2)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD
CidA_IV (β/1)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN
CidA_IV (β/2)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD
CidA_IV (γ/1)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN
CidA_IV (γ/2)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD
CidA_IV (δ/1)	R	R	K	PPGAF	SRGNPITDE	I	GG	R	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN
CidA_IV (δ/2)	R	R	K	PPGAF	SRGNPITDE	I	GG	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD
CidA_IV (ε/1)	R	R	K	PPGAF	GKTETEF GK	I	GG	R	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN
CidA_IV (ε/2)	R	R	K	PPGAF	GKTETEF GK	I	GG	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD



Upstream polymorphic region



Downstream polymorphic region

Supplementary Figure 1. Sequences of all the CidA_{IV} protein variants obtained to date

Alignment of protein sequences for the CidA_{IV} variants previously identified in natural populations from North Africa, China and Turkey (Bonneau *et al.*, 2018a) and in new variants detected in natural populations from North Italy.. The first sequence is used as a reference for defining the polymorphic region. For greater clarity, only the polymorphic positions of the alignment are represented, and amino-acid positions are not contiguous. When more than two contiguous amino acids were variable the “-” symbol is used between the first and last variable positions of the zone.

CidB positions (aa)

Variants Name	Upstream polymorphic region										Downstream polymorphic region					
	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	4 4	4	4 4	4	4 4	4 4
	4 4	5 5	5 5	5 5	5 5	6	6-7	7 7	7 7	5 5	5	6-6	7 7 7	7-8	4 4	
	2 3	0 1	3 4	7 9	2	8 0	6 8	6 8	6 8	2 3	9	2 6	2	5 6	9 1	
CidB_IV (a/1)	DR	RI	RE	P	I	I	QVL	N	N	EY	G	RVRVR	Y	GD	RHT	
CidB_IV (a/2)	DR	RI	RE	P	I	I	QVL	N	N	EY	*	EAPIT	H	EG	**C	
CidB_IV (a/3)	DR	RI	RE	P	I	I	QVL	N	N	HP	*	EAPIT	H	EG	**C	
CidB_IV (b/1)	NQ	SL	GD	E	V	T	AQI	G	S	EY	G	RVRVR	Y	GD	RHT	
CidB_IV (b/2)	NQ	SL	GD	E	V	T	AQI	G	S	EY	*	EAPIT	H	EG	**C	
CidB_IV (b/3)	NQ	SL	GD	E	V	T	AQI	G	S	HP	*	EAPIT	H	EG	**C	
CidB_IV (c/1)	NQ	SL	GD	E	V	T	AQI	N	N	EY	G	RVRVR	Y	GD	RHT	
CidB_IV (c/3)	NQ	SL	GD	E	V	T	AQI	N	N	HP	*	EAPIT	H	EG	**C	

Supplementary Figure 2. Sequences of all the CidB_IV protein variants obtained to date

Alignment of the protein sequences of the CidB_IV variants previously found in natural populations from North Africa, China and Turkey (Bonneau *et al.*, 2018a) and the new variants identified in natural populations from North Italy. The first sequence is used as a reference for defining the polymorphic region. For greater clarity, only polymorphic positions of the alignment are represented and amino-acid positions are not contiguous. When more than two contiguous amino acids were variable, the “-” symbol is used between the first and last variable positions of the zone.

Supplementary Table 1. Accession numbers.

Accession numbers for the *cidA* and *cidB* variants analyzed in Fig. 1 and Supplementary Figs. 1 and 2

Gene name	Variant sequence submission name	Accession number	Gene name	Variant sequence submission name	Accession number
cidA	cidA_IV(alpha/1)	MF444974	cidB	cidB_IV(a/1)	MF444991
	cidA_IV(alpha/2)	MF444975		cidB_IV(a/2)	MF444992
	cidA_IV(gamma/1)	MF444976		cidB_IV(b/1)	MF444993
	cidA_IV(gamma/2)	MF444977		cidB_IV(b/2)	MF444994
	cidA_IV(delta/1)	MF444978		cidB_IV(a/3)	MF444995
	cidA_IV(delta/2)	MF444979		cidB_IV(b/3)	MF444996
	cidA_IV(beta/1)	MF444980		cidB_IV(c/1)	MH809184
	cidA_IV(beta/2)	MF444981		cidB_IV(c/3)	MH809185
	cidA_IV(epsilon/1)	MH809182			
	cidA_IV(epsilon/2)	MH809183			

Rovere della Luna	Luna 10	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 11	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 12	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 13	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 14	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 15	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 17	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 18	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 20	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 21	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 23	Compatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	-	-
Rovere della Luna	Luna 24	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 25	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 26	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 27	Compatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	-	-
Rovere della Luna	Luna 28	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Rovere della Luna	Luna 30	Incompatible	46.249681,11.169314	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 1	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 2	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 3	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 4	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 5	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 6	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 7	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 8	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 9	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Mezzocorona	Mezzo 10	Incompatible	46.216701,11.118225	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Zambana	Zamb 1	Incompatible	46.151926,11.097357	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Zambana	Zamb 2	Incompatible	46.151926,11.097357	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Zambana	Zamb 4	Incompatible	46.151926,11.097357	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Zambana	Zamb 5	Incompatible	46.151926,11.097357	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)
Zambana	Zamb 6	Incompatible	46.151926,11.097357	<i>cidA_IV</i> (α)	-	<i>cidB_IV</i> (2)	<i>cidB_IV</i> (2)

Supplementary Dataset 1. *w*PipIV *mod* phenotype and occurrence of *cidA_IV*(α), *cidA_IV*(δ) and *cidB_IV*(2) variants in the 67 isofemale lines.

Note that the detection of *cidB_IV*(2) accounts for the presence of *cidB_IV*(a/2) and/or *cidB_IV*(b/2) variants. Two methods were used to detect *cidB_IV*(2) variants: a PCR-RFLP test and a specific PCR.

Supplementary Dataset 2. Hatching rates for crosses between males from the Michele26 and Luna1 isofemale lines and females from the Tunis line

Male line	eggs per eggs-raft	larvae per eggs-raft	unhatched eggs per eggs-raft	Hatching rate
Michele26	129	85	44	0.658914728682171
Michele26	145	99	46	0.682758620689655
Michele26	133	117	16	0.879699248120301
Michele26	89	82	7	0.921348314606742
Michele26	90	87	3	0.966666666666667
Michele26	119	110	9	0.92436974789916
Michele26	137	130	7	0.948905109489051
Michele26	142	85	57	0.598591549295775
Michele26	138	93	45	0.673913043478261
Michele26	124	115	9	0.92741935483871
Michele26	125	115	10	0.92
Michele26	124	121	3	0.975806451612903
Michele26	129	87	42	0.674418604651163
Michele26	125	118	7	0.944
Michele26	134	115	19	0.858208955223881
Michele26	118	112	6	0.949152542372881
Michele26	145	55	90	0.379310344827586
Michele26	116	84	32	0.724137931034483
Michele26	80	37	43	0.4625
Michele26	94	87	7	0.925531914893617
Michele26	123	90	33	0.731707317073171
Michele26	133	125	8	0.93984962406015
Michele26	110	103	7	0.936363636363636
Michele26	120	105	15	0.875
Michele26	155	130	25	0.838709677419355
Michele26	144	127	17	0.881944444444444
Michele26	125	75	50	0.6
Michele26	117	91	26	0.777777777777778
Michele26	154	119	35	0.772727272727273
Michele26	119	43	76	0.361344537815126
Luna8	101	85	16	0.841584158415841
Luna8	123	115	8	0.934959349593496
Luna8	109	84	25	0.770642201834862
Luna8	124	100	24	0.806451612903226
Luna8	98	73	25	0.744897959183674
Luna8	109	82	27	0.752293577981651
Luna8	132	105	27	0.795454545454545
Luna8	151	136	15	0.900662251655629
Luna8	111	100	11	0.900900900900901
Luna8	105	91	14	0.866666666666667
Luna8	109	101	8	0.926605504587156
Luna8	114	85	29	0.745614035087719
Luna8	155	126	29	0.812903225806452
Luna8	97	90	7	0.927835051546392
Luna8	90	45	45	0.5
Luna8	117	91	26	0.777777777777778
Luna8	150	138	12	0.92
Luna8	113	102	11	0.902654867256637
Luna8	131	111	20	0.847328244274809
Luna8	114	77	37	0.675438596491228
Luna8	79	54	25	0.683544303797468
Luna8	93	80	13	0.860215053763441
Luna8	144	113	31	0.784722222222222
Luna8	101	88	13	0.871287128712871
Luna8	113	108	5	0.955752212389381
Luna8	73	53	20	0.726027397260274
Luna8	77	60	17	0.779220779220779
Luna8	127	100	27	0.78740157480315
Luna8	135	124	11	0.918518518518519
Luna8	150	121	29	0.806666666666667

Supplementary Dataset 3. Levels of expression of the *cidA* and *cidB* genes and *cidB_IV(2)* variants of males from the Mezzo9 and Michele26 isofemale lines

Male Line	Individual Name	<i>cidA</i> expression	<i>cidB</i> expression	<i>cidB_IV(2)</i> expression
Mezzo9	Mezzo9 M1	0.9703	0.5076	0.0369
Mezzo9	Mezzo9 M2	1.424	0.6043	0.0671
Mezzo9	Mezzo9 M3	2.84	1.147	0.0807
Mezzo9	Mezzo9 M4	1.476	0.6518	0.0629
Mezzo9	Mezzo9 M5	1.056	0.5978	0.0521
Mezzo9	Mezzo9 M6	1.622	0.48	0.0543
Mezzo9	Mezzo9 M7	2.001	1.027	0.0758
Mezzo9	Mezzo9 M8	1.342	0.5742	0.0534
Mezzo9	Mezzo9 M9	1.808	0.9243	0.0687
Mezzo9	Mezzo9 M10	2.161	0.7129	0.0695
Michele26	Michele26 M1	1.552	0.7029	0.0563
Michele26	Michele26 M2	1.722	0.7091	0.0682
Michele26	Michele26 M3	1.131	0.5032	0.0354
Michele26	Michele26 M4	1.556	0.6348	0.0462
Michele26	Michele26 M6	0.8012	0.3741	0.0208
Michele26	Michele26 M7	1.32	0.6464	0.0343
Michele26	Michele26 M8	1.049	0.346	0.0181
Michele26	Michele26 M9	1.603	0.895	0.0454
Michele26	Michele26 M10	1.196	0.4336	0.0347

Supplementary Dataset 4. *Wolbachia* density in the testes of males from the Mezzo9 and Michele26 lines

Male Line	Pool Name	<i>Wolbachia</i> density (wsp/ace-2)
Michele26	Michele26 T1	7.484
Michele26	Michele26 T2	6.696
Michele26	Michele26 T3	10.63
Michele26	Michele26 T4	15.3
Michele26	Michele26 T5	11.12
Mezzo9	Mezzo9 T1	11.51
Mezzo9	Mezzo9 T2	17.81
Mezzo9	Mezzo9 T3	20.62
Mezzo9	Mezzo9 T4	16.43
Mezzo9	Mezzo9 T5	24.14

Annexes partie II

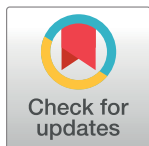
RESEARCH ARTICLE

The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity

Manon Bonneau¹, Frédéric Landmann², Pierrick Labbé¹, Fabienne Justy¹, Mylène Weill¹, Mathieu Sicard^{1*}

1 ISEM, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France, **2** CRBM, University of Montpellier, CNRS, Montpellier, France

* mathieu.sicard@umontpellier.fr



OPEN ACCESS

Citation: Bonneau M, Landmann F, Labbé P, Justy F, Weill M, Sicard M (2018) The cellular phenotype of cytoplasmic incompatibility in *Culex pipiens* in the light of *cidB* diversity. PLoS Pathog 14(10): e1007364. <https://doi.org/10.1371/journal.ppat.1007364>

Editor: Elizabeth Ann McGraw, Pennsylvania State University, UNITED STATES

Received: July 12, 2018

Accepted: September 28, 2018

Published: October 15, 2018

Copyright: © 2018 Bonneau et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its supporting Information files.

Funding: This work was funded by the French ANR (project "CIAWOL" ANR-16-CE02-0006-01 coordinated by MS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Wolbachia are maternally inherited endosymbiotic bacteria, widespread among arthropods thanks to host reproductive manipulations that increase their prevalence into host populations. The most commonly observed manipulation is cytoplasmic incompatibility (CI). CI leads to embryonic death in crosses between i) infected males and uninfected females and ii) individuals infected with incompatible *Wolbachia* strains. CI can be conceptualized as a toxin-antidote system where a toxin deposited by *Wolbachia* in the sperm would induce embryonic death unless countered by an antidote produced by *Wolbachia* present in the eggs. In *Drosophila melanogaster*, transgenic expression of *Wolbachia* effector *cidB* revealed its function of CI-inducing toxin. Moreover in *Culex pipiens*, the diversity of *cidB* variants present in *wPip* strains accounts for the diversity in crossing-types. We conducted cytological analyses to determine the CI mechanisms that lead to embryonic death in *C. pipiens*, and assess whether diversity in crossing-types could be based on variations in these mechanisms. We revealed that paternal chromatin condensation and segregation defects during the first embryonic division are always responsible for embryonic death. The strongest observed defects lead to an exclusion of the paternal chromatin from the first zygotic division, resulting in haploid embryos unable to hatch. The proportion of unhatched haploid embryos, developing with only maternal chromatin, which reflects the frequency of strong defects can be considered as a proxy of CI intensity at the cellular level. We thus studied the putative effect of variations in crossing types and *cidB* diversification on CI defects intensity. Incompatible crosses involving distinct *wPip* strains revealed that CI defects intensity depends on the *Wolbachia* strains hosted by the males and is linked to the diversity of *cidB* genes harbored in their genomes. These results support that, additionally to its implication in *C. pipiens* crossing type variability, *cidB* diversification also influences the strength of CI embryonic defects.

Author summary

In some crosses, mosquito males belonging to the species *Culex pipiens* prevent their females from having live progenies. This phenomenon called cytoplasmic incompatibility (CI) is caused by intracellular bacteria named *Wolbachia*. CI occurs when males infected with *Wolbachia* fertilize females infected with genetically distinct incompatible *Wolbachia* resulting in the death of all the embryos. At the world scale, crossing relationships between *C. pipiens* are quite puzzling. Despite this complexity in crossing relationships and the diversity of *cidB* genes involved in CI mechanisms in *C. pipiens*, we demonstrate a single shared CI cellular phenotype leading to the death of the embryos: the paternal chromatin exclusion from the first embryonic division. If paternal chromatin is fully excluded, embryos developed with haploid set of chromosomes. We show that the frequency of haploid development varies according to the *Wolbachia* strains hosted by the males which differ in the *cidB* variants harbored in their genomes. Absence of *Wolbachia* in the eggs totally block haploid development showing that maternal *Wolbachia* presence interplays with CI mechanisms in a way that allows haploid development to occur. Understanding CI mechanism in mosquitoes is the corner stone to build new sustainable and adaptable *Wolbachia* based strategies for vector control.

Introduction

Wolbachia are maternally-inherited endosymbionts, widespread among arthropods and filarial nematodes [1,2], and the most frequent endocytobiotic bacteria detected in arthropods [3]. This high prevalence is attributed to their ability to manipulate their host reproduction to spread within arthropod populations [1]. The main reproductive manipulation strategy used by *Wolbachia* is named cytoplasmic incompatibility (CI) [4]. CI is a form of conditional sterility resulting in embryonic lethality [5]. In most of the host species, CI occurs when males infected with *Wolbachia* fertilize uninfected females whereas the reciprocal cross remains compatible. This difference in the production of viable offspring between infected and uninfected female reproduction enhances the spread of *Wolbachia* in host's populations [6]. CI can also occur between males and females both infected with different and incompatible *Wolbachia* strains [7–12]. In such situations, incompatibility can be either unidirectional (one cross direction is compatible while the reciprocal one is incompatible) or bidirectional (both cross directions are incompatible) [7–9]. The penetrance of CI, *i.e.* the number of embryos affected by CI in a cross, varies depending on the *Wolbachia* strain and the host involved in the interaction. Indeed, in the same host *Drosophila simulans*, *w*Ri induces complete CI (*i.e.* crosses in which all the embryos are affected by CI), while *w*No and *w*Ha strains induce lower levels of CI, *i.e.* some embryos can develop [13]. Complete CI penetrance was also described in *Nasonia spp.* depending on the *Wolbachia* strain involved and in all *Culex pipiens* incompatible crosses [14,15]. The variability of CI penetrance has been correlated to different factors such as the *Wolbachia* density in the sperm and eggs [16,17], host age [18] and host nuclear genotype [19].

Cellular consequences of *Wolbachia*-induced CI have been monitored during embryogenesis in *D. simulans*, *D. melanogaster* and *N. vitripennis* [5,20–23]. These studies revealed common cellular defects in these three species: a delay in paternal chromatin condensation and segregation defect during the first mitotic division of the embryo [23–25]. In *D. melanogaster*, a delay in histone H3.3 deposition after protamine removal on the paternal chromatin was observed and linked to chromatin remodeling defects [22]. This remodeling defect was associated with the persistence of the DNA replication factor PCNA (Proliferating Cell Nuclear Antigen)

during mitosis, reflecting incomplete replication of paternal DNA [22]. It has been envisioned that these DNA replication defects might be responsible for the paternal chromatin bridges and segregation failure during the first mitotic division, which result in early embryonic arrest [22,24,25]. Nevertheless, some embryos reaching late development stages have been reported in CI crosses [21,25,26]. Late development in CI embryos was interpreted as resulting from a complete paternal chromatin exclusion during the first division, which allows successful maternal chromatin segregation and the formation of two haploid nuclei [25]. These haploid nuclei which further divide, lead to gynogenetic development (*i.e.* haploid development with only maternal genetic material) until late embryonic stages [27]. These haploid embryos are never viable in diploid species such as *D. simulans* [25]. However, in the haplodiploid parasitoid wasps *Leptopilina heterotoma* and *N. vitripennis*, CI-induced paternal chromosome defects can lead either i) to the death of the embryos or ii) to the production of healthy males [19,21,28,29]. It has been proposed that these two CI developmental outcomes could result from different degree of paternal chromatin defect (improper condensation) before the first division [29–31]. Severe defects would lead to complete elimination of male chromosomes from the first zygotic division resulting in haploidization and male development, whereas less severe defects would lead in partial exclusion of the paternal chromatin resulting in incomplete elimination of male chromosomes and early arrest of the aneuploid development [29–31]. In diploid species such as *C. pipiens* and *D. melanogaster*, the proportion of unhatched developed haploid embryos observed in fully incompatible CI crosses would be a proxy of the frequency of total paternal chromatin exclusion during embryogenesis due to strong CI intensity at the cellular level. However, this hypothesis is counter intuitive as one could expect that strong CI defects would prevent any development to occur while soft CI defect would allow development.

The molecular mechanism underlying CI can be conceptualized as a toxin-antidote system in which i) a toxin produced by *Wolbachia* in the testes, more generally called a “*mod* factor”, and introduced in the sperm during spermatogenesis would interfere (“modify”) with the paternal chromatin and induce embryonic perturbations, and ii) an antidote released by *Wolbachia* in the egg, more generally called “*resc* factor”, would “rescue” these paternal chromatin defects to allow normal embryogenesis to occur [32,33]. The recent discoveries of i) *Wolbachia* genes *cidA* and *cidB* ability to recapitulate the CI phenotypes when expressed in transgenic *Drosophila* [34,35], and ii) the link between specific allelic *cidA*^{wPip}/*cidB*^{wPip} variations in worldwide natural *C. pipiens* populations and the capacity of males to sterilize females [36], open new paths into understanding CI mechanisms. *CidA* and *cidB* genes are syntenic genes within the WO phage region (S6 Table) [34,35,37–39]. *CidB* encodes a deubiquitylating enzyme (DUB) and when a *cidB*^{wPip} construct bearing this catalytically inactivated DUB domain was expressed in *D. melanogaster* males, CI was no longer observed, showing the implication of the deubiquitylating activity in the *mod* function [34]. The role of *cidA* in the CI mechanism is more debated because i) both *cidA*^{wMel} and *cidB*^{wMel} are required to induce CI in transgenic *Drosophila* [35] and ii) in natural populations of *C. pipiens*, specific *cidA*^{wPip} allelic variations were found to be linked to *mod* variations [36]. However, the implication of *cidA* in the *resc* function is supported by the capacity of *cidA* to prevent *cidB* toxicity in yeast [34] and the capability of transgenic uninfected females expressing *cidA*^{wMel} throughout oogenesis to rescue the effect of *cidB*^{wMel} [40].

In *C. pipiens*, all individuals are infected with different *Wolbachia* strains belonging to the monophyletic wPip group, but divided in five subgroups wPipI to wPipV. MLST (Multi Locus Sequence Typing) genes from Baldo et al. (2006) [41] were not polymorphic between wPip strains, thus a wPip specific MLST with more polymorphic genes MutL, ank2, pk1, pk2, GP12, GP15, and RepA was used to resolve wPip phylogeny (S6 Table) [12]. Mosquitoes hosting

wPip from the same group are likely to be compatible with each other but incompatible with mosquitoes infected with *Wolbachia* from other *wPip* groups [42]. This diversity of *wPip* strains distributed all around the world is responsible for the unique complexity of CI crossing types described in this host species [12,43]. Unlike *wVitA* and *wMel*, which harbors only one *cidA/cidB* copy, and *wRi*, which harbors two identical copies of *cidA/cidB*, high intra and inter-genomic diversities of *cidA^{wPip}/cidB^{wPip}* genes were uncovered between and within all *wPip* strains studied [36]. This diversity certainly explains the unrivaled diversity of crossing types described in *C. pipiens* [36]. This *cidA^{wPip}/cidB^{wPip}* genes amplification and diversification within the same *Wolbachia* genome may also account for the impressive CI penetrance described in *C. pipiens*. Indeed, expression of multiple *cidA^{wPip}* and *cidB^{wPip}* variants in males could i) be responsible for differences in CI cellular phenotype(s) and ii) influence the penetrance of CI. Here, we investigated the putative impact of crossing type variations and *cidA^{wPip}/cidB^{wPip}* diversification on CI cellular phenotypes and CI intensity during *C. pipiens* embryogenesis. To this end, we monitored the development of embryos derived from various incompatible crosses involving males from *C. pipiens* lines infected with *Wolbachia* strains from distinct *wPip* groups and exhibiting different crossing types.

Results

A single cellular phenotype of CI in *C. pipiens*

Three different types of crosses were performed using different laboratory mosquito lines: i) fertile crosses between individuals from the same line, representing our control to monitor normal embryonic early development, ii) sterile crosses between mosquito lines harboring different *wPip* strains, and iii) sterile crosses between infected males and uninfected females (TC lines), to test the effect of *Wolbachia* absence on embryo development and CI cellular mechanism (S1 and S2 Tables).

The cellular phenotype during embryogenesis in fertile intra-line crosses is illustrated in Fig 1. To differentially visualize the paternal from the maternal chromatin, we used propidium iodide to mark both maternal and paternal chromatin and an anti-acetylated histone H4 labeling that preferentially marks the *de novo* assembled paternal chromatin after protamine removal [22]. Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant). After fertilization, maternal and paternal pronuclei migrated toward each other and apposed (documented embryos with confocal microscopy images $n = 4$, Fig 1A). Then, paternal and maternal chromatins condensed and entered into first mitotic division ($n = 3$, Fig 1B). During the first division, paternal and maternal chromosomes aligned in separate region at the metaphase plate ($n = 1$, Fig 1C). Both sets of chromosomes segregated equally during anaphase ($n = 3$, Fig 1D) to produce two diploid nuclei ($n = 1$, Fig 1E) that proliferate mitotically ($n = 16$, Fig 1F). After 24 hours of development, organogenesis was ongoing and segmentation was clearly visible ($n = 2$, Fig 1G).

In sterile crosses between two infected incompatible *C. pipiens* lines (Fig 2) as well as in crosses between infected males and uninfected females (Fig 3), paternal and maternal pronuclei migrated and apposed normally ($n = 2$, Figs 2A and 3A). However, during the early prophase, paternal chromatin appeared under-condensed compared to maternal chromatin ($n = 2$, Fig 2B and 2C). Then the paternal chromatin failed to segregate properly during anaphase ($n = 16$, Figs 2D, 2E, 3B and 3C). In telophase paternal chromatin can either i) form chromatin bridges between the two maternal nuclei ($n = 10$, Figs 2D and 3B), certainly causing the early arrest of embryogenesis and production of undeveloped embryos (Figs 2G1 and 3E), or ii) appear fully excluded ($n = 6$, Figs 2E and 3C), allowing maternal chromatin to

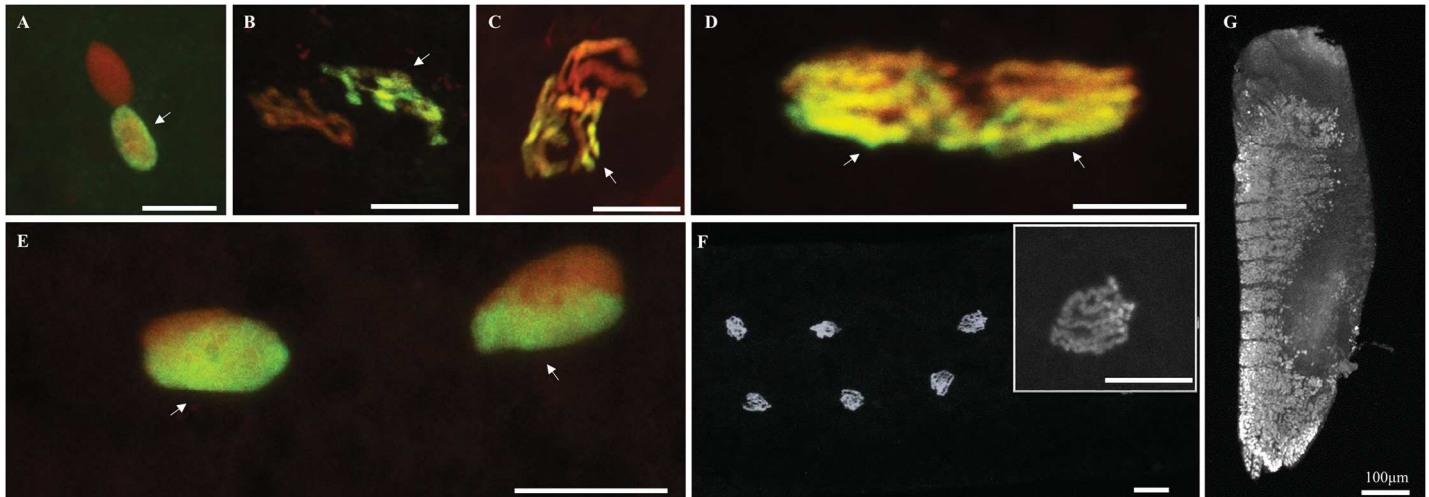


Fig 1. Normal embryogenesis in *C. pipiens*: From first nuclear divisions to segmentation. Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant). (A) apposition of maternal and paternal pronuclei, (B) chromatin under condensation, (C) condensed chromatin, (D) first mitotic division anaphase (maternal and paternal chromosome segregate independently), (E) two nuclei following the first division, (F) normal diploid development 1 hour after oviposition, six diploid nuclei are visible after 1h development, (G) segmented embryo after 24 hours of development. White arrows indicate the paternal chromatin. Confocal stacks were obtained on embryos from several fertile intra-line crosses due to the difficulty to obtain all the early embryonic stages from each cross (S2 Table). Scale bar is 10µm.

<https://doi.org/10.1371/journal.ppat.1007364.g001>

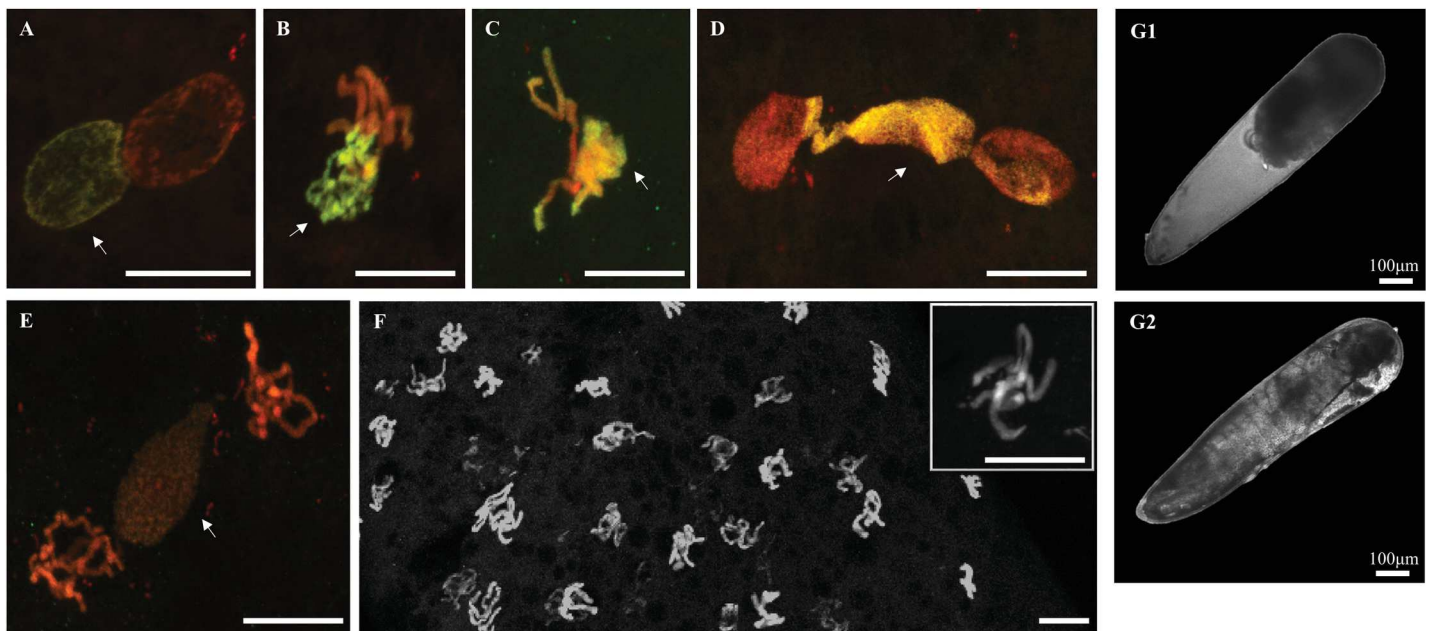


Fig 2. CI embryos from incompatible infected parents in *C. pipiens*: From first nuclear divisions to segmentation. Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant) (A) apposition of maternal and paternal pronuclei, (B) delay in paternal chromatin condensation, (C) condensed maternal chromatin and under-condensed paternal chromatin, (D) paternal chromatin failed to segregate and form a chromatin bridge between segregating maternal chromatin, (E) two nuclei containing mainly maternal chromatin while paternal chromatin do not segregate, (F) haploid development 2 hours after oviposition, (G) the two possible fates of development after 48 hours (1) non-viable embryo with no visible development, and (2) unhatched developed embryo with visible segments. White arrows indicate the paternal chromatin. Confocal stacks (panels A,B,C,D,E,F) and optical images (panels G1 and G2) were obtained on embryos from several CI crosses between infected males and females due to the difficulty to obtain all the early embryonic stages for each cross (S2 Table). Green dots are background noises likely due to the presence of residual antibodies. Scale bar is 10µm.

<https://doi.org/10.1371/journal.ppat.1007364.g002>

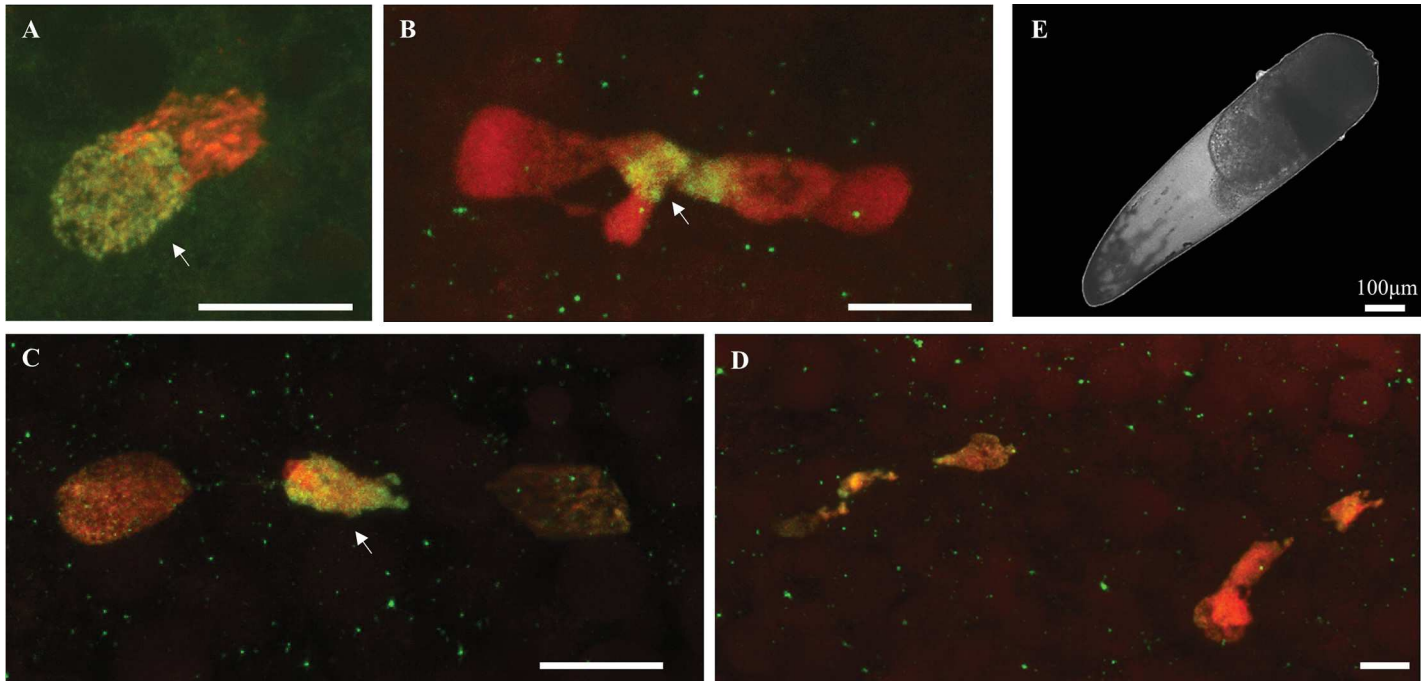


Fig 3. CI embryos from infected fathers and uninfected mothers: An arrest after the first nuclear divisions. Paternal chromatin appears in green/yellow (acetylated histone H4 labelling is dominant) and maternal chromatin appears in red (propidium iodide labelling is dominant) (A) apposition of maternal and paternal pronuclei, (B) paternal chromatin failed to segregate and form a chromatin bridge between segregated maternal chromatin, (C) two nuclei containing mainly maternal chromatin while paternal chromatin do not segregate, (D) abnormal development at 2 hours after oviposition: only few aborted divisions are observed, (E) none of the embryos from such crosses exhibited any visible development under microscope 48 hours post-oviposition. White arrows indicate the paternal chromatin. Confocal stacks (panels A,B,C,D,) and optical images (panels E) were obtained on embryos from several CI crosses between infected males and uninfected females due to the difficulty to obtain all the early embryonic stages for each cross (S2 Table). Green dots are background noises likely due to the presence of residual antibodies. Scale bar is 10µm.

<https://doi.org/10.1371/journal.ppat.1007364.g003>

successfully segregate and eventually formed unhatched haploid developed embryos presenting eyes and segments (Fig 2G2).

Eight distinct CI crosses were monitored between males and females infected with *wPip* strains belonging to different *wPip* groups, and two distinct CI crosses were monitored between infected males and uninfected females (S2 Table). However, despite this diversity of CI crosses, condensation and segregation defects of the paternal chromatin were the only observed cellular defects resulting in embryonic death, and were never observed in any embryo resulting from fertile crosses (seven intra-line fertile crosses, S2 Table).

Absence of *Wolbachia* in the oocytes blocks embryogenesis in CI embryos

2 hours after oviposition, some embryos resulting from CI crosses between infected lines pursued their embryogenesis ($n = 2$, Fig 2F), and after 48 hours these embryos exhibited visible development, as segmentation was clearly observable under optical microscope (Fig 2G2). However, more than 99.9% of these developed embryos did not hatch. In the sterile cross between ♂ Slab x ♀ Ichkeul 13, unhatched developed embryos only displayed maternal markers (see Material and methods, "Ploidy determination in CI developed embryos", S1 Fig), showing that they were composed of only haploid maternal DNA, as previously described in Duron and Weill (2006)[44].

All the seven different crosses performed between males infected with different *wPip* strains and uninfected females from different TC-treated lines produced 100% of non-developed embryos (Fig 3E and S3 Table). Confocal observations of these embryos showed that only few

and abnormal nuclei were observed in the cytoplasm 2 hours post oviposition ($n = 5$, Fig 3D), indicating an early arrest of the embryogenesis (S2 Table).

Males infected by different *wPip* with distinct *mod* profiles induced different CI defects intensities

It has been previously proposed that the production of haploid or aneuploid embryos in CI crosses represented a proxy of intensity of CI defects that leads to more or less complete paternal chromatin exclusion [29–31,44]. Severe defects would lead to the complete exclusion of the paternal chromatin during the first embryonic division (*i.e.* strong cellular CI intensity), which would allow maternal chromatin successful segregation and the production of a developed haploid embryo. Thus, unhatched developed haploid embryos reflect the occurrence during the first zygotic division of strong CI defects while unhatched non-developed embryos illustrate the occurrence of weak CI defects. We used this link between the degree of paternal chromatin exclusion (*i.e.* weak or strong cellular CI) and the proportion of unhatched developed embryos in eggs-rafts from incompatible crosses to investigate the variability of CI intensity (*i.e.* frequency of strong *versus* weak CI defects).

Using this proxy, we studied the variation in CI intensity between 20 incompatible crosses between infected lines (S3 Table). These 20 crosses involved i) males from four different isofemale lines (MAL lines) infected with *wPip* strains from different *wPip* groups all exhibiting distinct *mod* profiles, and ii) females from five isofemale lines (FEM lines) all harbouring *wPip* strains from the *wPip*IV group and exhibiting the same *resc* profile [36,42] (S4 Table). Significant differences were found regarding the proportion of unhatched developed embryos between these incompatible crosses (generalized linear model (GLM), $\chi^2 = 245.695$, $df = 19$, $p < 0.001$, Fig 4, Table 1 and S3 Table). While no effect of FEM lines was detected on this proportion (GLMM, $\chi^2 = 2.508$, $df = 4$, $p = 0.643$, Fig 4), the MAL lines involved in the crosses had a significant effect (GLMM, $\chi^2 = 16.211$, $df = 3$, $p = 0.001$, Table 1 and Fig 4). Males from Tunis (*wPip*I *mod* ii) and Slab (*wPip*III *mod* iii) lines induced the highest proportion of developed embryos (72% and 73%, respectively) but were not significantly different from one another (GLMM, $\chi^2 = 0.002$, $df = 1$, $p = 0.968$); males from Utique (*wPip*I *mod* iv) and Lavar (*wPip*II *mod* vi) lines induced significantly different and lower proportions of unhatched developed embryos (respectively 42% and 18%, Table 1). The nuclear genetic background of the males seems not to be involved in the variability of CI defects intensity: males from backcrossed line SI(*wPip*I-Tunis) and males from the Tunis line, which host the same *wPip*I strain in different genetic backgrounds, indeed induced similar unhatched developed embryos proportions when crossed with the five FEM lines (0.71 ± 0.22 and 0.72 ± 0.19 respectively; GLMM, $\chi^2 = 0.008$, $df = 1$, $p = 0.927$). Consequently variability in CI defects intensity appears to be only dictated by the *wPip* strain harbored by the different males.

Genetic investigations of cellular CI intensity variation

The results from the previous section indicate that the proportion of unhatched developed embryos in CI crosses likely depended on variations in the males' *mod* profiles. To investigate the sources of such variation in CI defects, we tested the putative influence of several variables: i) the density of *Wolbachia* in the testes, ii) the copy numbers of *cidA* and *cidB* genes in the different *wPip* genomes, iii) the expression levels of *cidA* and *cidB*, and iv) the *cidA* and *cidB* variants repertoires in the genomes of the different *wPip* strains hosted by the males.

***Wolbachia* density was significantly lower in Lavar males' testes.** Testicular *Wolbachia* densities were not significantly different between males from Tunis, Utique and Slab lines

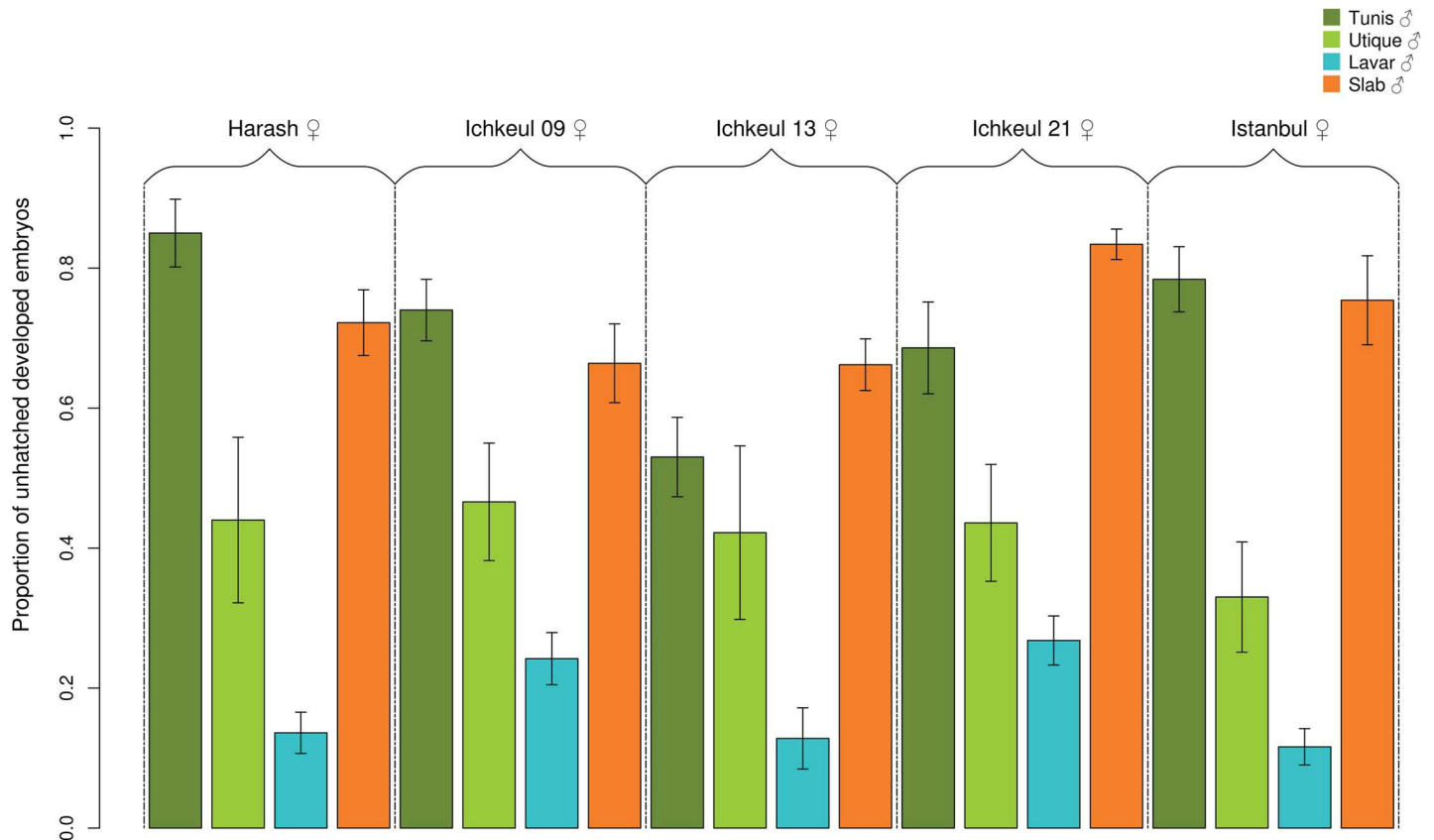


Fig 4. Variations in cellular CI intensity. Infected males from four lines exhibiting different *mod* profiles (MAL lines) and females from five lines with the same *resc* profile (FEM lines) were crossed and resulted in the 20 studied CI crosses (*i.e.* >99,9% of the embryos died before hatching). The vertical bars represent the proportion of unhatched developed embryos in each of these crosses, the four different colors represent the four different MAL lines, each group of four bars corresponding to the same FEM line. Error bars represent the standard error. The proportion of unhatched developed embryos was significantly different between crosses involving males infected with distinct wPip responsible for different *mod*; no significant effect of the FEM lines was detected.

<https://doi.org/10.1371/journal.ppat.1007364.g004>

(GLM, $F = 3.919$, $df = 1$, $p = 0.065$), but significantly lower in Lavar males (GLM, $F = 9.337$, $df = 3$, $p < 0.001$, Fig 5).

Table 1. Males with different *mod* profiles: Proportions of unhatched developed embryos, *cidA-cidB* copy number and expression.

Line	Proportion of unhatched developed embryo	<i>cidA</i> copy number	<i>cidB</i> copy number	<i>cidA/cidB</i> copy number	<i>cidA</i> expression	<i>cidB</i> expression	<i>cidA/cidB</i> expression
wPipI-Tunis	0.72 ± 0.19* (a)	5.98 ± 0.73 (a)	5.57 ± 0.56 (a)	1.08 ± 0.14 (a)	0.75 ± 0.18 (a)	0.50 ± 0.16 (a)	1.55 ± 0.30 (a)
wPipI-Utique	0.42 ± 0.31 (b)	5.02 ± 0.41 (b)	4.76 ± 1.00 (b)	1.08 ± 0.16 (a)	0.76 ± 0.27 (a)	0.49 ± 0.13 (a)	1.53 ± 0.32 (a)
wPipII-Lavar	0.18 ± 0.12 (c)	4.98 ± 0.83 (b)	4.14 ± 0.76 (c)	1.22 ± 0.22 (a)	0.97 ± 0.24 (a)	0.47 ± 0.15 (a)	2.14 ± 0.58 (b)
wPipIII-Slab	0.73 ± 0.16 (a)	4.69 ± 0.42 (b)	4.07 ± 0.47 (c)	1.17 ± 0.18 (a)	0.94 ± 0.26 (a)	0.60 ± 0.20 (a)	1.62 ± 0.32 (a)

*The average proportion of unhatched developed embryos, the number of copies of *cidA* and *cidB* and their ratio, as well as the expression levels of *cidA* and *cidB* and their ratio are indicated as the MAL line means ± standard deviations.

a, b, c letters represent statistical groups (*i.e.* means with the same letter are not significantly different).

<https://doi.org/10.1371/journal.ppat.1007364.t001>

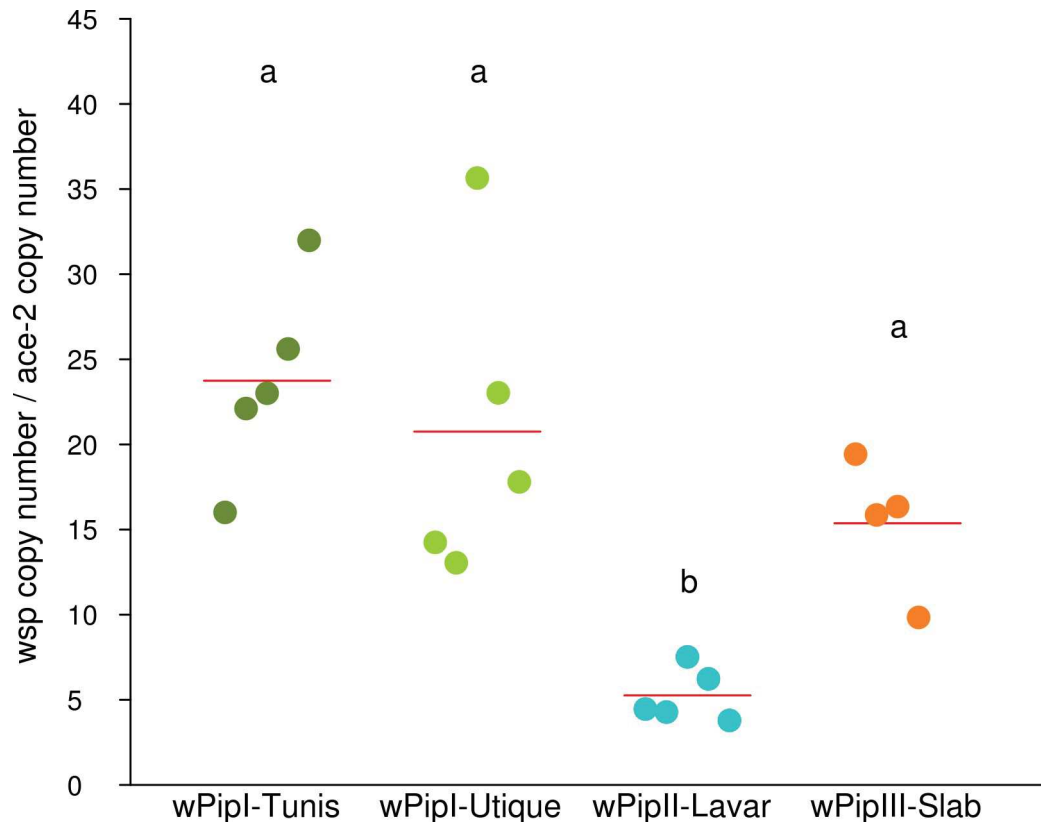


Fig 5. *Wolbachia* density in testes. *Wolbachia* densities in mosquito testes were measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia wsp* gene and the *C. pipiens ace-2* nuclear gene. The colored dots represent the average density of *Wolbachia* in a pool of three pairs of testes and the red strips represent the average *Wolbachia* density for each line/*wPip* strain. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different), showing that Lavar males harbored significantly less *Wolbachia* in their testes than the other males.

<https://doi.org/10.1371/journal.ppat.1007364.g005>

***CidA* and *cidB* copy numbers differed between the *wPip* strains.** Number of genomic copies of *cidA* per *wPip* strain (*i.e.* *Wolbachia* cells from the same *C. pipiens* line) varied between MAL lines from 4.68 ± 0.42 to 5.98 ± 0.73 copies, *wPipI-Tunis* displaying a significantly higher *cidA* copy number than the other strains (GLM, $F = 8.077$, $df = 3$, $p < 0.001$, Table 1, S5 Fig). Number of genomic copies of *cidB* also varied from 4.07 ± 0.47 to 5.57 ± 0.56 copies (GLM, $F = 9.142$, $df = 3$, $p < 0.001$), and was found significantly higher in *wPipI-Tunis* than in *wPipI-Utique*, *wPipIII-Slab* and *wPipII-Lavar* displaying significantly lower but similar copy numbers (Table 1, S6 Fig). Despite these differences, the *cidA/cidB* copy number ratios were not significantly different and close to one, for all *wPip* strains (GLM, $F = 1.504$, $df = 3$, $p = 0.230$, Table 1, S7 Fig).

Variation in *cidA* and *cidB* relative expression levels. The expression levels of both *cidA* and *cidB* were not different between the four *C. pipiens* MAL lines (GLM, *cidA*: $F = 2.409$, $df = 3$, $p = 0.083$, *cidB*: $F = 1.239$, $df = 3$, $p = 0.310$, Table 1, S8 and S9 Figs). *CidA* was found about 1.5 times more expressed than *cidB*, except for *wPipII-Lavar* which *cidA/cidB* expression level ratio appeared significantly higher (GLM, $F = 5.447$, $df = 3$, $p = 0.003$, Table 1, S10 Fig).

***CidA* and *cidB* variants repertoires were variable between *wPip* strains.** Cloning and Sanger sequencing revealed that the four MAL lines with different *mod* were infected with *Wolbachia* strains that harbored different *cidA* and *cidB* variant repertoires (S2 and S3 Figs). No

cidA or *cidB* nucleotide sequence variant were shared between the three *wPip* groups. However, the CidA_II($\alpha/1$) variant of *wPip*II-Lavar and the CidA_III($\beta/8$) variant of *wPip*III-Slab presented the same amino-acid sequence (S2 Fig). *wPip*III-Slab exhibited ten variants of *cidA* based on their nucleotide sequences (however only seven of them differed in their amino-acid sequences), *wPip*II-Lavar three, *wPip*I-Tunis four and *wPip*I-Utique seven (S2 Fig). Both *wPip*II-Lavar and *wPip*I-Utique carried two different variants of *cidB*, while *wPip*I-Tunis and *wPip*III-Slab carried four (S3 Fig).

Correlations between CI defects intensity and *wPip* genetic variations

Correlative analyses were conducted to assess the potential links between variations in CI defects intensity and genetic variations. We found no significant correlations between the proportion of unhatched developed embryos in CI crosses and i) *Wolbachia* density in the testes (Spearman, $\rho = 0.4$, $p = 0.750$), ii) *cidA* copy number (Spearman, $\rho = -0.2$, $p = 0.917$), iii) *cidB* copy number (Spearman, $\rho = -0.2$, $p = 0.917$), iv) *cidA/cidB* copy number ratio (Spearman, $\rho = -0.4$, $p = 0.750$), v) *cidA* expression levels (Spearman, $\rho = -0.4$, $p = 0.750$), vi) *cidB* expression levels (Spearman, $\rho = 1$, $p = 0.083$), vii) *cidA* over *cidB* expression levels (Spearman, $\rho = -0.2$, $p = 0.917$) and viii) the number of different *cidA* variants in the repertoires (Spearman, $\rho = 0.8$, $p = 0.333$). However, males infected with *wPip* strains with 4 *cidB* variants induced significantly higher proportions of unhatched developed embryos (*wPip*I-Tunis and *wPip*III-Slab mean: 0.72 ± 0.17) than males infected with *wPip* strains with only 2 *cidB* variants (*wPip*II-Lavar and *wPip*I-Utique; mean: 0.30 ± 0.26 , Wilcoxon, $W = 1159$, $p < 0.001$, S4 Fig).

Discussion

To investigate whether the high diversity of *cidA/cidB* variants within *wPip* could be responsible for variations in the cellular phenotype of CI, we studied the development of *C. pipiens* embryos resulting from various incompatible crosses. The early embryogenesis was assessed using fluorescence confocal microscopy in i) fertile intra-line crosses, ii) incompatible crosses between infected males and infected females, and iii) incompatible crosses between infected males and uninfected females. Despite the diversity of performed crosses between males and females infected with *wPip* strains harboring different *cidA/cidB* variants repertoires or uninfected female, a unique and recurrent embryonic phenotype was detected, consisting in paternal chromatin condensation and segregation defects during the first embryonic division (Figs 2B, 2E, 3B and 3C). This phenotype was never detected in any embryos derived from intra-line crosses (Fig 1). Hence the diversity of *cidA/cidB* variants repertoires describes in *C. pipiens* does not seem to influence the CI mechanism itself, which is consistent with all CidB variants carrying a conserved DUB domain [36]. Similar defects were already reported in both *Drosophila* and *Nasonia* [23,25], suggesting an universality of Cid induced-cellular CI mechanism whenever *cid* genes are diversified or not in the *Wolbachia* genome.

An unsolved question is the molecular pathway(s) targeted by CidA and CidB. Most protein domains within CidA and CidB remain to be characterized and how they interact with each other and host targets to induce CI remains unclear. However, a first tangible element is that the catalytically active DUB domain (involved in deubiquitination) in CidB proteins, which is considered as involved in the *mod* function, is necessary to induce CI in transgenic *Drosophila* [34]. Ubiquitination pathways have been shown to be crucial for many essential cellular processes, such as the regulation of the chromatin dynamics and the cell cycle progression [45]. Changes in ubiquitination could for instance directly or indirectly affect H3.3 histone incorporation after protamine removal and DNA replication as suggested by PCNA persistence on the paternal chromatin [22], which would result in an asynchronous mitotic entry of paternal and

maternal pronuclear chromatin [30]. Interestingly, *Cardinium*, an endosymbiont phylogenetically distant from *Wolbachia*, induces CI with quite similar embryogenesis defects in the hymenoptera *Encarsia suzanna* [46]. Moreover, an ubiquitin specific protease USP classified as a DUB protein has also been detected in *Cardinium* genome, suggesting a convergent implication of DUB in CI induced by insect endosymbionts [47]. However, some *Wolbachia* strains able to induce CI do not carry DUB domain (*i.e.* no *cid*) in their genomes, but display instead a paralog gene with a nuclease domain called *cinB* [34,35,38]. DUB (*Cid*) and Nuclease (*Cin*) domains do not have the same predicted functions suggesting that distinct molecular pathways may be responsible for CI [34,38]. The CI cellular defects caused by *Wolbachia* strains harboring only *cin* genes remain unknown and could differ from the one induced by *cid* genes. Our study showed that *wPip* strains, which carry both *cid* and *cin* genes in their genomes, induce similar defects during embryogenesis as *wMel*, which carries only a *cid* gene. This suggests that the association of *cid* and *cin* does not change the cellular phenotype of CI, but the molecular mechanism induces by DUB and Nuclease which must be different due to the biochemical nature of the proteins might converge on a similar cellular defect (*i.e.* paternal chromatin condensation defect). However, the presence of DUB and Nuclease domains in the same *Wolbachia* genome could still contribute to CI by modifying its penetrance: *wRi* (*D. simulans*) and *wPip* (*C. pipiens*), which harbor both paralogs, have indeed a strong CI penetrance (almost no hatched embryos), while *wNo* and *wHa* (*D. simulans*), which carry either *cin* or *cid* genes, respectively, induce lower CI penetrance [13,38].

Our cytological investigation in *C. pipiens* evidenced a link between the paternal chromatin exclusion degree during the first zygotic division and the existence of two developmental fates following first-division defects. In fact, unhatched embryos can either reach advanced developmental stages, exhibiting segments and visible eyes, or display no visible development (Figs 2G1, 2G2 and 6) [26,44]. We confirmed Duron and Weill (2006)[44] findings that the unhatched developed embryos resulting from CI were haploid, and carried genetic material from maternal origin only (Fig 6 and S1 Fig). Confocal observations showed that such haploid development likely occurred when paternal chromatin was fully excluded during the first zygotic division, allowing the successful segregation of the isolated maternal chromatin (Fig 6). In contrast, unhatched non-developed embryos would be due to partial exclusion of the paternal chromatin, which would result in aneuploid nuclei and early arrest of embryogenesis (Fig 6). It has been previously proposed for other arthropod models that the participation of paternal chromatin to the first division would depend on the intensity of paternal chromatin defects (*i.e.* improper condensation) [29,30,44,48]. Severe defects would lead to complete paternal exclusion (*i.e.* strong cellular CI) and to the production of haploid developed embryos, while less severe defects would lead to a partial paternal chromatin exclusion (*i.e.* weak cellular CI) and to the production of aneuploid non-developed embryos. We used this link between the degree of paternal chromatin exclusion and the ratio of unhatched developed and non-developed embryos in eggs-rafts from incompatible crosses to investigate the variability of cellular CI intensity between different incompatible crosses.

We first studied the variability of CI intensity using males and females both infected with incompatible *wPip* strains. Developed embryos were observed in all these incompatible crosses, with two possible outcomes: i) less than one per thousand of these embryos were apparently not affected by CI and hatched into diploid larvae [14,44], and ii) from 11% to 85% of the unhatched embryos, depending on the crosses, reached late embryonic developmental stages showing that they experienced strong CI defects (Figs 4 and 6). We then studied the influence of the absence of *Wolbachia* in the oocytes on the cellular CI intensity. As in Duron and Weill (2006)[44], i) we confirmed that not a single larvae was produced in such crosses, and ii) all the seven CI crosses between infected males and uninfected females (TC lines)

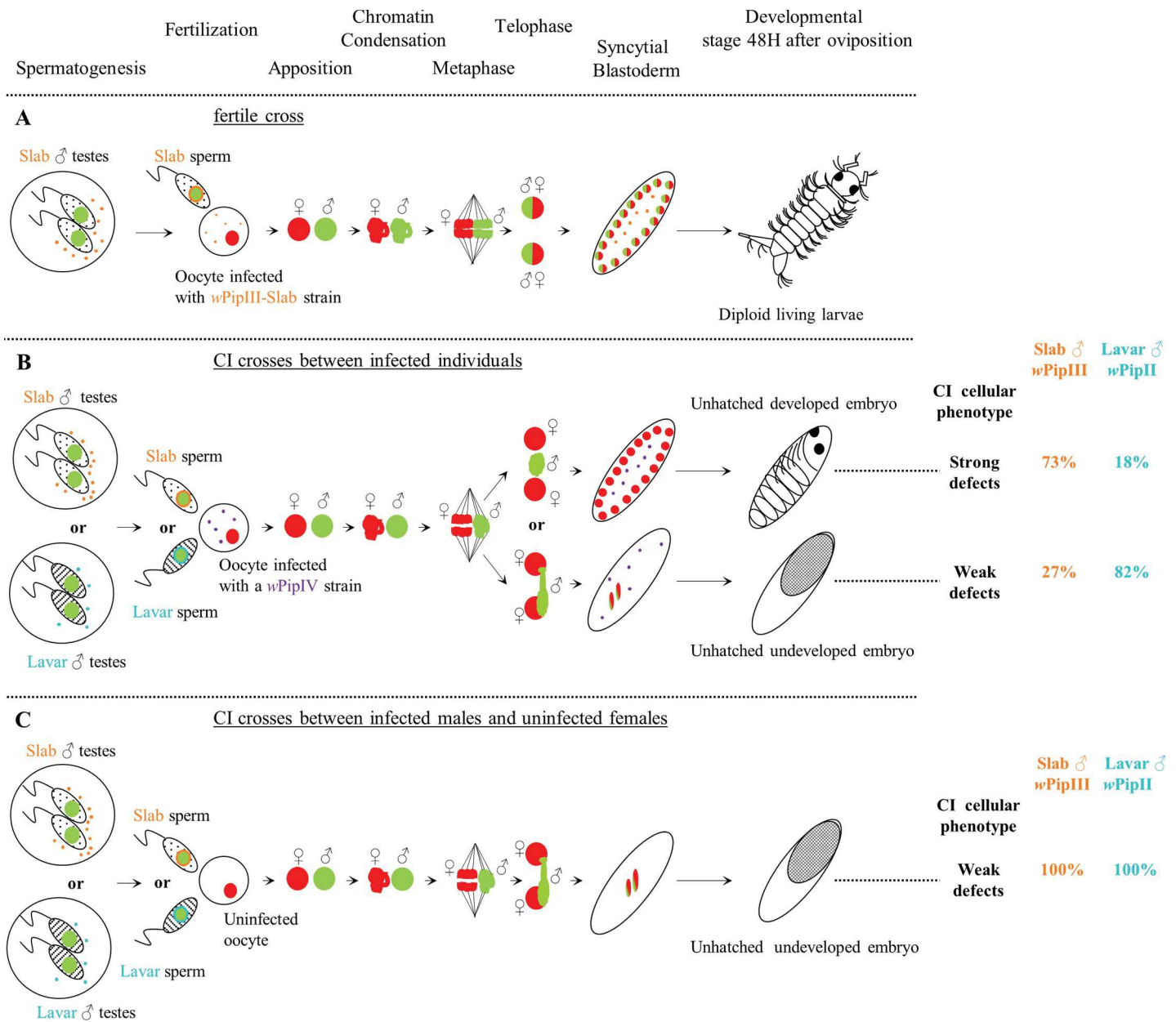


Fig 6. Cellular bases of different developmental fates in *C. pipiens*: From spermatogenesis to late development stages. Paternal and maternal DNA are represented in green and red respectively and *Wolbachia* cells are represented by the small dots (blue for *wPipII*; orange for *wPipIII* and purple for *wPipIV*). The two *MAL* lines Slab and Lavar have been chosen to illustrate the variability in CI defects intensity (*i.e.* frequency of strong and weak cellular CI). During spermatogenesis, the *Wolbachia* in male testes produce a toxin (*mod* factor). After fertilization, if females and males are infected with the same *Wolbachia* (Panel A), they can prevent the toxicity (*resc* function) and allow normal embryogenesis and the production of living diploid larvae. However, if the *Wolbachia* in the females are incompatible with the *Wolbachia* in the males (Panel B) or absent (Panel C) then the toxicity would not be prevented and paternal chromatin condensation delay and segregation defects occur. Two outcomes of the first zygotic division are possible regarding the degree of paternal chromatin exclusion. If the paternal chromatin is partially excluded, chromatin bridges would be formed resulting in aneuploid nuclei which might go through few mitotic divisions but will arrest the embryogenesis at early stages producing embryos with no visible development. If the paternal chromatin is fully excluded from the first zygotic division, maternal chromatin can segregate and produce two haploid nuclei which will divide further to produce non-viable haploid embryos exhibiting advanced development stages with eyes and segmentation clearly visible. The proportion of unhatched haploid embryos is influenced by different *wPip* with distinct *mod* profiles in the *MAL* lines involved in the crosses. However when the egg is not infected by *Wolbachia* (Panel C), all embryos exhibited an absence of further development after the first zygotic divisions whatever the *mod* induced by the *wPip* hosted by the *MAL* lines.

<https://doi.org/10.1371/journal.ppat.1007364.g006>

resulted in 100% of non-developed embryos suggesting that in such crosses, CI phenotype was always weak (Fig 6).

In crosses between infected individuals, it clearly appeared that MAL lines harboring *Wolbachia* from different wPip groups (wPip I, II, III) and displaying distinct *mod* induced significant variation in CI defects intensity when crossed with females harboring distinct wPipIV strains displaying the same *resc* (Figs 4 and 6). Variation in CI defects intensity has already been reported in *Nasonia* species, where the production of haploid viable males in *N. vitripennis* was interpreted as resulting from severe paternal chromatin defects, while the production of unviable aneuploid embryos in *N. longicornis* and *N. giraulti* was interpreted as resulting from weak paternal chromatin defects. However, variation of CI intensity in these host species was not associated with the different *Wolbachia* strains, but to variation in host genetic backgrounds [19]. The backcross experiment performed in the present study suggests that CI intensity is not impacted by nuclear genetic variations in *C. pipiens*. While it was already established that *Wolbachia* drives alone the observed variation in crossing types in *C. pipiens* [11,49,50], *Wolbachia* also seems to dictate the intensity of CI defects. Consequently, the variation in CI intensity observed when two infected individuals are crossed seems to be under the major influence of the wPip strain infecting the MAL line via the degree of paternal chromatin exclusion they trigger.

In *C. pipiens*, when females from tetracycline-cleared lines (TC females) were crossed with the four MAL lines, 100% of unhatched non-developed embryos only exhibiting few degenerated nuclei were observed, even 2 hours after oviposition (Figs 3D and 6). Such CI phenotype suggests that the defects caused by the wPip infecting all the MAL lines are always weak (Fig 6). This result is counter intuitive because one would expect that when *Wolbachia* is absent from the eggs CI should be always strong and many haploid embryos should be produced. We mentioned above that all the MAL lines can induce strong CI defects in variable proportion of the embryos when crossed with infected females. Consequently, the constant weak CI phenotype observed when females are not infected is linked to the absence of *Wolbachia* during egg maturation. Our results suggest that in incompatible crosses between infected *C. pipiens* individuals, the presence of maternal *Wolbachia* somehow interferes with early embryogenesis allowing haploid development to occur. It seems very unlikely that the presence of incompatible *Wolbachia* in the egg would enhance the mechanisms leading ultimately to paternal chromosome condensation defects (*i.e.* accentuate the *mod* function) to result in its total exclusion during the first embryonic division. Instead, the presence of incompatible *Wolbachia* in the eggs may have an additive effect on the incompatibility between pronuclei, not by directly affecting the paternal chromatin but by influencing the cell cycle timing. For instance, maternal *Wolbachia* could modulate the maternal kinetics for DNA replication or the mitotic entry during early development, increasing the incompatibility between pronuclei and therefore favoring the haploid development. Thus, while paternal *Wolbachia*-induced CI defects always occur regardless of the infection status of the eggs, the absence of incompatible maternal *Wolbachia* would block haploid development resulting in weak CI phenotype.

We then investigated the putative genetic determinism of CI intensity variation in embryos derived from infected parents. We assessed whether it could result from difference in *Wolbachia* density, *cidA-cidB* gene expression, copy numbers, or variant diversity between the wPip strains. As previously described in *Drosophila* [35,38], we found in *C. pipiens* that *cidA* was always significantly more expressed than *cidB*, whatever the wPip strain (Table 1). This is in accordance with the hypothesis that *cidA* and *cidB* form a toxin-antidote system where CidA is the antidote of CidB [34,36]. Indeed, in such system the antidote was always found more expressed than the toxin to prevent the host from toxicity [51]. No significant difference between MAL lines was found for *cidA* and *cidB* expression levels per *Wolbachia* cell (Table 1),

suggesting that the *cidA* and *cidB* expression does not influence CI defects intensity. However, while the *cidA* and *cidB* expression levels per *Wolbachia* cell did not significantly vary between *C. pipiens* lines, the total amount of CidA and CidB proteins in the host mainly depends on the density of *Wolbachia*. Since the *mod* factors are most likely deposited on the sperm in the testes during spermatogenesis [32,33], we measured the density of *Wolbachia* in the male gonads. We found that Lavar males hosted significantly less *Wolbachia* in their testes than males from the three other lines (Fig 5); Lavar males were also those that generated the lowest proportion of unhatched developed embryos in their offspring, whatever the FEM lines (Fig 4). Due to lower *Wolbachia* density in the testes, the global amount of CidB protein could be lower in Lavar line compared to the other lines. This low dosage of CidB would more likely result in weak CI defects leading to only few haploid development. However, this hypothesis relies on a single line and requires more *C. pipiens* lines with distinct testicular *Wolbachia* densities to be confirmed. Lavar was also the line with the highest *cidA* expression relatively to *cidB* (Table 1); as CidA has been proposed as the CidB antidote [34,36], its overexpression could reduce CidB-induced CI defects, and contribute to the low frequency of developed haploid embryos observed in crosses involving males from Lavar line.

We previously demonstrated that the amplification followed by the diversification of *cidA* and *cidB* variants in *wPip* certainly constitutes the source for CI diversity profiles in *C. pipiens* while *cinA* and *cinB* did not exhibit any polymorphism [36]. Indeed, specific variations in *cidA* and *cidB* repertoires (number and/or nature of the variants) clearly seemed to determine the compatibility outcome of crossings between *wPip*IV-infected males and any infected females, pointing out the putative role of these variations in the prodigious CI complexity recorded in this species [36,42]. Here, we tested the putative consequence of *cidA* and *cidB* gene amplification (*i.e.* number of copies per genome) on variation of CI defects intensity, and demonstrated no significant correlation between the two parameters. When the quantification of genomic copies obtained by q-PCR are put in relation to the number of different variants in the same isofemale line obtained by cloning-sequencing, some of *cidA* results appear discordant. This is especially true for the Slab line, which exhibits ten distinct *cidA* variants for ~5 copies per genomes quantified (Table 1 and S2 Fig). Even taking into account technical limits of q-PCR to quantify high level of gene amplification, this discordance suggests that, at least in the Slab line, some of the *Wolbachia* cells do not harbor the same *cidA* variants.

We found that the different *wPip* strains carried by the four MAL lines exhibiting different *mod* profiles harbored distinct *cidB* variants. Any variant of this gene could certainly trigger CI alone, as the DUB domain is perfectly conserved between all variants [36]. However, their diversity can modulate CI defects intensity. We thus tested whether *cidB* repertoire diversity could play a role in CI intensity variability. Supporting this hypothesis, we found that males from the two *C. pipiens* lines harboring *wPip* strains with four different *cidB* variants induced higher proportions of unhatched developed embryos compared to lines harboring *wPip* with only two different *cidB* variants (S4 Fig). Each distinct *cidB* variants could differentially impact the paternal chromatin (*i.e.* like different locks), putatively leading to an additive *mod* effect: the more different *cidB* variants present in a *wPip* strain, the more likely strong CI defects. However, more *wPip* strains varying in their diversity of *cidB* are required to further test this hypothesis.

In conclusion, despite the diversity of crossing types observed in *C. pipiens*, linked to the diversity of *cidA/cidB* variants repertoires, a single cellular phenotype of CI, was observed in this species. In all crosses (*i.e.* uni-bidirectional), CI results in early developmental defects in the paternal chromatin condensation and segregation during the first zygotic division similar to that observed in other insects. Our study demonstrates that in CI crosses between two infected individuals, the CI intensity (*i.e.* frequency of strong and weak CI defects) is

influenced by the male-carried *wPip*. However, when the female is not infected, and despite the variability of the distinct *wPip* strains carried by the males, no unhatched developed embryos (strong cellular CI) were ever found, suggesting that the weak CI phenotype observed in such crosses is instead due to the absence of *Wolbachia* in the eggs. Genetic investigation reveals that the variability of CI defects intensity may be linked to *cidB* variant diversity in *wPip* strains. While the putative functional role and the singularity of *cidB* amplification and diversification in *wPip* remains yet to be fully solved, it clearly appears that it deeply modifies the *wPip*-induced CI phenotype at different scales, from crossing types [36] to its intensity at the cellular level.

Materials and methods

Culex pipiens lines

***C. pipiens* lines hosting different *Wolbachia* strains or without *Wolbachia*.** Ten laboratory isofemale lines belonging to *C. pipiens s.l.*, from our laboratory, were used; they differ in their geographical origins and in the *Wolbachia* strain hosted (S1 Table). To determine the *mod* and *resc* profiles of the different mosquito lines, crosses with four reference lines (4-ref cytotypes) were realized similarly to Atyame et al. (2014)[42] (S4 Table). To study the effect of the absence of *Wolbachia* on embryogenesis, tetracycline-treated *Wolbachia*-free lines (TC lines) SlabTC, IstanbulTC, Ichkeul 21TC, and Ichkeul 13TC were obtained respectively from Slab, Istanbul, Ichkeul 21 and Ichkeul 13 *wPip* infected lines, as described in Duron et al. (2006)[11]. After TC treatment, PCR amplifications of a fragment of *wsp* gene using the primers designed in Berticat et al. (2002)[52] allowed controlling the absence of *Wolbachia* in DNA extracted from a larvae pool (Dneasy Blood & Tissue Spin-Column protocol Kit; Qiagen; Bench protocol: Animal Tissues). To prevent possible side-effects of the treatment, TC-treated lines were raised in standard laboratory conditions for at least four generations without tetracycline before the beginning of experiments.

***C. pipiens* lines with the same host genetic background but different *Wolbachia* strains.** *wPipI* strain from Tunis line was introduced into Slab line nuclear genetic background through 8 backcrosses. For the first generation, 200 females from Tunis line were crossed with 100 males of the SlabTC line. Then, for each of the seven following generations, 200 females from the previous generation were crossed with 100 males from the SlabTC line. This led to a progressive replacement (over 97%) of the maternal nuclear genome (Tunis) by the paternal nuclear genome (SlabTC), with retention of the maternal cytoplasm, including the Tunis *Wolbachia* strain [this line was called SI(*wPipI*-Tunis)].

Cellular study of early embryogenesis

To characterize CI cellular phenotype(s) in *C. pipiens*, several crosses were performed (S2 Table). For every crosses, to avoid confounding age effects, two-day old adults were released in cages. Cages containing 100 females and 50 males were then put into a closet at 25°C where day-night cycle was inverted to allow collection of early developmental stage eggs during the day. After six days in these cages, females were fed with turkey blood in heparin sodium (bcl Wholly Wild World) using a Hemotek membrane feeding system (Discovery Workshops, United Kingdom). Five days after blood meal, water-pots were placed into the cages to collect the eggs-rafts. For *C. pipiens* eggs, at 25°C, the meiosis is approximatively completed 30 minutes after the oviposition and the first mitotic nucleus division 15 minutes after the end of the meiosis, while four hours after oviposition the embryos normally reach the syncytial blastoderm stage [53]. Since, the CI defects described in *D. simulans* [25] and *N. vitripennis* [23,24] occurred during the first nucleus mitotic division, we mainly collected eggs aged from 30 minutes to

1 hour. Older eggs were also harvested to monitor further developmental stages in both fertile and sterile crosses. Eggs-rafts were then placed into commercial bleach (active ingredient, 9.6% of sodium hypochlorite) to dissociate eggs, and then washed in distilled water. They were then fixed by being shaken for 2 hours in a solution of 3.2% para-formaldehyde in PBS 1X with Tween 0.02% (PBS-T) and washed with PBS 1X. For each fixed egg, the chorion was removed manually with a needle under an optical microscope (Leica MZ 8). Dechorionated embryos were then collected and treated with RNase A (10 mg/mL, Sigma) overnight.

To differentially visualize the paternal from the maternal chromatin, we used propidium iodide to mark both chromatin and an anti-acetylated histone H4 labelling that preferentially marks the *de novo* assembled paternal chromatin after protamine removal [22]. Thus maternal and paternal chromatin will be respectively predominantly marked with propidium iodide (mostly red fluorescence) and with anti-acetylated histone H4 antibodies (mostly green fluorescence). For immunolabeling, embryos were first incubated overnight at 4°C with primary antibodies (Polyclonal anti-acetylated histone H4 primary antibody (1:1000, Upstate)), washed during one day with PBS-T 1X, then incubated overnight at 4°C with the secondary antibody (Alexa Fluor 488 goat anti-rabbit IgG secondary antibodies (1:250, Invitrogen)) then washed with PBS-T 1X. Embryos were then incubated in PBS-T 1X for 20 minutes with propidium iodide a DNA intercalating agent (Molecular Probes, 10 μL/1 mL). Finally, embryos were washed for 5 minutes and mounted between slide and coverslip in Fluoroshield Mounting Medium (Vector). Confocal microscope images were captured on an inverted photomicroscope (DMIRB; Leitz) equipped with a laser confocal imaging system (TCS SP5; Leica) using an HCX PL APO 1.4 NA 63 oil objective (Leica). Images from fixed, immunostained embryos are merged confocal z-stacks taken sequentially in the green and red channels for the anti-acetylated histone H4 labelling and the propidium iodide signal respectively. Crosses from which confocal microscope images were obtained (Figs 1–3) are listed in S3 Table.

Proportion of unhatched developed embryos in CI crosses

To study the proportion of unhatched developed embryo in CI crosses, we performed a total of 32 crosses: 20 crosses involving four lines for the males (MAL lines) and five lines for the females (FEM lines), 5 involving *Sl(wPipI-Tunis)* for the males and the five FEM lines, and 7 involving the four MAL lines and females from different TC lines (S2 Table). All these crosses were performed using 50 females and 25 males. After 6 days in the cages, females were blood-fed and after 5 days eggs-rafts were collected in water pot and deposited into 24 wells plates. As hatching normally occurs approximately 48 hours after oviposition, developmental status in non-viable rafts was characterized at least two days after eggs-rafts collection. To attribute a developmental status to each egg, eggs-rafts were mounted between slide and coverslip, observed and documented with an optic microscope (Axiophot2 equipped with a CCD camera, Zeiss). Two developmental statuses were discriminated i) unhatched embryos harboring no visible development (Fig 2G1), or ii) unhatched embryos with visible development (Fig 2G2). For each cross, we calculated the proportion of embryos showing development for 50 embryos per eggs-raft in 10 eggs-rafts (total of 500 eggs observed per cross).

Ploidy determination in CI developed embryos

To assess the ploidy status in unhatched developed embryos, we used a PCR/RFLP diagnosis *kdr/RsaI* that allowed discriminating between *C. pipiens* and *C. quinquefasciatus* lines, as previously described in Duron and Weill (2006)[44]. Slab (*C. quinquefasciatus*) and Ichkeul 13 (*C. pipiens*) were chosen because they exhibit an unidirectional sterile cross: fertile in the direction (♂Ichkeul 13 x ♀Slab) and sterile in the other direction (♂Slab x ♀Ichkeul 13). This PCR/RFLP test was performed

on DNA extracted as describe above from i) a pool of larvae from Slab and Ichkeul 13 parental lines and ii) from eggs-rafts resulting from the two reciprocal crosses between those two lines.

Real Time Quantitative PCR

Quantification of *Wolbachia* density in male testes. In order to test the influence of *Wolbachia* densities in testes on the CI cellular intensity, we quantified them with Real Time Quantitative PCR using the LightCycler 480 system (Roche). Specific primers and procedures were described in Berticat et al. (2002)[52]. Testes from 6-day old males of Tunis, Utique, Lavar and Slab lines were sampled. Each DNA template were obtained from pools of three testis pairs and extracted as described above. Five independent DNA templates were realized for each line. To estimate the number of *Wolbachia* per mosquito testes, we amplified two different genes on each sample, the *C. pipiens* specific *ace-2* locus [54] and the *Wolbachia* specific monocopy *wsp* locus [52]. Standard curves were performed using dilutions of a pBluescriptKS vector containing a unique *ace-2* and *wsp* gene copy. Each DNA template was analyzed in triplicate for both *wsp* and *ace-2* quantification. As both genes are present as single copies per haploid genome, the ratio of *wsp* over *ace-2* signals allowed estimating the relative number of *Wolbachia* genomes per *Culex* genome, thus correcting for mosquito size and DNA quality.

Amplification of *cidA* and *cidB* genes within *wPip* genome. For each *C. pipiens*-*Wolbachia* line (Tunis, Utique, Lavar and Slab hosting a different *wPip* strains belonging to group I, II or III) and exhibiting different *mod* profiles, quantitative PCRs were carried out to estimate the number of copies of *cidA* and *cidB* genes per *wPip* genome. Three different quantitative PCRs were performed on DNA samples extracted from ten 6-day old males per line following the procedure described in Berticat et al. (2002)[52]: i) specific of the locus *wsp*, ii) specific of a 189bp fragment of the *cidA* gene conserved between all *wPip* strains using primers *wPip_0282_QPCR_2_Dir* (5'-AGG-TCC-TGT-ATT-TGA-TTT-CTG-GA) and *wPip_0282_QPCR_2_Rev* (5'-TGA-ACG-CGA-GAA-AGA-GCA-AG), and iii) specific of a 135bp fragment of the *cidB* gene conserved between all *wPip* strains using primers *wPip_0283_QPCR_1_Dir* (5'-TGA-GTG-TTT-GGA-GAA-TGA-AGG-A) and *wPip_0283_QPCR_1_Rev* (5'-TTC-CCA-AAA-GCA-AAA-CCA-GTT). Standard curves of *cidA*, *cidB* and *wsp* genes were performed using dilution of the PCR product of these three genes previously quantified using the fluorometre-QuBit (Invitrogen). Each DNA template was analysed in triplicate for *wsp*, *cidA* and *cidB* locus. *CidA* and *cidB* copy numbers were estimated using the ratio of *cidA* or *cidB* estimated copy number over *wsp* estimated copy number, to obtain a copy number per *Wolbachia* genome since *wsp* is present in one copy in *Wolbachia* genome.

Expression of *cidA* and *cidB* genes. For each line Tunis, Utique, Lavar and Slab, ten 6-day old males were used for RNA extraction with Trizol (Life Technologies) and treated with DNase with the TURBO DNA-free Kit (Life Technologies), in accordance with the manufacturer's instructions. 2–5 µg of each total RNA sample were reverse-transcribed into cDNA with the SuperScript III Reverse Transcriptase Kit and 30 ng of random oligomer primers ((RP)10; Invitrogen, Life Technologies). Three different quantitative PCRs were performed: i) specific of *wsp* locus, ii) specific of *cidA* locus, and iii) specific of *cidB* locus as describe above. Each DNA template was analysed in triplicate for *wsp* *cidA* and *cidB* locus. Levels of expression of *cidA* and *cidB* genes were estimated relatively to *wsp* genes by using the ratio of expression of these two genes over *wsp*.

Determination of *cidA/cidB* variants repertoire in the different *wPip* hosted by males

To describe the diversity of *cidA*^{*wPip*}/*cidB*^{*wPip*} repertoires for the two *C. pipiens* lines Utique and Slab not yet investigated, cloning and Sanger sequencing of the *cidA* and *cidB* variants

were performed as described in Bonneau et al. (2018)[36] on DNA from pools of larvae extracted as described above. Variant sequences were aligned, using the Muscle algorithm implemented in Seaview 6.4.1 software [55].

Statistical analysis

Variability of unhatched developed embryo proportion in sterile crosses was analyzed using a generalized linear model (GLM): $U_{DEP} = CROSS + \epsilon$, with U_{DEP} the proportion of unhatched developed embryos for each cross ($CROSS$, which represent the interaction between the MAL and FEM lines) and ϵ the error parameter, following a binomial distribution. To test the specific effect of the four MAL lines and the five FEM lines separately, GLMs with mixed effects (GLMM) were used: $U_{DEP} = MALE + FEMALE + 1|CROSS + \epsilon$ with $MALE$ and $FEMALE$ respectively the MAL and FEM lines involved in each cross as fixed effects, with $CROSS$ as the interaction between MAL and FEM lines as a random effect (as crosses to produce embryos necessary require an interaction between females and males), and ϵ the error parameter, following a binomial distribution. To test for a specific effect of the host genetic background in crosses involving males from SI(*wPipI*-Tunis) and Tunis lines which host the same *Wolbachia* in two different genetic background we used a GLMM: $U_{DEP} = MALBACK + FEMALE + 1|CROSS + \epsilon$ with U_{DEP} the unhatched developed embryos proportion for each cross involving males from SI (*wPipI*-Tunis) and Tunis lines ($MALBACK$) and the five FEM lines ($FEMALE$) as fixed effects, with $CROSS$ as a random effect, and with ϵ the error parameter, following a binomial distribution.

For several variables (*Wolbachia* density in testes, *cidA* and *cidB* expressions and copy number) obtained with q-PCR, variability between the four MAL lines was analyzed using GLMs in the form $VAR = MALE + \epsilon$, with VAR one of the estimated variable of the MAL line ($MALE$) and ϵ the error parameter, following a Gaussian distribution.

Spearman correlation tests [56] were used to test for correlation between these variables (*Wolbachia* density in testes, *cidA* and *cidB* expressions and copy number) and the proportion of unhatched developed embryos for each MAL line. We did the same for the relation between the number of different *cidA* variants and the proportion of unhatched developed embryos for each MAL line. Finally, Wilcoxon test [57] was used to compare mean proportions of unhatched developed embryos between the two MAL lines harboring only two different *cidB* variants and the two MAL lines harboring four different *cidB* variants.

All computations were performed using the R version 3.4.4 [58]. Computed models were simplified by testing the significance of the different terms using likelihood ratio tests (LRT) and starting from the higher-order terms, as described in Crawley [59]. Factor levels of qualitative variables that were not different in their estimates (using LRTs) were grouped as described by Crawley [59]. The normality of the residuals was tested using Shapiro test for models with Gaussian error [60]. For models with Binomial error, overdispersion was calculated using the “dispersion_glm” function from the package blmeo for GLMM, and by dividing the residual deviance by the residuals degree of freedom of the model for GLM [61]; when detected, overdispersion was taken into account in the LRTs [62,63].

Supporting information

S1 Table. Presentation of the ten *C. pipiens* lines used in this study.
(DOCX)

S2 Table. The different crosses from which the CI cellular phenotype in *C. pipiens* was studied. Three different types of crosses were performed to study the cellular phenotype responsible for embryonic death in sterile crosses: i) sterile crosses between males and females

infected with different *Wolbachia* strains, ii) sterile crosses between infected males and uninfected females, and iii) fertile crosses between males and females from the same mosquito line infected or not by *Wolbachia*. Crosses from which confocal and optical microscopy pictures were taken are indicated next to the cross (Figs 1–3).

(DOCX)

S3 Table. Proportion of unhatched developed embryos from CI crosses involving either infected or uninfected (TC) females. Proportion of unhatched developed embryos are given as the mean proportion measured on fifty eggs for 10 rafts per cross (500 eggs observed per cross) \pm standard deviation. No unhatched developed embryos were found in any of the seven different crosses performed between infected males and uninfected females.

(DOCX)

S4 Table. *mod* and *resc* profiles of the lines used in crosses experiments. *mod* profiles were determined by crossing males from the four MAL lines with the females of the 4 ref-cytotypes lines Atyame et al. (2014). *resc* profiles were determined by crossing females of the five FEM lines with males of the 4 ref-cytotypes lines. MAL lines harbored *wPip* strains with different *mod* profiles while *wPip* strains from the five FEM lines presented the same *resc* profile.

(DOCX)

S5 Table. Accession numbers. Accession numbers for *cidA* *cidB* variants analyzed in S2 and S3 Figs.

(DOCX)

S6 Table. Explanation of the nomenclature used in this paper.

(DOCX)

S1 Fig. Unhatched developed embryos are haploid. Restriction profile of *kdr* PCR products by *RsaI* enzyme from single mosquito extracted DNA. M: molecular weight marker. 1 *wPip*-pIII-Slab line; 2/3: *wPip*IV-Ichkeul 13 line; 4/5: eggs-raft containing non-viable developed embryos from a CI cross between ♂ Slab x ♀ Ichkeul 13 (embryos display only maternal markers); 6: eggs-raft containing viable embryos from the fertile cross between ♂ Ichkeul 13 x ♀ Slab.

(TIF)

S2 Fig. Repertoires of CidA protein variants in the four MAL line *wPip* strains. Protein sequences alignment of the CidA variants found in the four *Wolbachia* strains *wPip*I-Tunis, *wPip*I-Utique, *wPip*II-Lavar and *wPip*III-Slab (MAL lines). The first sequence is used as a reference to determine the polymorphic region. For more clarity, only polymorphic positions are represented, thus amino-acid positions are not continuous. When more than two contiguous amino-acids were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic *wPip* group (I to III). No *cidA* or *cidB* nucleotide sequence variant was shared between the three *wPip* groups. However, the *wPip*II-Lavar CidA_{II}(α /1) variant and the *wPip*III-Slab CidA_{III}(β /8) variant presented the same amino-acid sequence. Based on their nucleotide sequences *wPip*III-Slab exhibited ten variants of *cidA*, *wPip*II-Lavar three, *wPip*I-Tunis four and *wPip*I Utique seven. However, *wPip*III-Slab exhibited only seven variants that differ in their amino-acid sequences since *cidA*_{III}(χ /6) and *cidA*_{III}(χ /7), *cidA*_{III}(δ /6) and *cidA*_{III}(δ /7), *cidA*_{III}(δ /8) and *cidA*_{III}(δ /9) have respectively identical amino-acid sequences (*i.e.* nucleotide polymorphic positions between them are synonymous).

(TIF)

S3 Fig. The repertoire of CidB protein variants in the four MAL line wPip strains. Protein sequences alignment of the CidB variants found in the four *Wolbachia* strains wPipI-Tunis, wPipI-Utique, wPipII-Lavar and wPipIII-Slab (MAL lines). The first sequence is used as a reference to determine the polymorphic region. For more clarity, only polymorphic positions are represented, thus amino-acid positions are not continuous. When more than two contiguous amino-acids were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic wPip group (I to III). However, no variant (*i.e.* complete CidB sequence) is common to wPip strains from different groups.

(TIF)

S4 Fig. Correlation between *cidB* variants diversity in wPip genomes and CI cellular intensity. Lighter gray bar plot accounts for the 10 crosses involving males from Lavar and Utique lines hosting wPip, which harbor 2 different variants of *cidB* in their genomes, while darker gray bar plot accounts for the 10 crosses involving males from Tunis and Slab lines, both infected with wPip strains harboring 4 different variants of *cidB*. Error bars represent the standard error. The proportion of unhatched developed embryos was significantly higher for males hosting four-variants wPip strains than for males hosting two-variants wPip strains (Wilcoxon, $W = 1159$, $p < 0.001$).

(TIF)

S5 Fig. *cidA* copy number in the wPip strains infecting the four MAL lines. *cidA* copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidA* gene and the *Wolbachia wsp* gene. The colored dots represent the *cidA* copy number per wPip genome in a male and the red strips represent the average *cidA* copy number per wPip genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different).

(TIF)

S6 Fig. *cidB* copy number in the wPip strains infecting the four MAL lines. *cidB* copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidB* gene and the *Wolbachia wsp* gene. The colored dots represent the *cidB* copy number per wPip genome in a male and the red strips represent the average *cidB* copy number per wPip genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different).

(TIF)

S7 Fig. *cidA/cidB* copy number in the wPip strains infecting the four MAL lines. *cidA/cidB* copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidA* gene and the *Wolbachia cidB* gene. The colored dots represent the *cidA/cidB* copy number per wPip genome in a male and the red strips represent the average *cidA/cidB* copy number per wPip genome for ten males per MAL lines. *cidA/cidB* copy number were not significantly different between the four wPip strains infecting the four MAL lines.

(TIFF)

S8 Fig. *cidA* expression level in the wPip strains infecting the four MAL lines. *cidA* expression was measured by quantitative PCR as the ratio between the *Wolbachia cidA* gene expression and the *Wolbachia wsp* gene expression. The colored dots represent the *cidA* expression level per wPip genome in a male and the red strips represent the average *cidA* expression level per wPip genome for ten males per MAL lines. Expression levels of *cidA* genes were not

significantly different between the four *wPip* strains infecting the four MAL lines. (TIFF)

S9 Fig. *cidB* expression level in the *wPip* strains infecting the four MAL lines. *cidB* expression was measured by quantitative PCR as the ratio between the *Wolbachia cidB* gene expression and the *Wolbachia wsp* gene expression. The colored dots represent the *cidB* expression level per *wPip* genome in a male and the red strips represent the average *cidB* expression level per *wPip* genome for ten males per MAL lines. Expression levels of *cidB* genes were not significantly different between the four *wPip* strains infecting the four MAL lines. (TIFF)

S10 Fig. *cidA/cidB* expression level in the *wPip* strains infecting the four MAL lines. *cidA/cidB* expression levels was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidA* gene and the *Wolbachia cidB* gene. The colored dots represent the *cidA/cidB* expression level per *wPip* genome in a male and the red strips represent the average *cidA/cidB* expression per *wPip* genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different). (TIFF)

S1 Dataset. Proportion of unhatched developed haploid embryos of performed CI crosses. (CSV)

S2 Dataset. *Wolbachia testes* density of the four MAL lines estimated with Real Time Quantitative PCR. (CSV)

S3 Dataset. *Wolbachia cidA* and *cidB* copy number of the four MAL lines estimated with Real Time Quantitative PCR. (CSV)

S4 Dataset. *Wolbachia cidA* and *cidB* expression of the four MAL lines estimated with Real Time Quantitative PCR. (CSV)

Acknowledgments

We thank Sandra Unal, Adeline Toro and Marco Perriat-Sanguinet for technical support in the crossing experiments, Dr Rudy Caparros for technical support in cloning and Sanger sequencing and Mine Altinli for providing us with some of the TC-treated lines. We also thank Dr Philippe Clair for his help in the Real Time Quantitative PCR experiments which were performed through the technical facility of the “qPCR Haut Debit (qPHD) Montpellier génomiX” platform. Sequencing data were produced through the technical facilities of the LabEX “Centre Méditerranéen de l’Environnement et de la Biodiversité” in the GENSEQ platform. Confocal microscopy imaging was performed in the MRI-CRBM platform. Optical observations of embryos 48 hours after oviposition were performed at the CytoEvol facilities of UMR ISEM—CBGP of the LabEx CeMEB.

Author Contributions

Conceptualization: Manon Bonneau, Frédéric Landmann, Mylène Weill, Mathieu Sicard.

Data curation: Manon Bonneau.

Formal analysis: Manon Bonneau, Pierrick Labbé.

Funding acquisition: Mathieu Sicard.

Investigation: Manon Bonneau, Fabienne Justy.

Methodology: Manon Bonneau, Frédéric Landmann, Pierrick Labbé, Fabienne Justy, Mylène Weill, Mathieu Sicard.

Project administration: Mathieu Sicard.

Supervision: Mylène Weill, Mathieu Sicard.

Validation: Frédéric Landmann.

Visualization: Manon Bonneau, Frédéric Landmann.

Writing – original draft: Manon Bonneau, Frédéric Landmann, Pierrick Labbé, Mylène Weill, Mathieu Sicard.

Writing – review & editing: Manon Bonneau, Mylène Weill, Mathieu Sicard.

References

1. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol.* 2008; 6: 741–751. <https://doi.org/10.1038/nrmicro1969> PMID: 18794912
2. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. How many species are infected with *Wolbachia*? a statistical analysis of current data. *FEMS Microbiol Lett.* 2008; 281: 215–220. <https://doi.org/10.1111/j.1574-6968.2008.01110.x> PMID: 18312577
3. Zug R, Hammerstein P. Still a Host of Hosts for *Wolbachia*: Analysis of Recent Data Suggests That 40% of Terrestrial Arthropod Species Are Infected. *PLoS One.* 2012; 7: e38544. <https://doi.org/10.1371/journal.pone.0038544> PMID: 22685581
4. Yen JH, Barr AR. The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *J Invertebr Pathol.* 1973; 22: 242–250. [https://doi.org/10.1016/0022-2011\(73\)90141-9](https://doi.org/10.1016/0022-2011(73)90141-9) PMID: 4206296
5. Callaini G, Riparbelli MG, Giordano R, Dallai R. Mitotic Defects Associated with Cytoplasmic Incompatibility in *Drosophila simulans*. *J Invertebr Pathol.* 1996; 67: 55–64. <https://doi.org/10.1006/jipa.1996.0009>
6. Engelstädter J, Telschow A. Cytoplasmic incompatibility and host population structure. *Heredity.* 2009; 103: 196–207. <https://doi.org/10.1038/hdy.2009.53> PMID: 19436325
7. Laven H. Speciation and evolution in *Culex pipiens*. In: Wright J, Pal R, editors. *Genetics of Insect Vectors of Disease.* Elsevier: Amsterdam; 1967. pp. 251–275.
8. O'Neill SL, Karr TL. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. *Nature.* 1990; 348: 178–180. <https://doi.org/10.1038/348178a0> PMID: 2234083
9. Breeuwer JAJ, Werren JH. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature.* 1990; 346: 558–560. <https://doi.org/10.1038/346558a0> PMID: 2377229
10. Bordenstein SR, O'Hara FP, Werren JH. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature.* 2001; 409: 707–710. <https://doi.org/10.1038/35055543> PMID: 11217858
11. Duron OI, Bernard C, Unal S, Berthomieu A, Berticat C, Weill M. Tracking factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. *Mol Ecol.* 2006; 15: 3061–3071. <https://doi.org/10.1111/j.1365-294X.2006.02996.x> PMID: 16911221
12. Atyame CM, Delsuc F, Pasteur N, Weill M, Duron O. Diversification of *Wolbachia* Endosymbiont in the *Culex pipiens* Mosquito. *Mol Biol Evol.* 2011; 28: 2761–2772. <https://doi.org/10.1093/molbev/msr083> PMID: 21515811
13. Merçot H, Charlat S. *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. *Genetica.* 2004; 120: 51–59. https://doi.org/10.1007/978-94-007-0965-2_4 PMID: 15088646
14. Duron O, Fort P, Weill M. Influence of aging on cytoplasmic incompatibility, sperm modification and *Wolbachia* density in *Culex pipiens* mosquitoes. *Heredity.* 2007; 98: 368–374. <https://doi.org/10.1038/sj.hdy.6800948> PMID: 17519957

15. Bordenstein SR, Werren JH. Bidirectional incompatibility among divergent *Wolbachia* and incompatibility level differences among closely related *Wolbachia* in *Nasonia*. *Heredity*. 2007; 99: 278–287. <https://doi.org/10.1038/sj.hdy.6800994> PMID: [17519968](https://pubmed.ncbi.nlm.nih.gov/17519968/)
16. Bourtzis K, Nirgianaki A, Markakis G, Savakis C. *Wolbachia* infection and cytoplasmic incompatibility in *Drosophila* species. *Genetics*. 1996; 144: 1063–73. PMID: [8913750](https://pubmed.ncbi.nlm.nih.gov/8913750/)
17. Breeuwer JA, Werren JH. Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. *Genetics*. 1993; 135: 565–74. PMID: [8244014](https://pubmed.ncbi.nlm.nih.gov/8244014/)
18. Bressac C, Rousset F. The Reproductive Incompatibility System in *Drosophila simulans*: Dapi-Staining Analysis of the *Wolbachia* Symbionts in Sperm Cysts. *J Invertebr Pathol*. 1993; 61: 226–230. <https://doi.org/10.1006/jipa.1993.1044> PMID: [7689622](https://pubmed.ncbi.nlm.nih.gov/7689622/)
19. Bordenstein SR, Uy JJ, Werren JH. Host genotype determines cytoplasmic incompatibility type in the haplodiploid genus *Nasonia*. *Genetics*. 2003; 164: 223–33. PMID: [12750334](https://pubmed.ncbi.nlm.nih.gov/12750334/)
20. Lassy CW, Karr TL. Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*. *Mech Dev*. 1996; 57: 47–58. [https://doi.org/10.1016/0925-4773\(96\)00527-8](https://doi.org/10.1016/0925-4773(96)00527-8) PMID: [8817452](https://pubmed.ncbi.nlm.nih.gov/8817452/)
21. Ryan SL, Saul GB. Post-fertilization effect of incompatibility factors in *Mormoniella*. *MGG Mol Gen Genet*. 1968; 103: 29–36. <https://doi.org/10.1007/BF00271154> PMID: [5753233](https://pubmed.ncbi.nlm.nih.gov/5753233/)
22. Landmann F, Orsi GA, Loppin B, Sullivan W. *Wolbachia*-Mediated Cytoplasmic Incompatibility Is Associated with Impaired Histone Deposition in the Male Pronucleus. *PLoS Pathog*. 2009; 5: e1000343. <https://doi.org/10.1371/journal.ppat.1000343> PMID: [19300496](https://pubmed.ncbi.nlm.nih.gov/19300496/)
23. Tram U, Sullivan W. Role of Delayed Nuclear Envelope Breakdown and Mitosis in *Wolbachia*-Induced Cytoplasmic Incompatibility. *Science*. 2002; 296: 1124–1126. <https://doi.org/10.1126/science.1070536> PMID: [12004132](https://pubmed.ncbi.nlm.nih.gov/12004132/)
24. Reed KM, Werren JH. Induction of paternal genome loss by the paternal-sex-ratio chromosome and cytoplasmic incompatibility bacteria (*Wolbachia*): A comparative study of early embryonic events. *Mol Reprod Dev*. 1995; 40: 408–418. <https://doi.org/10.1002/mrd.1080400404> PMID: [7598906](https://pubmed.ncbi.nlm.nih.gov/7598906/)
25. Callaini G, Dallai R, Riparbelli MG. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. *J Cell Sci*. 1997; 110: 271–80. PMID: [9044057](https://pubmed.ncbi.nlm.nih.gov/9044057/)
26. Jost E. Untersuchungen zur Inkompatibilität im *Culex pipiens*-Komplex. *Wilhelm Roux Arch Entwickl Mech Org*. 1970; 166: 173–188. <https://doi.org/10.1007/BF00576992> PMID: [28304791](https://pubmed.ncbi.nlm.nih.gov/28304791/)
27. Tram U, Fredrick K, Werren JH, Sullivan W. Paternal chromosome segregation during the first mitotic division determines *Wolbachia*-induced cytoplasmic incompatibility phenotype. *J Cell Sci*. 2006; 119: 3655–3663. <https://doi.org/10.1242/jcs.03095> PMID: [16912076](https://pubmed.ncbi.nlm.nih.gov/16912076/)
28. Vavre F, Fleury F, Varaldi J, Fouillet P, Bouleatreau M. Evidence for female mortality in *Wolbachia*-mediated cytoplasmic incompatibility in haplodiploid insects: epidemiologic and evolutionary consequences. *Evolution*. 2000; 54: 191–200. <https://doi.org/10.1111/j.0014-3820.2000.tb00019.x> PMID: [10937195](https://pubmed.ncbi.nlm.nih.gov/10937195/)
29. Vavre F, Dedeine F, Quillon M, Fouillet P, Fleury F, Boulétreau M. Within-species diversity of *Wolbachia*-induced cytoplasmic incompatibility in haplodiploid insects. *Evolution*. 2001; 55: 1710–1714. <https://doi.org/10.1111/j.0014-3820.2001.tb00691.x> PMID: [11580031](https://pubmed.ncbi.nlm.nih.gov/11580031/)
30. Tram U, Ferree PM, Sullivan W. Identification of *Wolbachia*-host interacting factors through cytological analysis. *Microbes Infect*. 2003; 5: 999–1011. [https://doi.org/10.1016/S1286-4579\(03\)00192-8](https://doi.org/10.1016/S1286-4579(03)00192-8) PMID: [12941392](https://pubmed.ncbi.nlm.nih.gov/12941392/)
31. Breeuwer JAJ. *Wolbachia* and cytoplasmic incompatibility in the spider mites *Tetranychus urticae* and *T. turkestanii*. *Heredity*. 1997; 79: 41–47. <https://doi.org/10.1038/sj.hdy.6881810>
32. Werren JH. Biology of *Wolbachia*. *Annu Rev Entomol*. 1997; 42: 587–609. <https://doi.org/10.1146/annurev.ento.42.1.587> PMID: [15012323](https://pubmed.ncbi.nlm.nih.gov/15012323/)
33. Hurst LD. The evolution of cytoplasmic incompatibility or when spite can be successful. *J Theor Biol*. 1991; 148: 269–277. [https://doi.org/10.1016/S0022-5193\(05\)80344-3](https://doi.org/10.1016/S0022-5193(05)80344-3) PMID: [2016892](https://pubmed.ncbi.nlm.nih.gov/2016892/)
34. Beckmann JF, Ronau JA, Hochstrasser M. A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat Microbiol*. 2017; 2: 17007. <https://doi.org/10.1038/nmicrobiol.2017.7> PMID: [28248294](https://pubmed.ncbi.nlm.nih.gov/28248294/)
35. LePage DP, Metcalf JA, Bordenstein SR, On J, Perlmutter JI, Shropshire JD, et al. Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature*. 2017; 543: 243–247. <https://doi.org/10.1038/nature21391> PMID: [28241146](https://pubmed.ncbi.nlm.nih.gov/28241146/)
36. Bonneau M, Atyame C, Beji M, Justy F, Cohen-Gonsaud M, Sicard M, et al. *Culex pipiens* crossing type diversity is governed by an amplified and polymorphic operon of *Wolbachia*. *Nat Commun*. 2018; 9. <https://doi.org/10.1038/s41467-017-02749-w> PMID: [29358578](https://pubmed.ncbi.nlm.nih.gov/29358578/)

37. Beckmann JF, Fallon AM. Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: Implications for cytoplasmic incompatibility. *Insect Biochem Mol Biol.* 2013; 43: 867–878. <https://doi.org/10.1016/j.ibmb.2013.07.002> PMID: 23856508
38. Lindsey ARI, Rice DW, Bordenstein SR, Brooks AW, Bordenstein SR, Newton ILG. Evolutionary Genetics of Cytoplasmic Incompatibility Genes *cifA* and *cifB* in Prophage WO of *Wolbachia*. *Genome Biol Evol.* 2018; 10: 434–451. <https://doi.org/10.1093/gbe/evy012> PMID: 29351633
39. Klasson L, Walker T, Sebahia M, Sanders MJ, Quail, Lord A, et al. Genome evolution of *Wolbachia* strain *wPip* from the *Culex pipiens* group. *Mol Biol Evol.* 2008; 25: 1877–1887. <https://doi.org/10.1093/molbev/msn133> PMID: 18550617
40. Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc Natl Acad Sci.* 2018; 115: 4987–4991. <https://doi.org/10.1073/pnas.1800650115> PMID: 29686091
41. Baldo L, Hotopp JCD, Jolley K a., Bordenstein SR, Biber S a., Choudhury RR, et al. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol.* 2006; 72: 7098–7110. <https://doi.org/10.1128/AEM.00731-06> PMID: 16936055
42. Atyame CM, Labbé P, Dumas E, Milesi P, Charlat S, Fort P, et al. *Wolbachia* Divergence and the Evolution of Cytoplasmic Incompatibility in *Culex pipiens*. *PLoS One.* 2014; 9: e87336. <https://doi.org/10.1371/journal.pone.0087336> PMID: 24498078
43. Dumas E, Atyame CM, Milesi P, Fonseca DM, Shaikovich E V, Unal S, et al. Population structure of *Wolbachia* and cytoplasmic introgression in a complex of mosquito species. *BMC Evol Biol.* 2013; 13: 181. <https://doi.org/10.1186/1471-2148-13-181> PMID: 24006922
44. Duron O, Weill M. *Wolbachia* infection influences the development of *Culex pipiens* embryo in incompatible crosses. *Heredity.* 2006; 96: 493–500. <https://doi.org/10.1038/sj.hdy.6800831> PMID: 16639421
45. Atanassov BS, Koutelou E, Dent SY. The role of deubiquitinating enzymes in chromatin regulation. *FEBS Lett.* 2011; 585: 2016–2023. <https://doi.org/10.1016/j.febslet.2010.10.042> PMID: 20974139
46. Gebiola M, Giorgini M, Kelly SE, Doremus MR, Ferree PM, Hunter MS. Cytological analysis of cytoplasmic incompatibility induced by *Cardinium* suggests convergent evolution with its distant cousin *Wolbachia*. *Proc R Soc B Biol Sci.* 2017; 284: 20171433. <https://doi.org/10.1098/rspb.2017.1433> PMID: 28878066
47. Penz T, Schmitz-Esser S, Kelly SE, Cass BN, Müller A, Woyke T, et al. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in *Cardinium hertigii*. Moran NA, editor. *PLoS Genet.* 2012; 8: e1003012. <https://doi.org/10.1371/journal.pgen.1003012> PMID: 23133394
48. Breeuwer JAJ, Werren JH. Effect of genotype on cytoplasmic incompatibility between two species of *Nasonia*. *Heredity.* 1993; 70: 428–436. <https://doi.org/10.1038/hdy.1993.60>
49. Atyame CM, Duron O, Tortosa P, Pasteur N, Fort P, Weill M. Multiple *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in *Culex pipiens* mosquito populations. *Mol Ecol.* 2011; 20: 286–298. <https://doi.org/10.1111/j.1365-294X.2010.04937.x> PMID: 21114563
50. Duron O, Bernard J, Atyame CM, Dumas E, Weill M. Rapid evolution of *Wolbachia* incompatibility types. *Proc R Soc B Biol Sci.* 2012; 279: 4473–4480. <https://doi.org/10.1098/rspb.2012.1368> PMID: 22951738
51. Aakre CD, Herrou J, Phung TN, Perchuk BS, Crosson S, Laub MT. Evolving New Protein-Protein Interaction Specificity through Promiscuous Intermediates. *Cell.* 2015; 163: 594–606. <https://doi.org/10.1016/j.cell.2015.09.055> PMID: 26478181
52. Berticat C, Rousset F, Raymond M, Berthomieu A, Weill M. High *Wolbachia* density in insecticide-resistant mosquitoes. *Proc R Soc B Biol Sci.* 2002; 269: 1413–1416. <https://doi.org/10.1098/rspb.2002.2022> PMID: 12079666
53. Clements AN. *The Biology of Mosquitoes: Development, Nutrition and Reproduction.* Wallingford: CAB International; 1992.
54. Weill M, Berticat C, Raymond M, Chevillon C. Quantitative Polymerase Chain Reaction to Estimate the Number of Amplified Esterase Genes in Insecticide-Resistant Mosquitoes. *Anal Biochem.* 2000; 285: 267–270. <https://doi.org/10.1006/abio.2000.4781> PMID: 11017713
55. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol Evol.* 2010; 27: 221–224. <https://doi.org/10.1093/molbev/msp259> PMID: 19854763
56. Best D, Roberts D. Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. *J R Stat Soc.* 1975; 24: 377–379.
57. Bauer DF. Constructing confidence sets using rank statistics. *J Am Stat Assoc.* 1972; 67: 687–690.
58. R Core Team. *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing; 2018.

59. Crawley M. The R Book. John Wiley, editor. Chichester UK: John Wiley & Sons, Ltd; 2007.
60. Royston P. An extension of Shapiro and Wilk's W test for normality to large samples. *Appl Stat.* 1982; 31: 115–124.
61. Korner-Nievergelt F, Roth T, Felten Sv, Guelat J AB and K-NP. Bayesian Data Analysis in Ecology using Linear Models with R, BUGS and Stan. New York: Academic Press 2015.
62. Lebreton J-D, Burnham KP, Clobert J, Anderson DR. Modeling Survival and Testing Biological Hypotheses Using Marked Animals: A Unified Approach with Case Studies. *Ecol Monogr.* 1992; 62: 67–118. <https://doi.org/10.2307/2937171>
63. Anderson DR, Burnham KP, White GC. AIC Model Selection in Overdispersed Capture-Recapture Data. *Ecology.* 1994; 75: 1780–1793. <https://doi.org/10.2307/1939637>

S1 Table. Presentation of the ten *C. pipiens* lines used in this study.

Line	<i>Wolbachia</i> group	<i>Culex</i> subspecies	Country	Year of collection	Reference
Tunis	wPipI	<i>pipiens</i>	Tunisia	1995	Ben Cheikh <i>et al.</i> , 1998
Utique	wPipI	<i>pipiens</i>	Tunisia	2014	This study
Lavar	wPipII	<i>pipiens</i>	France	2003	Duron <i>et al.</i> , 2005
Slab	wPipIII	<i>quinquefasciatus</i>	USA	1954	Georghiou <i>et al.</i> , 1966
Maclo	wPipIII	<i>quinquefasciatus</i>	USA	1984	Duron <i>et al.</i> , 2006
Istanbul	wPipIV	<i>pipiens</i>	Turkey	2003	Duron <i>et al.</i> , 2005
Harash	wPipIV	<i>pipiens</i>	Algeria	2006	Alout <i>et al.</i> , 2009
Ichkeul 09	wPipIV	<i>pipiens</i>	Tunisia	2010-2011	Bonneau <i>et al.</i> , 2018
Ichkeul 13	wPipIV	<i>pipiens</i>	Tunisia	2010-2011	Bonneau <i>et al.</i> , 2018
Ichkeul 21	wPipIV	<i>pipiens</i>	Tunisia	2010-2011	Bonneau <i>et al.</i> , 2018

S2 Table. The different crosses from which the CI cellular phenotype in *C. pipiens* was studied.

Inter-lines CI crosses		Intra-line fertile crosses	
Between infected males and infected females	Between infected males and uninfected females	Between infected males and infected females	Between uninfected males and uninfected females
♂ Utique x ♀ Harash (Fig 2F)	♂ Utique x ♀ IstanbulTC (Fig 3B & 3C & 3D)	♂ Tunis x ♀ Tunis (Fig1C)	♂ SlabTC x ♀ SlabTC (Fig 1A & 1D)
♂ Utique x ♀ Istanbul (Fig 2G1 & 2G2)	♂ Lavar x ♀ SlabTC (Fig 3A & 3E)	♂ Harash x ♀ Harash (Fig 1G)	♂ Istanbul TC x ♀ Istanbul TC
♂ Utique x ♀ Ichkeul 09 (Fig 2C)		♂ Istanbul x ♀ Istanbul (Fig 1F)	
♂ Tunis x ♀ Ichkeul 21		♂ Slab x ♀ Slab (Fig 1B & 1 E)	
♂ Lavar x ♀ Slab (Fig 2D & 2E)		♂ Maclo x ♀ Maclo	
♂ Ichkeul 21 x ♀ Utique (Fig 2A & 2B)			
♂ Ichkeul 09 x ♀ Tunis			
♂ Ichkeul 09 x ♀ Utique			

Three different types of crosses were performed to study the cellular phenotype responsible for embryonic death in sterile crosses i) sterile crosses between males and females infected with different *Wolbachia* strains, ii) sterile crosses between infected males and uninfected females, and iii) fertile crosses between males and females from the same mosquito line infected or not by *Wolbachia*. Crosses from which confocal and optical microscopy picture were taken are written next to the cross (Figs 1-3).

S3 Table. Proportions of unhatched developed embryos from CI crosses involving either infected or uninfected (TC) females.

Crosses		unhatched developed embryos rate
♂ Tunis	♀ Harash	0.850 ± 0.153
♂ Tunis	♀ Ichkeul 09	0.740 ± 0.139
♂ Tunis	♀ Ichkeul 13	0.530 ± 0.180
♂ Tunis	♀ Ichkeul 13TC	0.000 ± 0.000
♂ Tunis	♀ Ichkeul 21	0.686 ± 0.207
♂ Tunis	♀ Ichkeul 21TC	0.000 ± 0.000
♂ Tunis	♀ Istanbul	0.784 ± 0.148
♂ Tunis	♀ IstanbulTC	0.000 ± 0.000
♂ Tunis	♀ SlabTC	0.000 ± 0.000
♂ SI(wPipI-Tunis)	♀ Harash	0.816 ± 0.056
♂ SI(wPipI-Tunis)	♀ Ichkeul 09	0.776 ± 0.107
♂ SI(wPipI-Tunis)	♀ Ichkeul 13	0.536 ± 0.338
♂ SI(wPipI-Tunis)	♀ Ichkeul 21	0.600 ± 0.156
♂ SI(wPipI-Tunis)	♀ Istanbul	0.842 ± 0.122
♂ Utique	♀ Harash	0.440 ± 0.374
♂ Utique	♀ Ichkeul 09	0.466 ± 0.265
♂ Utique	♀ Ichkeul 13	0.422 ± 0.392
♂ Utique	♀ Ichkeul 21	0.436 ± 0.264
♂ Utique	♀ Istanbul	0.330 ± 0.250
♂ Utique	♀ SlabTC	0.000 ± 0.000
♂ Lavar	♀ Harash	0.136 ± 0.093
♂ Lavar	♀ Ichkeul 09	0.242 ± 0.118
♂ Lavar	♀ Ichkeul 13	0.128 ± 0.138
♂ Lavar	♀ Ichkeul 21	0.268 ± 0.111
♂ Lavar	♀ Istanbul	0.116 ± 0.082
♂ Lavar	♀ SlabTC	0.000 ± 0.000
♂ Slab	♀ Harash	0.722 ± 0.148
♂ Slab	♀ Ichkeul 09	0.664 ± 0.178
♂ Slab	♀ Ichkeul 13	0.662 ± 0.116
♂ Slab	♀ Ichkeul 21	0.834 ± 0.069
♂ Slab	♀ Istanbul	0.754 ± 0.201
♂ Slab	♀ SlabTC	0.000 ± 0.000

Proportions of unhatched developed embryos are given as the mean proportion measured on fifty eggs for 10 rafts per cross (500 eggs observed per cross) ± standard deviation. No unhatched developed embryos were found in any of the seven different crosses performed between infected males and uninfected females.

S4 Table. *mod* and *resc* profiles of the lines used in crosses experiments.

4-ref cytotypes ♀						
MAL Lines	Lavar	Maclo	Slab	Istanbul	<i>mod</i> type	Reference
wPipI-Tunis	IC	C	IC	IC	ii	Atyame <i>et al.</i> , 2014
wPipI-Utique	C	IC	C	IC	iv	This study
wPipII-Lavar	C	C	IC	IC	vi	Atyame <i>et al.</i> , 2014
wPipIII-Slab	IC	C	C	IC	iii	Atyame <i>et al.</i> , 2014
4-ref cytotypes ♂						
FEM Lines	Lavar	Maclo	Slab	Istanbul	<i>resc</i> type	Reference
wPipIV-Harash	IC	C	IC	C	3	Atyame <i>et al.</i> , 2014
wPipIV-Istanbul	IC	C	IC	C	3	Atyame <i>et al.</i> , 2014
wPipIV-Ichkeul 13	IC	C	IC	C	3	This study
wPipIV-Ichkeul 09	IC	C	IC	C	3	This study
wPipIV-Ichkeul 21	IC	C	IC	C	3	This study

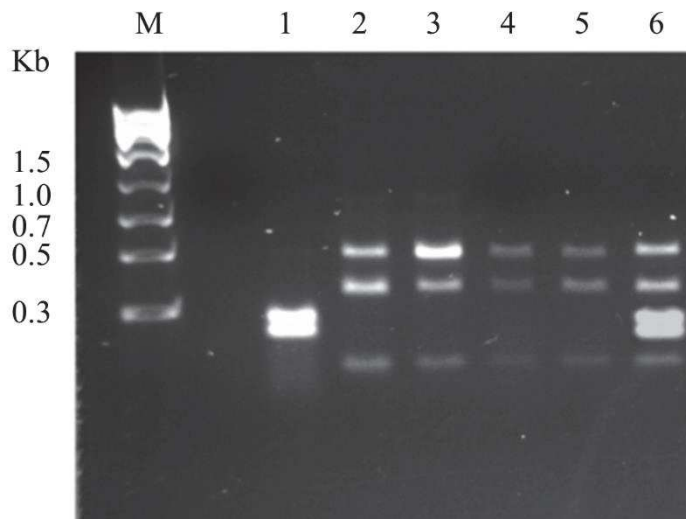
mod profiles were determined by crossing males from the four MAL lines with the females of the 4 ref-cytotypes lines (Atyame *et al.*, 2014). *resc* profiles were determined by crossing females of the five FEM lines with males of the 4 ref-cytotypes lines. MAL lines harbored wPip strains with different *mod* profiles while wPip strains from the five FEM lines presented the same *resc* profile.

S5 Table. Accession numbers.Accession numbers for *cidA cidB* variants analyzed in S2 and S3 Figs

Gene name	Variant sequence submission name	Accession number	Gene name	Variant sequence submission name	Accession number
<i>cidA</i>	cidA_I(alpha/1)	MF444963	<i>cidB</i>	cidB_I(a/1)	MF444982
	cidA_I(alpha/2)	MH544806		cidB_I(a/2)	MF444983
	cidA_I(beta/2)	MF444966		cidB_I(b/1)	MF444984
	cidA_I(gamma/1)	MF444964		cidB_I(b/2)	MF444985
	cidA_I(gamma/2)	MF444965		cidB_II(a/2)	MF444986
	cidA_I(zeta/1)	MH544807		cidB_II(a/1)	MF444987
	cidA_I(zeta/2)	MH544808		cidB_III(d/1)	MH544820
	cidA_I(zeta/5)	MH544809		cidB_III(d/2)	MH544821
	cidA_II(alpha/1)	MF444967		cidB_III(e/1)	MH544822
	cidA_II(alpha/2)	MF444968		cidB_III(e/2)	MH544823
	cidA_II(beta/2)	MF444969			
	cidA_III(beta/8)	MH544810			
	cidA_III(gamma/4)	MH544811			
	cidA_III(gamma/5)	MH544812			
	cidA_III(gamma/6)	MH544813			
	cidA_III(gamma/7)	MH544814			
	cidA_III(gamma/8)	MH544815			
	cidA_III(delta/6)	MH544816			
	cidA_III(delta/7)	MH544817			
	cidA_III(delta/8)	MH544818			
cidA_III(delta/9)	MH544819				

S6 Table. Explanation of the nomenclature used in this paper.

DUB	Deubiquitylating enzyme: enzyme that remove ubiquitin from ubiquitinated protein.
CI penetrance	Percentage of embryos that will die from CI
CI cellular phenotypes	Cellular consequences of CI during early embryogenesis
CI intensity at the cellular level	Severity of the defect suffered by the paternal chromatin resulting in its partial or total exclusion during first embryonic division
<i>cidA</i>	<i>cidA</i> gene is involved in CI induction in <i>Drosophila melanogaster</i> transinfected males [35,37] and in CI rescue in <i>D. melanogaster</i> females [40]. This gene is putatively involved in CI diversity in <i>Culex pipiens</i> [36]
CidA	The protein produce by <i>cidA</i> gene
<i>cidB</i>	<i>cidB</i> gene is involved in CI induction in <i>D. melanogaster</i> transinfected males and CI diversity in <i>C. pipiens</i> [35,37,36].
CidB	The protein produce by <i>cidB</i> gene
<i>cidA</i> ^{wPip} , <i>cidA</i> ^{wMel}	<i>cidA</i> homologues respectively found in the wPip and wMel genomes. Protein sequences are divergent (66% of identity).
<i>cidB</i> ^{wPip} , <i>cidB</i> ^{wMel}	<i>cidB</i> homologues respectively found in the wPip and wMel genomes. Protein sequences are divergent (76% of identity).
<i>cinA</i>	Paralogue of <i>cidA</i> gene likely involved in CI [34, 37] but not in <i>C. pipiens</i> CI diversity [36].
<i>cinB</i>	Paralogue of the <i>cidB</i> gene likely involved in CI [34, 37] but not in CI diversity in <i>C. pipiens</i> [36].
<i>Culex pipiens</i> crossing type	Compatibility and incompatibility relationships between males and females from different <i>C. pipiens</i> lines.
MLST	Multi Locus Sequence Typing: list of genes created by Baldo et al. 2006 that are used to classify the <i>Wolbachia</i> strains into the different super-groups.
MLST specific to wPip	Group of hypervariable genes that are used to classify the different wPip strains in the five wPip phylogenetic groups.
<i>Mod</i> profile	Compatibility or incompatibility profile when males from a given line are crossed with females from other lines.
<i>Resc</i> profile	Compatibility or incompatibility profile when females from a given line are crossed with males from other lines.
Toxin-antidote system	Theoretical model on CI mechanism proposed by Hurst (1991)
WO phage	Bacteriophage name WO after <i>Wolbachia</i> that is present as a prophage in almost all <i>Wolbachia</i> genomes.
wPip	<i>Wolbachia</i> strains infecting <i>Culex pipiens</i>
wPipI, wPipII, wPipIII, wPipIV and wPipV	The different phylogenetic wPip groups obtained from wPip specific MLST
wRi, wNo, wHa	<i>Wolbachia</i> strains infecting <i>Drosophila simulans</i>
wVitA	<i>Wolbachia</i> strain infecting <i>Nasonia vitripennis</i>



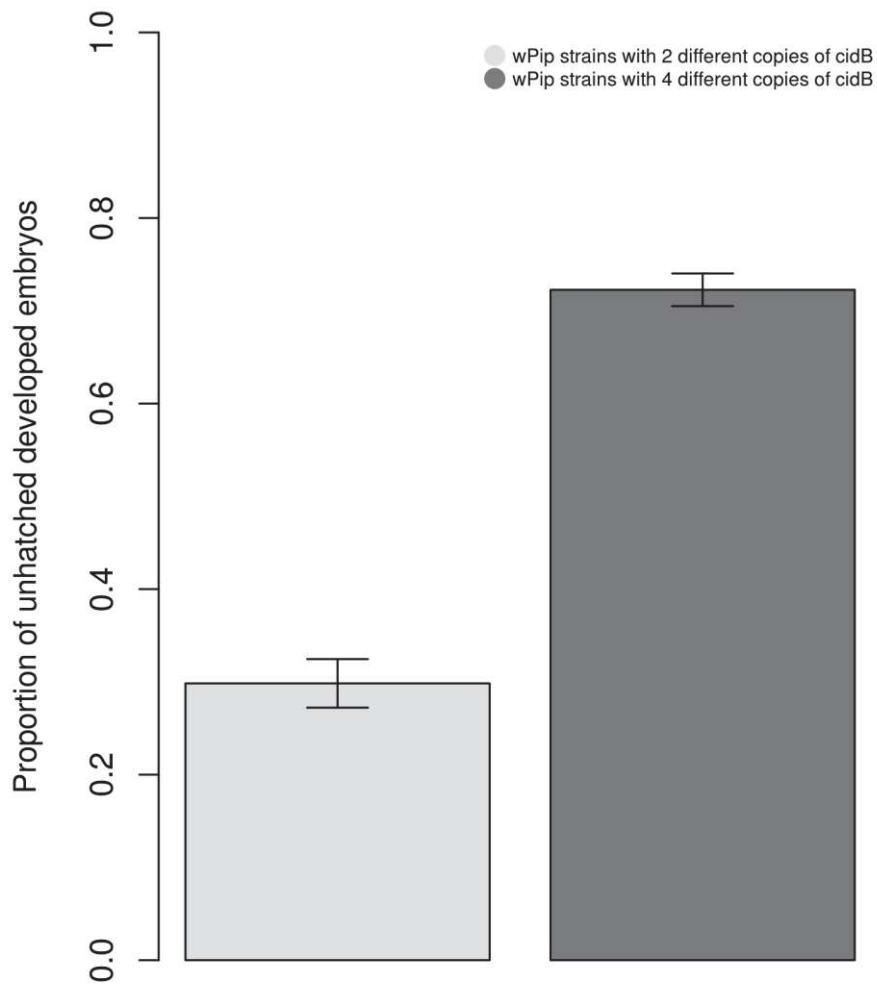
S1 Fig. Unhatched developed embryos are haploid.

Restriction profile of *kdr* PCR products by *RsaI* enzyme from single mosquito extracted DNA. M: molecular weight marker. 1 wPipIII-Slab line; 2/3: wPipIV-Ichkeul 13 line; 4/5: eggs-raft containing non-viable developed embryos from a CI cross between ♂ Slab x ♀ Ichkeul 13 (embryos display only maternal markers); 6: eggs-raft containing viable embryos from the fertile cross between ♂ Ichkeul 13 x ♀ Slab.

Variants Name	CidA positions (aa)																																															
	1	1	1	1	1	1	1	1	1	1	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4										
CidA_I (α1)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	F	V	G	H	E	T	LS	F	E	LYQ	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_I (γ1)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	V	G	H	E	T	LS	F	E	LYQ	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_I (γ2)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	V	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_I (β2)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	F	A	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_I (α1)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	F	V	G	H	E	T	LS	F	E	LYQ	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_I (α2)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	F	V	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_I (γ1)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	V	G	H	E	T	LS	F	E	LYQ	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_I (γ2)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	V	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_I (ζ1)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	A	G	H	E	T	LS	F	E	LYQ	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_I (ζ2)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	A	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_I (ζ5)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	F	A	G	H	E	T	LS	F	E	LYQ	R	*V	YTV	TS	F	R	T	L	NS	R	F	EK	KN	KR																
CidA_II (α1)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	EK																
CidA_II (α2)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	L	A	R	C	D	V	SN	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_II (β2)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	L	A	R	C	D	V	SN	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR																
CidA_III (β8)	A	K	G	SSDEL	SRGNPITDE	I	GG	R	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	EK																
CidA_III (γ4)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	L	A	G	H	E	T	LS	V	SN	F	E	LYQ	R	*V	YTV	TS	F	R	T	S	DC	K	L	KL	RD	KR														
CidA_III (γ5)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	L	A	G	H	E	T	LS	V	SN	I	Q	FSE	K	II	GRI	IK	L	K	T	S	DC	K	L	KL	RD	KR														
CidA_III (γ6)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_III (γ7)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_III (γ8)	R	R	K	PPGAF	GKTETEF GK	V	GG	R	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	EK																
CidA_III (δ6)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_III (δ7)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	KR																
CidA_III (δ8)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	EK																
CidA_III (δ9)	R	R	K	PPGAF	GKTETEF GK	V	DS	K	L	A	G	H	E	T	LS	I	Q	FSE	K	II	GRI	IK	L	K	P	L	NS	R	F	EK	KN	EK																

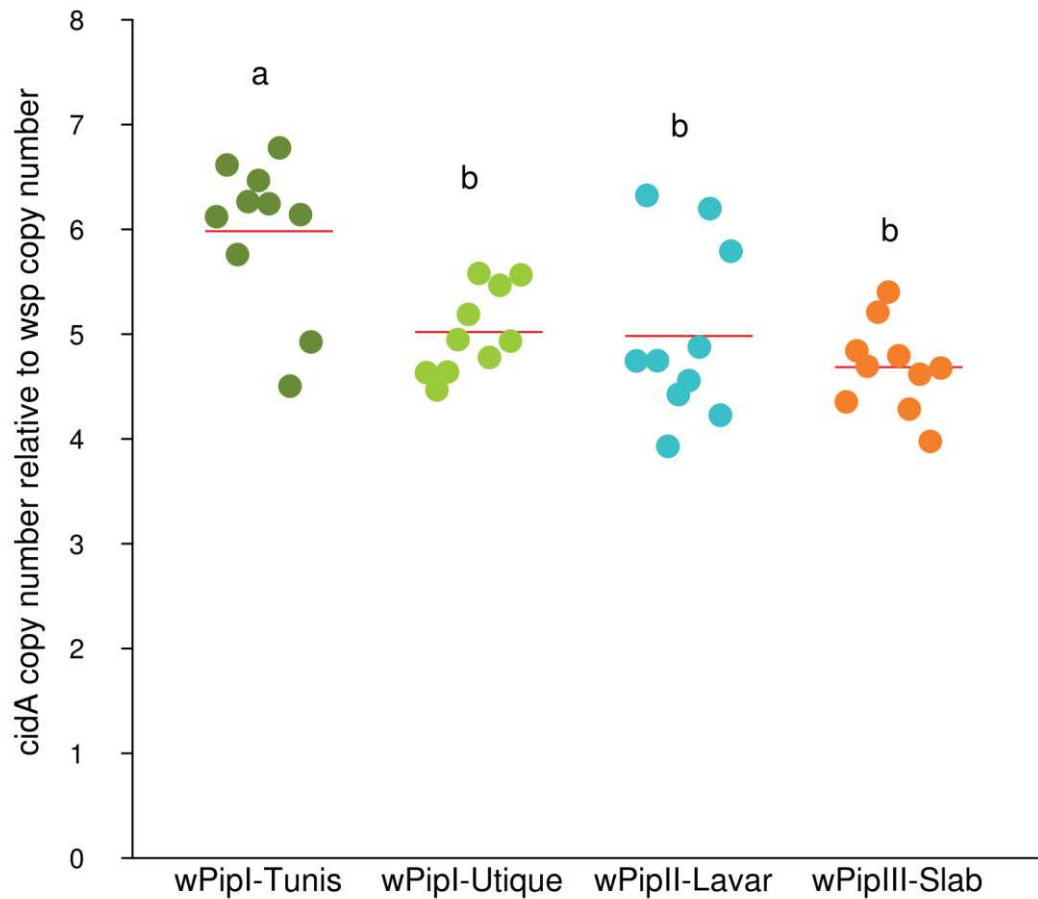
S2 Fig. Repertoires of CidA protein variants in the four MAL line *w*Pip strains.

Protein sequences alignment of the CidA variants found in the four *Wolbachia* strains *w*PipI-Tunis, *w*PipI-Utique, *w*PipII-Lavar and *w*PipIII-Slab (MAL lines). The first sequence is used as a reference to determine the polymorphic region. For more clarity, only polymorphic positions are represented, thus amino-acid positions are not continuous. When more than two contiguous amino-acids were variable the “-” symbol was used between the first and the last variable position of the zone. Colors show polymorphic blocks of residues present in variants regardless of their phylogenetic *w*Pip group (I to III). No *cidA* or *cidB* nucleotide sequence variant was shared between the three *w*Pip groups. However, the *w*PipII-Lavar CidA_II(α1) variant and the *w*PipIII-Slab CidA_III(β8) variant presented the same amino-acid sequence. Based on their nucleotide sequences *w*PipIII-Slab exhibited ten variants of *cidA*, *w*PipII-Lavar three, *w*PipI-Tunis four and *w*PipI Utique seven. However, *w*PipIII-Slab exhibited only seven variants that differ in their amino-acid sequences since *cidA_III*(γ6) and *cidA_III*(γ7), *cidA_III*(δ6) and *cidA_III*(δ7), *cidA_III*(δ8) and *cidA_III*(δ9) have respectively identical amino-acid sequences (*i.e.* nucleotide polymorphic positions between them are synonymous).



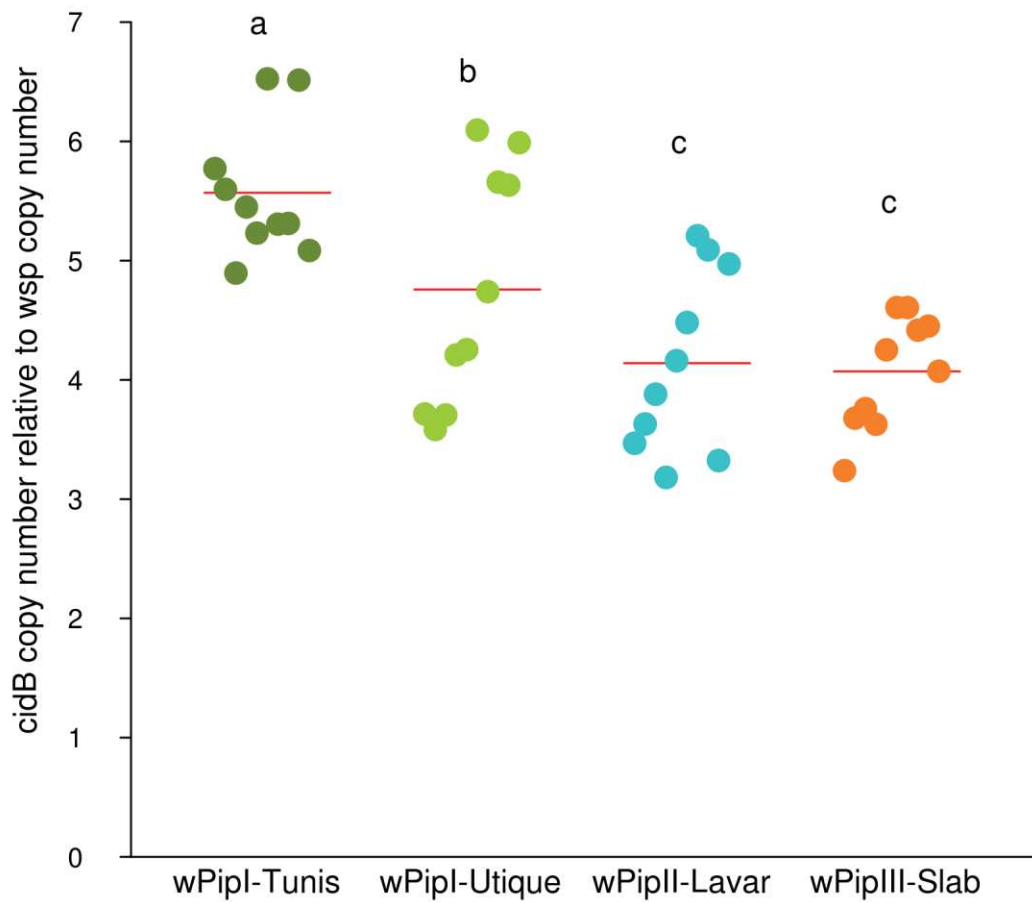
S4 Fig. Correlation between *cidB* variants diversity in *wPip* genomes and CI cellular intensity.

Lighter gray bar plot accounts for the 10 crosses involving males from Lavar and Utique lines hosting *wPip*, which harbor 2 different variants of *cidB* in their genomes, while darker gray bar plot accounts for the 10 crosses involving males from Tunis and Slab lines, both infected with *wPip* strain harboring 4 different variants of *cidB*. Error bars represent the standard error. The proportion of unhatched developed embryos was significantly higher for males hosting four-variants *wPip* strains than for males two-variants *wPip* strains (Wilcoxon, $W = 1159$, $p < 0.001$).



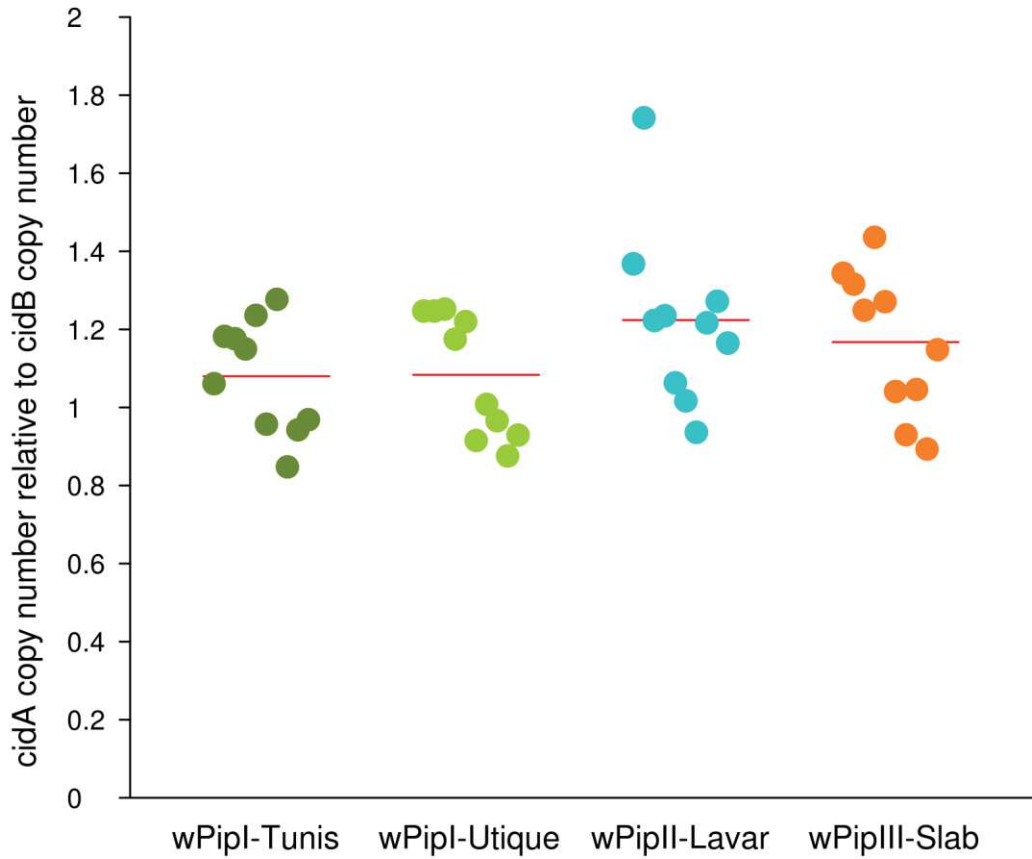
S5 Fig. *cidA* copy number in the *wPip* strains infecting the four MAL lines.

cidA copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidA* gene and the *Wolbachia wsp* gene. The colored dots represent the *cidA* copy number per *wPip* genome in a male and the red strips represent the average *cidA* copy number per *wPip* genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different).



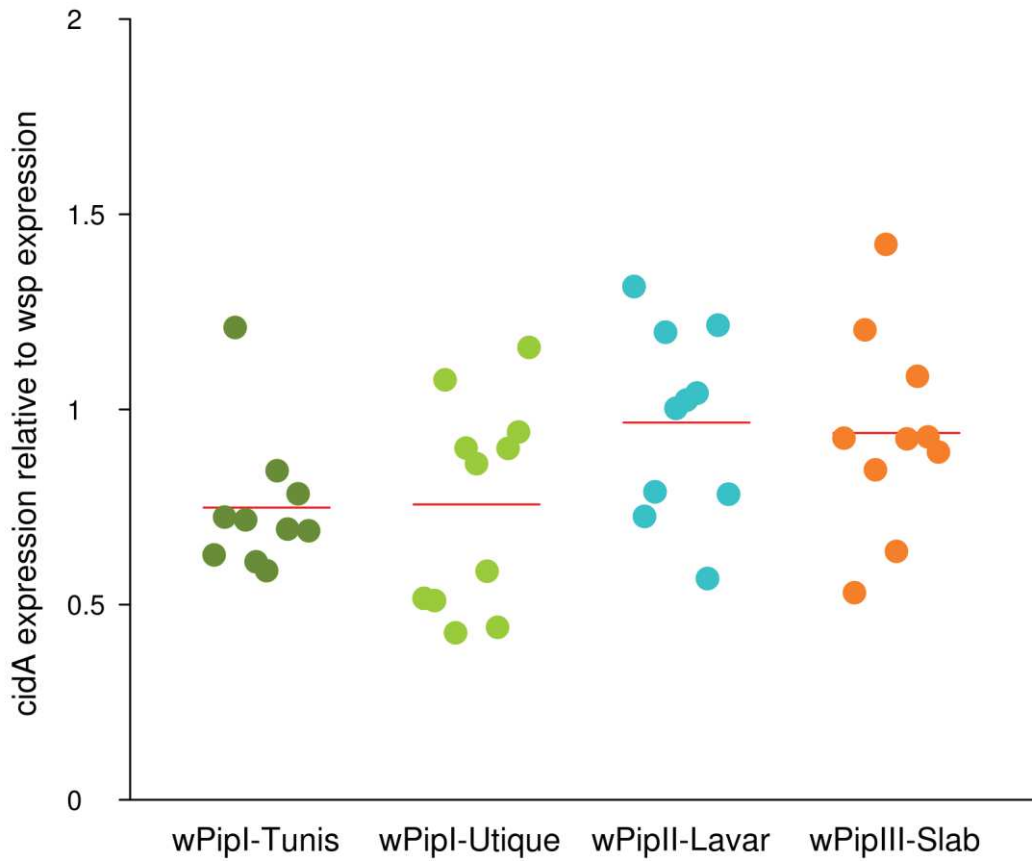
S6 Fig. *cidB* copy number in the *wPip* strains infecting the four MAL lines.

cidB copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidB* gene and the *Wolbachia wsp* gene. The colored dots represent the *cidB* copy number per *wPip* genome in a male and the red strips represent the average *cidB* copy number per *wPip* genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different).



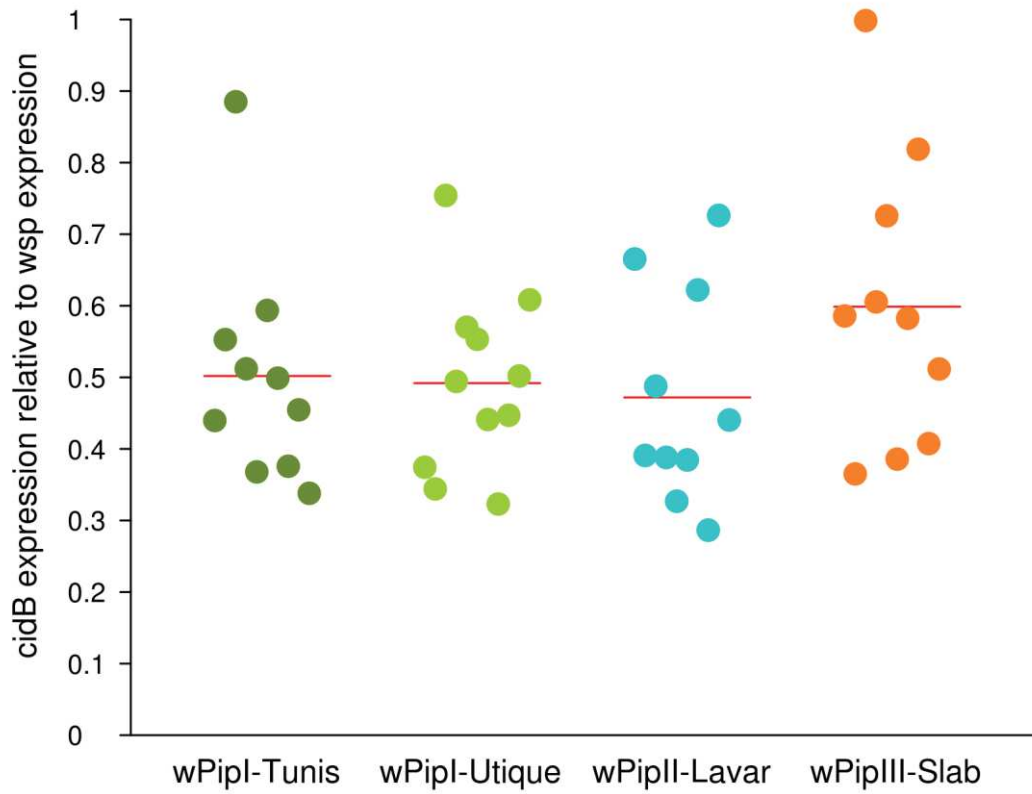
S7 Fig. *cidA/cidB* copy number in the *wPip* strains infecting the four MAL lines.

cidA/cidB copy number was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidB* gene and the *Wolbachia cidB* gene. The colored dots represent the *cidA/cidB* copy number per *wPip* genome in a male and the red strips represent the average *cidA/cidB* copy number per *wPip* genome for ten males per MAL lines. *cidA/cidB* copy number were not significantly different between the four *wPip* strains infecting the four MAL lines.



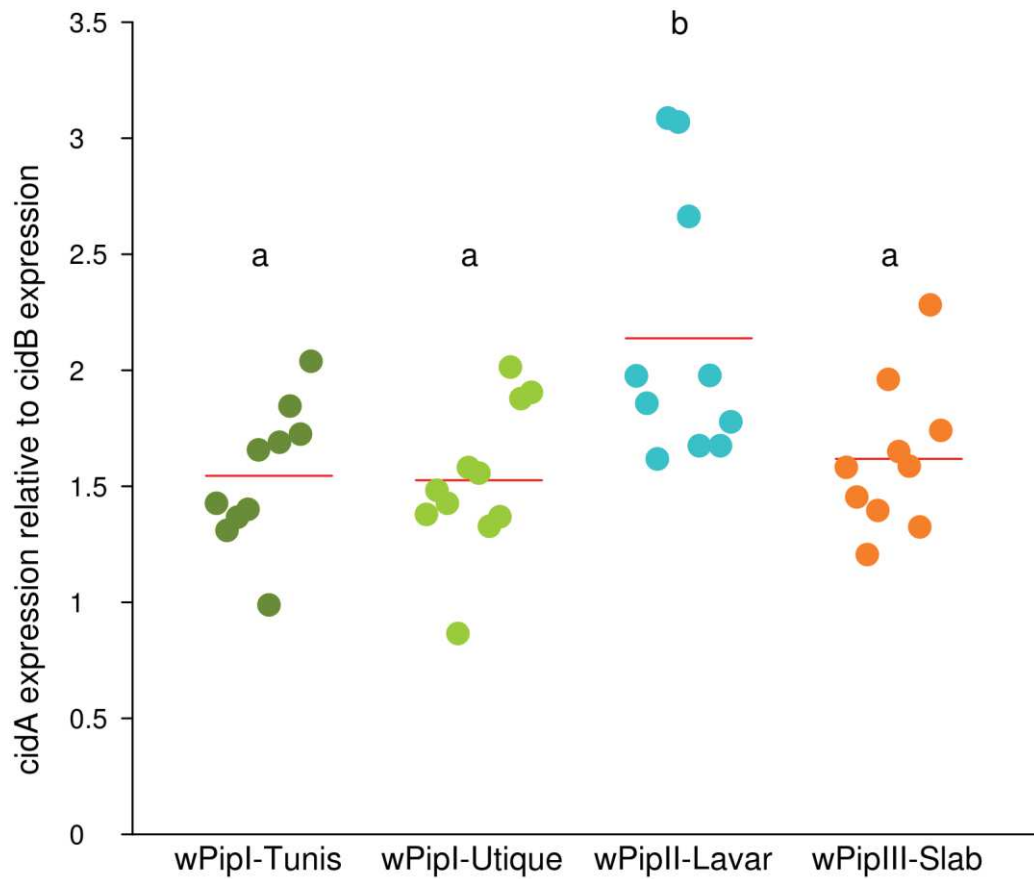
S8 Fig. *cidA* expression level in the *wPip* strains infecting the four MAL lines.

cidA expression was measured by quantitative PCR as the ratio between the *Wolbachia cidA* gene expression and the *Wolbachia wsp* gene expression. The colored dots represent the *cidA* expression level per *wPip* genome in a male and the red strips represent the average *cidA* expression level per *wPip* genome for ten males per MAL lines. Expression levels of *cidA* genes were not significantly different between the four *wPip* strains infecting the four MAL lines.



S9 Fig. *cidB* expression level in the *wPip* strains infecting the four MAL lines.

cidB expression was measured by quantitative PCR as the ratio between the *Wolbachia cidB* gene expression and the *Wolbachia wsp* gene expression. The colored dots represent the *cidB* expression level per *wPip* genome in a male and the red strips represent the average *cidB* expression level per *wPip* genome for ten males per MAL lines. Expression levels of *cidB* genes were not significantly different between the four *wPip* strains infecting the four MAL lines.



S10 Fig. *cidA/cidB* expression level in the *wPip* strains infecting the four MAL lines.

cidA/cidB expression levels was measured by quantitative PCR as the ratio between the number of copies of the *Wolbachia cidA* gene and the *Wolbachia cidB* gene. The colored dots represent the *cidA/cidB* expression level per *wPip* genome in a male and the red strips represent the average *cidA/cidB* expression per *wPip* genome for ten males per MAL lines. Letters represent the different statistical groups (*i.e.* means with the same letter are not significantly different).

Remerciements la suite

Best athletes ever

Merci à pierre [Olivier Antoine](#) pour avoir organisé le Marathon plusieurs années de suite, c'étaient toujours des moments géniaux et merci à tous les coureurs !!!

Merci à toute l'équipe de Volley Ball pour ces matchs de folies sous 40° : [Rémi](#), [Grand Yoann](#), [Yo](#), [Marjo](#), [Manon](#), [Andréa](#), [Benjamin](#), [Max](#), [Sergio](#), [Paul](#), [Fabien](#), [Manu](#)

Best PhD students and Postdocs Ever

Je ne pense pas que j'aurais pu passer trois aussi bonnes années de thèse sans vous tous. Vous avez tous contribué à ce que la vie à l'ISEM soit super et qu'il s'y passe tellement de choses inoubliables !

Merci à [Marjo](#), ma catholique repentie préférée, pour avoir été à mes côtés pendant toute la thèse, pour avoir accepté que je te taquine sans jamais bouder, pour être venu bosser le samedi avec moi, pour m'avoir soutenu pendant les moments difficiles, pour toutes ces fois où on est allée courir en râlant sur la thèse, pour m'avoir vendu du rêve à Evolution. Merci pour tout !

Merci à [Maud](#) d'avoir toi aussi été là pour m'écouter me plaindre et râler, pour tous les bons moments qu'on a passé ensemble à faire la fête ou juste à discuter !

Merci à [Clem](#) d'avoir été là quand la vie était belle mais aussi quand elle était pourrie et de m'avoir aidé à relativiser sur la thèse.

Merci à [Yo](#), pour tous ces moments de ouf passés à jouer à des jeux, à faire du sport, à faire la fête et à danser jusqu'au bout de la night sur de la transe ! Ces cinq ans à Montpellier n'auraient pas été les mêmes sans toi.

Merci à [Paul](#), d'avoir fait la démo de dissection des souris pour mes étudiants et merci pour toutes ces pauses clopes pendant lesquelles on a refait le monde et la science.

Merci à [Maeva](#) pour adorer autant faire la fête même si on sait toujours d'avance comment ça se termine !

Merci à [Sergio](#), parce qu'il faut toujours avoir un ami portugais qui fait des blagues de merde et qui aime les dinosaures et les nazis ^^

Merci à [Manu](#) d'avoir rendu Mine si heureuse

Merci à [Sam](#) d'être toujours chaud d'écouter du gros son !!!

Merci à [Cécile](#) de m'avoir appris à jouer du ukulélé, je n'ai pas été une étudiante très assidue mais maintenant que la thèse est finie je vais pouvoir m'y remettre à fond !

Merci à [Myriam](#) d'avoir toujours été motivée pour faire la fête même si tu avais kayak ;))

Merci à [Andrea](#) pour avoir apporté la calor espagnole à l'ISEM !

Merci à [Rémi](#) et [Grand Yoann](#) pour avoir monté l'équipe de Volley grâce à laquelle j'ai passé de si bons moments !

Merci à [Max](#) de m'avoir prêté son bouquin de Lovecraft (je ne l'ai pas encore ouvert mais promis après la thèse je le lis)

Merci à tous les autres doctorants : [Laura](#), [Alexis](#), [Quentin V](#), [Quentin M](#) (je travaille sur une recette de savon sans huile de palme :p), [Pierre](#), [Alexandre S](#), [Alexandre G](#), [Alain](#), [Michael](#) et [Alice](#).

Best Master Ever

Parce qu'avant d'arriver en thèse il faut passer par le master ! Merci à tous, deux super-années avec vous !

Best Friends Ever

Parce que trois ans de thèse c'est aussi ce qui se passe en dehors et heureusement qu'il y a un dehors !

Merci à ma femme **Ninon** pour être à mes côtés depuis 25 ans maintenant et avoir toujours été là pour le meilleur et pour le pire ! Et Merci à **Marion** la copine de ma femme pour la rendre si heureuse !

Merci à **Clem** parce que toi aussi tu sais ce que ça fait d'être en thèse tout en aimant écouter de la transe jusqu'au bout de la nuit dans des clubs underground de Vienne.

Merci à **Leslie** pour avoir fait le trajet jusqu'à Montpellier pour venir me voir et ne pas râler parce que je ne rentrais pas souvent à Bordeaux city beach.

Merci à **Alan**, une des plus belles rencontres que j'ai faites pendant ces 5 ans à Montpellier ! Merci pour tout (je ne peux pas lister il y a tant de choses à dire). Merci d'être tel que tu es and seen you soon in Denmark !

Merci à **Pauline**, **Adrien**, **Faustine** et **Adrien** pour m'avoir motivé à faire de l'escalade et avoir cru en ma capacité de faire des rouges et peut-être des blanches un jour qui sait !

Merci à **Cédric**, pour toutes ces soirées de folies passées au 5 rue de Lorraine et pour en avoir été avec laulau de dignes descendants quand je suis partie de la colloc' !

Merci à l'équipe de D&D : **Marjo**, **Yoann**, **Iago**, **Paul**, **Coco** pour ces longues heures de folie passées à pourfendre des gobelins dans des contrées lointaines.

Merci aux Boby **Alicia** et **Marine** pour ce merveilleux voyage en Norvège et la tentative d'apprentissage du Tamure.

Merci à la colloc' du 1706 Avenue de Maurin : **Iago**, **Charly**, **Marjo**, **Yo**, **Alain**, **Val**, **Marianne** pour toutes ces soirées et après-midis piscines jeux de société !

Merci à tous les autres copains de Montpellier et d'ailleurs : **Matthis**, **Juliette**, **Yon**, **Marion**, **Vincent**, **Martin**, **Alexandre**, **Emelyne**, les copains du master : **Julie**, **Guillaume**, **Lucas**, **Salomé**, **Marine S**, **Coline**, la bande des M1 2014-2015 : **Max**, **Suzane**, **Laure**, **Raph**, **Léo**, **Valentin**, **Laulau**, **Estelle**.

Bases génétiques et mécanismes cytologiques de la diversité de l'incompatibilité cytoplasmique induite par *Wolbachia* chez le moustique *Culex pipiens*.

Les *Wolbachia* sont des alpha-protéobactéries intracellulaires transmises verticalement de la mère aux descendants via les ovocytes. Du fait de ce mode de transmission, des stratégies de manipulation de la reproduction favorisant leur propagation ont été sélectionnées chez ces bactéries. La plus communément utilisée par *Wolbachia* s'appelle l'incompatibilité cytoplasmique (IC). L'IC a lieu lorsque des mâles infectés copulent avec des femelles non infectées ou infectées par des *Wolbachia* incompatibles et se traduit par la mort des descendants avant l'éclosion. L'IC est généralement conceptualisée comme un modèle *mod/resc* ou toxine/antidote dans lequel les *Wolbachia* présentes chez les mâles introduiraient dans les spermatozoïdes une toxine (fonction *mod*) qui, après la fécondation, entrainerait la mort des embryons, sauf si les *Wolbachia* présentes dans l'œuf produisent un antidote (fonction *resc*). C'est chez l'espèce de moustique *Culex pipiens* que la plus grande diversité de phénotypes d'IC a été décrite. Cette diversité repose uniquement sur la diversité des souches de *Wolbachia* hébergées par *C. pipiens*. Dans cette thèse, nous avons mené la première étude des mécanismes cytologiques responsables de la mort des embryons dans les croisements incompatibles chez *C. pipiens*. Nous avons montré que des défauts de condensation et de ségrégation de la chromatine paternelle lors de la première division embryonnaire entraînent la mort des embryons dans tous les croisements incompatibles étudiés. Ces défauts cellulaires sont les seuls observés, ce qui indique que la diversité de relations d'IC décrite chez *C. pipiens* ne repose pas sur une diversité de défauts cellulaires. L'étude, chez plusieurs souches de *wPips*, de l'opéron *cidA/cidB* dont l'implication fonctionnelle dans l'IC a récemment été mise en évidence chez la drosophile, nous a permis de montrer que cet opéron est amplifié et polymorphe dans tous les génomes de *wPips* séquencés. L'exploration des variants de cet opéron dans les génomes de *Wolbachia* infectant des populations naturelles de *C. pipiens*, à l'aide de plus de 250 lignées isofemelles, a permis d'associer de manière robuste une variation de *cidB* avec un changement dans le phénotype *mod* de certains mâles. En outre, la présence d'un variant ubiquitaire de *cidA*, supporte le rôle de ce gène dans la fonction *resc*. Ainsi, chez *C. pipiens*, l'opéron *cidA/cidB* grâce à son amplification et sa diversification est impliqué dans la diversité des phénotypes d'IC et fonctionnerait comme un système toxine-*mod* /antidote-*resc* : *cidB* étant impliqué dans la fonction *mod* et *cidA* dans la fonction *resc*.

Abstract

Genetic bases and cytological mechanisms underlying cytoplasmic incompatibility diversity induced by *Wolbachia* in the *Culex pipiens* mosquito.

Wolbachia are intracellular alpha-proteobacteria vertically transmitted from mothers to their offspring through oocytes. As a consequence of this transmission mode, reproductive manipulation strategies that promote bacteria spread in host populations have been selected. The most common manipulation used by *Wolbachia* is called cytoplasmic incompatibility (CI). CI occurs when infected males mate with uninfected or incompatible *Wolbachia*-infected females and results in the death of offspring before hatching. CI is generally conceptualized as a *mod/resc* or toxin/antidote model in which paternal *Wolbachia* would introduce a toxin (*mod* function) in sperms which would, after fertilization, induce embryonic death unless an antidote produced by maternal *Wolbachia* in the egg counteracts its effect (*resc* function). A to date unique diversity of CI phenotypes has been described in the mosquito species *Culex pipiens*. This diversity is based solely on the diversity of *Wolbachia* strains hosted by *C. pipiens*. In this PhD, we conducted, in *C. pipiens*, the first study of the cytological mechanism behind embryonic mortality in CI crosses. We showed that paternal chromatin condensation and segregation defects during the first embryonic division were responsible for embryonic death in all CI crosses. These CI defects were the only ones observed indicating that the diversity of CI phenotypes in *C. pipiens* is not based on a diversity of cellular mechanisms. We then studied the *cidA/cidB* operon in several *wPip* strains as the functional involvement of this operon in CI was recently demonstrated in *Drosophila*. We showed that this operon is amplified and polymorphic in all genomes of sequenced *wPip*. Investigation of *cidA/cidB* variants in *Wolbachia* genomes infecting natural populations of *C. pipiens*, using more than 250 isofemale lines, enabled us to reveal a robust association between *cidB* variations and change in *mod* phenotype. In addition, the presence of an ubiquitous *cidA* variant supports the role of this gene in the *resc* function. In *C. pipiens*, the *cidA/cidB* operon, through its amplification and diversification, is involved in the CI phenotypes diversity and would operate as a toxin-*mod* / antidote-*resc* system: *cidB* being involved in the *mod* function and *cidA* in the *resc* function.