UPMC
SORBONNE UNIVERSITÉS

Université Pierre et Marie Curie

Ecole doctorale : Complexité du vivant

Equipe : Evolution des Caenorhabditis

Base génétique de la sensibilité au virus d'Orsay au sein des populations naturelles de *Caenorhabditis elegans*

Présenté par **Tony Bélicard**

Thèse de doctorat en Génétique évolutive

Dirigée par Mme Marie-Anne Félix

Présentée et soutenue publiquement le 18 Septembre 2014

Devant un jury composé de :

Mme Tatiana Giraud	Rapportrice
M. Jonathan Ewbank	Rapporteur
M. Louis Lambrechts	Examinateur
M. Dominique Higuet	Examinateur
Membre invité : Mme Marie-Anne Félix	Directrice de thèse

Abstract

Caenorhabditis elegans is a commonly studied animal model in laboratories. The discovery of the first natural viral infections of *Caenorhabditis* brings new models to study animal-virus interactions.

The Orsay virus, specifically infecting *C. elegans*, and the Santeuil virus, specifically infecting *C. briggsae*, are positive single strand RNA viruses (Hepatites, Chikungunya, Coronavirus etc…) disrupting the structure of intestinal cells of their host. However, we observed a strong variability in the sensitivity to those viruses at the intraspecific level.

To identify the genetic basis of the sensitivity, we performed a genome wide association study on 97 wild isolates of *C. elegans*. We were able to identify the center of chromosome IV as a region containing the locus responsible for this sensitivity. A deletion in the *drh-1* gene, coding for a RIG-I-Like protein, confers sensitivity to their carrier. RIG-I is known to recognize viral RNA and to trigger an antiviral response through the production of interferons in vertebrates. However, *C. elegans* is not able to produce interferons but it appears that DRH-1 initiates a viral specific siRNA pathway.

Immunity genes are under strong selective pressure. Thus, it is surprising that such an important protein for the antiviral pathway appears to be disrupted in 23% of the wild isolates. This deletion shows high linkage disequilibrium with a broader region of 3Mb, suggesting that the deletion propagates with this region. However, this region does not seem to provide any advantage to their owner under laboratory conditions.

Résumé

Caenorhabditis elegans fait partie des modèles animaux les plus étudiés en laboratoire. La découverte des premiers virus infectant naturellement des *Caenorhabditis* apporte de nouveaux modèles animal-virus très prometteurs.

Le virus d'Orsay, infectant spécifiquement *C. elegans*, et le virus de Santeuil, infectant spécifiquement *C. briggsae*, sont des virus à ARN positif simple brins (Hépatites A, C et E, Chikungunya, Coronavirus etc…) qui provoquent une désorganisation de la structure des cellules intestinales de leur hôte. Par ailleurs, une variation dans la sensibilité à ce virus au sein des espèces est observée.

Grâce à une étude statistique d'association pangénomique portant sur la sensibilité au virus d'Orsay de 97 isolats naturels de *C. elegans*, nous avons pu mettre en évidence une région chromosomique qui contient un polymorphisme responsable de cette sensibilité. Il s'agit une délétion dans le gène *drh-1* qui code une protéine similaire à la protéine RIG-I, connue pour initier une réponse antivirale chez les vertébrés. Cependant, *C. elegans* est incapable de produire la même réponse antivirale que ces derniers. Ainsi, DRH-1 est impliquée dans la reconnaissance des ARN viraux et dans la production d'une réponse spécifique à l'infection virale par les petits ARN. Les gènes impliqués dans l'immunité sont soumis à une forte pression de sélection.

Pourtant, de manière surprenante, la délétion du gène *drh-1*, appartenant à une région invariante de 3 MB, est présente dans 23% des isolats étudiés. La propension qu'a cet allèle à se propager dans les populations naturelles sans pour autant fournir un avantage sélectif évident en conditions de laboratoire sera discuté.

Table des matières

INTRODUCTION *11*

I. Importance écologique du parasitisme **15**

I.A. Définition du parasitisme 15

I.B. Diversité systématique des parasites 15

I.B.1. Les virus 17

I.B.2. Les procaryotes 19

I.B.3. Les eucaryotes 20

I.C. Notion de spécificité d'hôte 21

I.D. Cycle de vie et traits d'histoire de vie des parasites 23

I.D.1. Dispersion 23

I.D.2. Infection et établissement dans l'hôte 25

I.D.3. Exploitation de l'hôte, croissance et reproduction (endoparasites) 25

I.E. Séquence de défense de l'hôte 27

I.E.1. Défenses antérieures à l'infection 27

I.E.2. Défenses postérieures à l'infection 31

I. Conclusion 32

II. Diversité et évolution des défenses immunitaires **33**

II.A. Les effecteurs de l'immunité ou "comment neutraliser les parasites". 33

II.A.1. Les différentes réponses cellulaires 33

II.A.2. Les différentes réponses humorales 40

II.B. Reconnaissance du soi et réponses adaptées à l'infection 43

II.B.1. Les Récepteurs riches en répétitions de leucines 43

II.B.2. Les voies de signalisations indépendantes de domaines LRR 44

II.C. Une immunité innée et une immunité acquise pour tous 47

II.C.1. Acquisition d'une spécificité immunitaire 49
II.C.2. Maintien/mise en mémoire d'une spécificité immunitaire 50

II. Conclusion 50

III. Bases génétiques et microévolution des interactions hôtes-parasites 53

III.A. Bases génétiques des interactions hôte-parasite 53

III.A.1. Architecture génétique de la résistance aux parasites chez l'hôte 53
III.A.2. Bases génétiques de l'infection: interactions entre le génotype du parasite et le génotype de l'hôte 57

III.B. Impact de la coévolution sur la diversité génétique des hôtes et des parasites 60

III.B.1. Le modèle gène-pour-gène: une succession de balayages sélectifs 60
III.B.2. Le modèle "matching allele": sélection négative dépendante de la fréquence 63
III.B.3. Adaptation locale 64

III.C. Signatures de la sélection par les parasites 64

III.C.1. Comment détecter la sélection dans les génomes ? 65
III.C.3. Quel type de sélection sur le système immunitaire? 68

III. Conclusion 69

IV. Le nématode *Caenorhabditis elegans* 70

IV.A. Place de *C. elegans* dans le vivant 70

IV.A.1. Les nématodes dans l'arbre du vivant 70
IV.A.2. Phylogénie des nématodes 72

IV.B. Anatomie et cycle de vie de *C. elegans* 72

IV.B.1. Anatomie des adultes 72
IV.B.2. Cycle de vie 79

IV.C. Ecologie et génétique des populations de *C. elegans* 80

IV.C.1. Où trouver *C. elegans*? 80
IV.C.2. Quelles sont les meilleures conditions pour trouver *C. elegans*? 81
IV.C.3. Génétique des populations 81
IV.C.4. Interactions avec les autres organismes 85

IV.D. Systèmes de défense chez *C. elegans* 89

IV.D.1. Défenses intestinales _____ 90

IV.D.2. Défenses hypodermiques _____ 93

IV.D.3. Défenses du rectum _____ 95

IV.D.4. Rôle du système nerveux _____ 95

IV.D.5. Les voies des petits ARN chez *C. elegans* _____ 97

IV. Conclusion 101

RESULTATS ***..... 103***

I. Infection naturelle et expérimentale de nématodes *Caenorhabditis* par de nouveaux virus proches des Nodavirus 104

II. Une délétion naturelle de l'homologue de RIG-I abolit le clivage de l'ARN viral et l'immunité antivirale chez *Caenorhabditis elegans* 127

DISCUSSION 163

I. Les premières infections virales naturelles de *Caenorhabditis* : implications 165

I.A. Un nouveau groupe de virus proche des nodavirus 165

I.A.1. Structure des virus de *Caenorhabditis* _____ 165

I.A.2. Relation phylogénétique avec les nodavirus _____ 167

I.A.3. Phénotypes liés à l'infection _____ 167

I.B. Les *Caenorhabditis* et leur virus : un modèle pour l'évolution des interactions animal-virus 171

I.C. Un rôle majeur des petits ARN dans la lutte contre les infections virales 173

I.C.1. La voie siARN amorcée par DRH-1 constitue une réponse antivirale chez *C. elegans*. _____ 173

I.C.2. Les RIG-I-Like Receptors chez les mammifères set les DRH chez *C. elegans*: un même combat, des armes différentes. _____ 177

II. Histoire évolutive de l'interaction *Caenorhabditis* – Virus naturels 180

II.A. Les virus sont spécifiques à l'échelle de l'espèce 180

II.B. La capacité de réplication du virus d'Orsay varie entre souches sauvages de *C. elegans* 180

II.C. Histoire évolutive de *niDf250* 184

II.C.1. Répartition de l'allèle niDf250 dans les populations naturelles _____ 184

II.C.2. Des traces de balayage sélectif récent autour de niDf250 _____ 185

II.C.3. Quelle pression de sélection pourrait avoir conduit à un balayage sélectif dans cette région ? _____ 185

II.D. Le polymorphisme de *drh-1* n'explique pas totalement la sensibilité au virus d'Orsay 187

Conclusion.....192

BIBLIOGRAPHIE 193

ANNEXE.....218

Index des figures et tableaux

Figure 1. Pressions de sélections exercées sur une espèce dans la nature. 12
Figure 2. Exemple d'apparitions indépendantes du parasitisme.. 16
Figure 3. Classification de Baltimore. 18
Figure 4. Facteurs affectant la gamme d'hôte d'un parasite. 22
Figure 5. Patrons d'évolution des hôtes et des parasites associés à l'échelle macroévolutive. 24
Figure 6. Différents scénarios permettent d'expliquer les écarts à la congruence des phylogénies. 26
Figure 7. Cycle de vie typique d'un parasite. 28
Figure 8. Diversité des mécanismes de défense de l'hôte face à un parasite. 30
Figure 9. Processus de phagocytose dans l'immunité. 34
Figure 10. Mécanisme simplifié du système immunitaire acquis contre un parasite extra-cellulaire chez les vertébrés. 36
Figure 11. Réponse antivirale par les petits ARN chez *Arabidopsis thaliana* et *Drosophila melanogaster*. 38
Figure 12. Récepteurs membranaires et voies de signalisations impliquées dans les défenses antiparasitaires. 42
Figure 13. Récepteurs cytoplasmiques associée à une infection virale chez les mammifères. 46
Figure 14. Les différents types d'immunités rencontrés dans le vivant. 48
Tableau 1. Evolution du système immunitaire. 51
Figure 15. Infection hypothétique de deux génotypes d'hôte par deux génotypes de parasite. 55
Figure 16. Mise en évidence d'interactions entre génotypes de l'hôte et du parasite. 56
Figure 17. Modèles d'interaction entre les génotypes de l'hôte et du parasite..... 58
Figure 18. Evolution des fréquences allèliques au cours du temps dans un système d'interactions antagonistes..... 61
Figure 19. Exemple de sélection négative dépendante de la fréquence et de balayage sélectif. 62
Figure 20. Phylogénie des espèces de *Caenorhabditis* cultivées en laboratoire. 71
Figure 21. Anatomie *de C. elegans*..... 74
Figure 22. Appareil digestif chez *C. elegans*. 76
Figure 23. Cycle de vie d'un hermaphrodite de *C. elegans* au laboratoire à 20°C. 78
Figure 24. Habitat naturel et comportement migratoire chez *C. elegans*. 82
Figure 25. Polymorphisme le long des chromosomes chez *C. elegans*. 86

Figure 26. Parasites naturels de *C. elegans*. 88
Tableau 2. Parasites capables d'infecter *C. elegans*. 91
Figure 27. Défense immunitaire de différents tissus chez *C. elegans*. 94
Figure 28. Réponse antivirale par les petits ARN chez *C. elegans*. 96
Figure 29. Analyse phylogénétique des virus de *Caenorhabditis* et des Nodaviridae, selon la méthode « neighbour-joining ».. 166
Figure 30. Localisation subcellulaire des protéines codées par le virus d'Orsay. 168
Figure 31. Cycle de vie des virus de *Caenorhabditis* tel qu'il est connu aujourd'hui... 170
Figure 32. Arbre phylogénétique des domaines DExD/H hélicases à ARN. 176
Figure 33. Suivie de l'infection du virus d'Orsay à partir d'infections effectuées au stade L1. 181
Figure 34. Détection du RNA1 du virus d'Orsay par FISH dans des mutants *rde-1*. ... 182
Figure 35. Effet de la délétion niDf250 sur l'infection par les microsporidies en présence ou en absence du virus d'Orsay. 188
Figure 36. Modèle de compétition entre la voie des microARN et des siARN antiviraux chez JU1580. 189
Figure 37. Courbes de survie de *Caenorhabditis elegans* de la souche JU1580 préalablement infectés ou non par le virus d'Orsay nourris exclusivement par *Pseudomonas aeruginosa*. 191

Abréviations

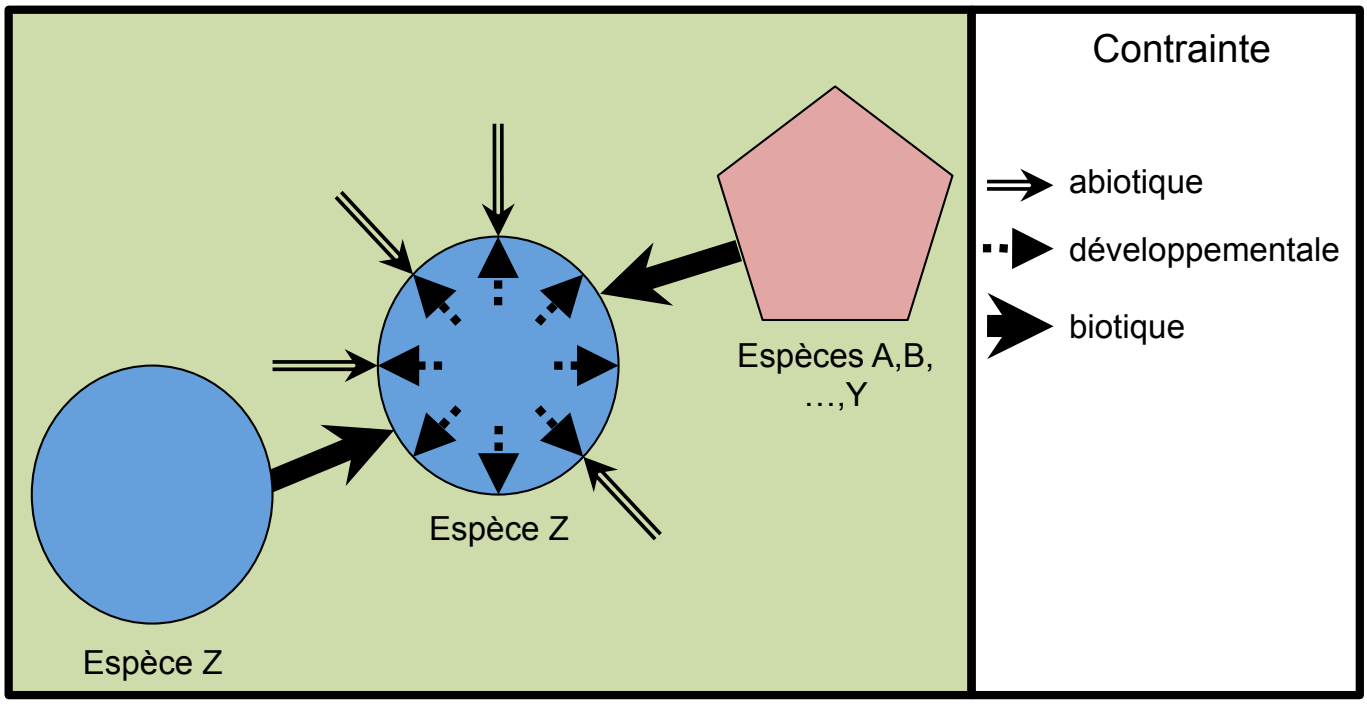
ABF	AntiMicrobial Factor
ADN	Acide désoxyribonucléique
AFLP	Amplification Fragment Length Polymorphism
AGO	ArGOnaute
AMP	AntiMicrobial Peptide
ARN	Acide RiboNucléique
ARNm	ARN méssager
ARNr 18S	ARN ribosomique de la petite sous-unité
BCR	B-Cell Receptor
CARD	Caspase activation et recruitment *domain*
Cas	CRISPR-associated
Cer1	*C. elegans* retrotransposon 1
CLR	CATERPILLAR
CMH	Complexe Majeur d'Histocompatibilité
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
INS-7	INSulin related 7
DAF-16/FOXO	abnormal DAuer Formation 16/ FOrkhead boX O
DAF-2/InsR	abnormal DAuer Formation 2/Insulin Receptor
DAF-4	abnormal DAuer Formation 4
DAG	DiAcylGlycerol
DBL-1	DPP/BMP-Like 1
DCL	DiCer-Like
DCR	DiCeR
DKF-2	D Kinase Family 2
DRH	Dicer-Related Helicase
Dscam	Down syndrom adhesion molecules
EAP	Etude d'Association Pangénomique
EGL-8	EGg Laying defective 8
EGO-1	Enhancer of Glp-One 1
EKL-1	Enhancer of Ksr-1 Lethality 1
FHV	Flock House Virus
FISH	Fluorescent In Situ Hybridization
FREP	Fibrinogen Related Protein
FSH	Follicule Stimulating Hormone
FSHR-1	Follicule Stimulating Hormone Receptor 1
F_{st}	Indice de fixation

GFP	Green Fluoresent Protein
GPA-12	G Protein, Alpha subunit 12
GWAS	Genome Wide Association Study
Imd	Immuno deficiency
IRAK	IL-1 Receptor Associated Kinase
IRF	Interferon Regulatory transcription Factor
Jak-STAT	Janus Kinase-Signal Transducer et Activator of Transcription
LCQ	Loci de Caractère Quantitatif
LGP2	Laboratory of Genetics et Physiology 2
LIN-45	abnormal cell LINeage 45
LRR	Leucin Rich Repetition
MAPK	Mitogen-Activated Protein Kinase
MDA-5	Melanoma Differentiation-Associated protein 5
MEK-2	MAP kinase kinase or Erk Kinase 2
MPK-1/ERK	MAP Kinase 1 / Extracellular signal-Regulated Kinases
MyD88	Myeloid Differentiation rimary response gene 88
NBS-LRR	Nuclear Binding Site - Leucin Rich Repeat
NFDS	Negative Frequency Dependant Selection
NF-κB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
NK	Natural Killer
NLP	Neuropeptide-Like Peptide
NSY-1	Neuronal Symmetry 1
ORF	Open reading Frame
PAM	Protospacer Adjacent Motif
PAMP	Pathogen Associated Molecular Pattern
PGRP	PeptidoGlycan Recognition Proteins
PKC-3	Protein Kinase C 3
PLC-3	PhosphoLipase C 3
PMK-1	P38 Map Kinase family 1
PPO	ProPhenolOxydase
PRR	Pattern Recognition Receptor
QTL	Quantitative Trait Loci
RACK-1	Receptor of Activated C Kinase 1
RAD	Restriction Associtated DNA
RAG-1	Recombinase-Activating Gene 1
RAG-2	Recombinase-Activating Gene 2
RDE-1	RNAi DEfective
RDRP	Polymérase à ARN Dépendante de l'ARN
RIG-I	Retinoic acid-Inducible Gene 1
RISC	RNA-induced silencing complex
RLR	RIG-I-Like-Receptor

ROS	Reactive Oxygen Species
RRF-1	RNA-dependent RNA polymerase Family 1
RT-PCR	Reverse Transcription Polymerase Chaine reaction
SAR	Systemic Aquired Resistance
SEK-1	SAPK/ERK kinase 1
siARN	small interfering RNA
SNF-12	Sodium Neurotransmitter symporter Family 12
SNP	Single Nucloetide Polymorphism
TALEN	Transcription Activator-Like Effector Nuclease
TCR	T-Cell Receptor
TGF-β	Transforming Growth Factor beta
TIR	Toll-Interleukin Receptor
TLR	Toll-Like Receptor
TNF	Tumor Necrosis Factor
TNFR	Tumor Necrosis Factor Receptor
TPA-1	Tetradecanoyl Phorbol Acetate resistant 1
UPR	Unfolded Protein Response
VSR	Viral Suppressor of RNAi
VSV	Vesicular stomatistis Virus
WAGO	Worm ArGOnaute protein

INTRODUCTION

A



B

Sp. Z / *Sp. Y*	Nuisible	Neutre	Bénéficiaire
Nuisible	Compétition symétrique	Amensalisme	Prédation, Parasitisme, Herbivorisme
Neutre	-	Neutralisme	Commensalisme
Bénéficiaire	-	-	Mutualisme

Figure 1. Pressions de sélections exercées sur une espèce dans la nature. A: Les contraintes exercées sur une espèce Z proviennent à la fois de son environnement physico-chimique (abiotique), des espèces avec qui elle interagit (biotique) et de facteurs internes à l'espèce (développementale).
B: Les types d'interaction entre les espèces Y et Z sont caractérisés par l'effet qu'a chacune des espèces sur l'autre.

Jusqu'à peu, l'écologie et la génétique sont restées deux domaines si distincts que la première n'est toujours pas considérée à l'heure actuelle par certains comme partie intégrante de la biologie. Tandis que l'écologie s'intéresse aux interactions entre les organismes et leur environnement, la génétique s'occupe de comprendre le fonctionnement interne des organismes. La génétique des populations tend à unifier ces deux approches. L'avènement du séquençage à haut débit a permis, entre autre, un développement rapide de cette discipline et il est aujourd'hui de plus en plus facile d'élucider les bases génétiques d'un processus écologique. Par conséquent, il est désormais envisageable d'avoir une approche intégrative de l'évolution des organismes en partant des processus moléculaires (les mutations *sensu largo*) jusqu'aux processus démographiques (migration, dérive) en passant par les processus phénotypiques (sélection).

Les facteurs environnementaux influençant l'évolution des être vivants sont de différentes natures (Figure 1A): A) les facteurs abiotiques, définis essentiellement par des paramètres physico-chimiques (humidité, température, pH etc.), généralement stables sur une échelle évolutive, B) les facteurs biotiques (interactions intra ou interspécifiques), beaucoup plus changeants et C) une combinaison des deux facteurs. La sélection naturelle énoncée par Charles Darwin (Darwin, 1859), repose sur le fait que les ressources naturelles, nécessaires à la croissance et à la reproduction, sont limitées. Ainsi, les organismes sont en permanence soumis à la sélection exercée par les autres organismes constituant leur environnement (Figure 1B). Cette sélection peut provoquer l'extinction de certaines espèces ou la spécialisation de certaines autres pour de nouvelles niches écologiques. Les rapports entre organismes peuvent être multiples. Par exemple, les interactions mutualistes résultent de la coopération entre deux protagonistes, chacun des deux organismes étant alors bénéficiaire (absorption de nutriments, protection etc.…). Dans le cas du commensalisme, un seul des protagonistes est bénéficiaire sans pour autant avoir un impact nuisible sur l'autre. Enfin, il est fréquent qu'un des deux protagonistes exploite l'autre. C'est le cas des interactions antagonistes que sont la prédation et le parasitisme. Le prédateur/parasite se nourrit ou utilise les ressources aux dépens de sa proie/son hôte.

Les parasites ont toujours présenté un intérêt majeur que ce soit pour la recherche fondamentale mais aussi bien évidemment pour la recherche médicale, agronomique ou vétérinaire. L'enjeu principal est de réussir à éradiquer les parasites nocifs pour l'homme ou pour ses activités. L'étude des interactions hôte-parasite passe par l'utilisation d'organismes modèles qui permettent la compréhension des mécanismes moléculaires impliqués dans l'établissement de l'infection par un agent pathogène et dans la réponse de l'hôte. La découverte du premier virus naturel du modèle *Caenorhabditis elegans*, le virus d'Orsay, ouvre de nouvelles possibilités quant à l'étude des interactions animal-virus. Les systèmes animal-virus sont peu représentés en écologie évolutive et les outils

précédemment développés chez *C. elegans* vont permettre une étude plus approfondie des interactions entre les deux protagonistes.
L'objectif de cette introduction est de faire une synthèse de ce qui est connu sur l'impact du parasitisme sur l'évolution des êtres vivants et de présenter le modèle d'étude.

I. Importance écologique du parasitisme

I.A. Définition du parasitisme

A l'origine, le concept de parasitisme définit l'utilisation d'un organisme, l'hôte, comme un habitat et une source de nourriture par un autre organisme, le parasite (Leuckart, 1879). Cette définition est principalement basée sur les relations trophiques entre deux espèces. Bien que juste, elle reste imprécise quant à la nature des interactions entre les deux organismes. Au-delà de l'aspect trophique, il est aussi important de considérer l'effet réciproque des deux antagonistes. Sont donc considérés comme parasites exclusivement les symbiotes ayant un effet négatif sur la fitness de leur hôte, ce dernier étant indispensable à la survie du parasite. Une définition plus mécanistique du parasitisme veut que la survie du parasite soit dépendante d'au moins un gène ou du produit d'au moins un gène de son hôte. Les interactions entre les parasites et leurs hôtes sont tellement diverses et variées qu'il est difficile de définir très précisément ce concept. Déterminer si un organisme est un parasite dépend particulièrement de l'environnement dans lequel il est observé. Par exemple, l'infection par certains virus végétaux (virus de la mosaïque) procure à leurs hôtes une tolérance accrue à la sécheresse ou au gel en augmentant la production d'osmoprotecteurs et d'antioxydants (Xu et al., 2008). *Wolbachia pipientis*, une bactérie à Gram- endosymbiotique féminisante présente chez un grand nombre d'arthropodes, dont 65% des insectes, procure une protection contre l'infection par le Drosophila C virus chez *Drosophila melanogaster* (Osborne et al., 2009). L'utilisation de certaines souches de *Wolbachia* est aujourd'hui envisagée pour réduire la transmission du virus de la dengue par le moustique *Aedes aegypti* (Bull et Turelli, 2013).

I.B. Diversité systématique des parasites

L'émergence du mode de vie parasitaire, à cause de son ancienneté, est difficile à évaluer précisément. Les premières traces fossiles répertoriées de parasitisme datent du Cambrien (570 Ma) (Walossek et Müller, 1994). On considère que plus de la moitié des espèces sur Terre sont parasitaires pour au moins une étape de leur cycle de vie (May, 1988). Le parasitisme n'est pas caractéristique d'un groupe monophylétique particulier

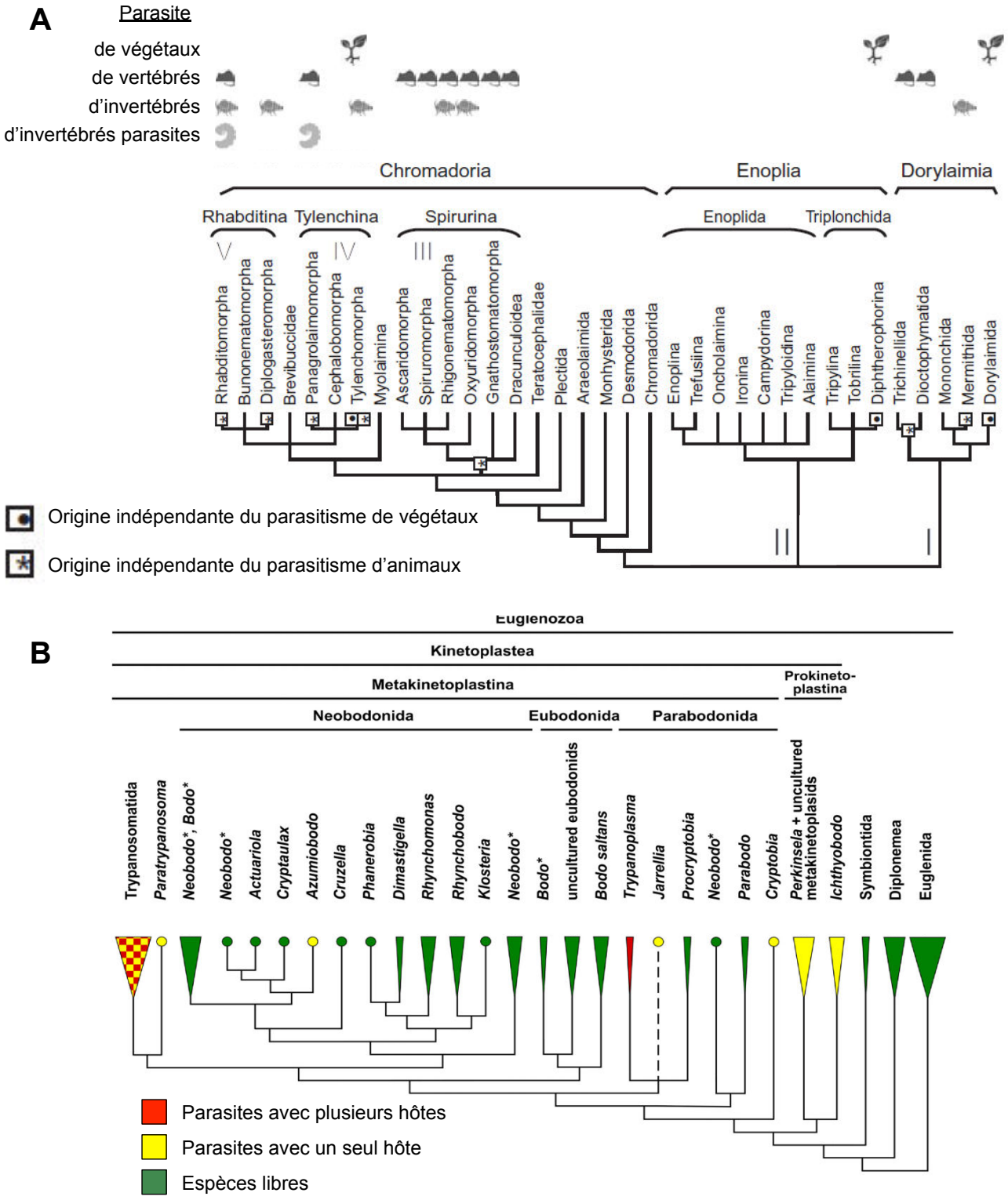


Figure 2. Exemple d'apparitions indépendantes du parasitisme. A: Phylogénie des nématodes réalisée à partir de séquences de petites sous-unités ribosomales de 300 espèces de nématodes, complétées par des caractères morphologiques. Selon cette phylogénie, le mode de vie parasitaire est apparu au moins dix fois indépendamment chez les nématodes. Adapté de Blaxter (2003).

B: Phylogénie des kinétoplastides réalisée à partir de séquences de petites sous-unités ribosomales. Selon cette phylogénie, le parasitisme est apparu au moins six fois indépendamment chez les kinétoplastides. Adapté de Lukeš et al. (2014).

mais se retrouve régulièrement dans l'arbre du vivant. Le mode de vie parasitaire, considéré comme un état dérivé par rapport à l'état dit libre, serait donc apparu plusieurs fois indépendamment lors de l'évolution (exemple Figure 2 avec (Blaxter, 2003; Lukeš et al., 2014). Tous les phylums ne sont pas concernés de manière égale par ce mode de vie. Ainsi, la plupart des virus (si considérés comme vivants) sont des parasites. Parmi les eucaryotes, une grande proportion des plathelminthes sont parasitaires (Campos et al., 1998) tandis qu'aucun échinoderme n'a été décrit comme tel. D'autres phylums, comme les nématodes ou les champignons contiennent environ 50% d'espèces libres et 50% d'espèces parasitaires.
Toutes les espèces existantes, parasitaires comprises, peuvent potentiellement être la cible de parasites. On sait maintenant que les virus eux-mêmes peuvent être parasités. Par exemple, le virus Sputnik, un virophage, exploite la machinerie de réplication du Mimivirus lorsque celui-ci se réplique dans une amibe (La Scola et al., 2008).

I.B.1. Les virus

La classification de Baltimore définit sept classes de virus (I à VII) basées sur la description de la production de l'ARNm viral à partir du génome initial (Figure 3); (Baltimore, 1971).

I.B.1.α. Les virus à ADN

Ces virus détournent l'ADN polymérase d'organismes eucaryotes pour pouvoir se répliquer lors de la division cellulaire à l'intérieur du noyau. Ils sont parfois capables de déréguler le cycle cellulaire, provoquant ainsi certains cancers ((Oncogenic DNA viruses, 2001).
La classe I est composée des virus dont le génome est constitué d'ADN double brin circulaire ou linéaire. La production d'ARNm se fait directement par transcription classique.
La classe II contient les virus à ADN simple brin circulaire ou linéaire. Après une étape de réplication, ce virus devient alors double brin et peut donc être répliqué et transcrit de manière classique.

I.B.1.β. Les virus à ARN

Ces virus ne passent pas par un intermédiaire ADN et le détournement de la machinerie de transcription de l'hôte n'est pas nécessaire car ils codent pour leur propre Polymérase à ARN Dépendante de l'ARN (RDRP). La réplication des génomes viraux est cytoplasmique.
La classe III regroupe les virus à ARN double brin.

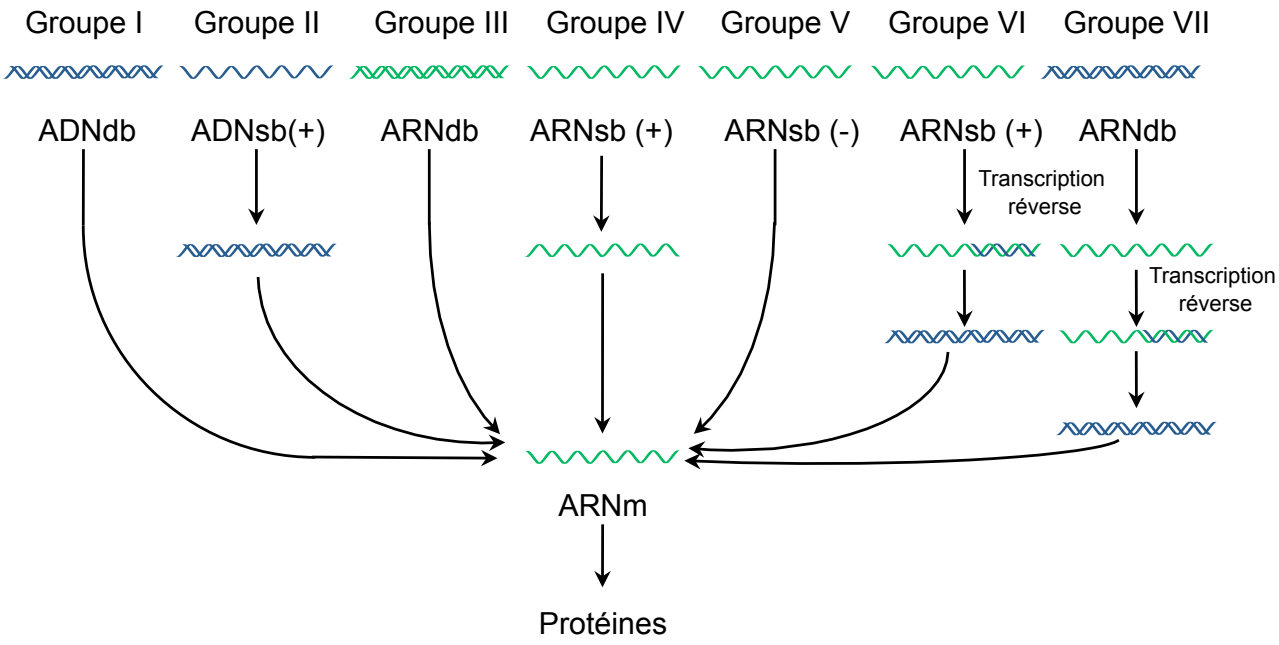


Figure 3. Classification de Baltimore. Les virus sont classés à partir du mécanisme de production de l'ARNm viral à partir du génome initial. ADNdb: double brin d'ADN; ADNsb(+): simple brin d'ADN positif; ARNdb: double brin d'ARN; ARNsb(+): simple brin d'ARN positif; ARNsb(-): simple brin d'ARN négatif; ARNm: ARN messager.

La classe IV réunit les virus à ARN simple brin positif. Leur génome peut directement être traduit mais la réplication est dépendante de RDRP.
La classe V est constituée des virus à ARN simple brin négatif. La production d'ARN simple brin positif, nécessaire à la traduction, est assurée par une RDRP. Cette dernière est aussi nécessaire à la réplication du génome viral.

I.B.1.δ. Les rétrovirus

Les rétrovirus ne peuvent que se répliquer en s'intégrant dans le génome de leurs hôtes. Les virus de classe VI possèdent un génome à ARN simple brin. Une transcriptase réverse permet de synthétiser un complexe ARN/ADN à partir de l'ARN. Ce dernier est ensuite digéré par une RNAse tandis qu'un ADN complémentaire de l'ADN nouvellement synthétisé est produit. L'ADN double brin est intégré au génome de l'hôte à l'aide d'une intégrase produite par le virus.
Le génome des virus de classe VII, peu fréquents et peu étudiés, se compose d'ADN double brin (comme les virus de classe I) mais se réplique en utilisant un intermédiaire ARN. L'ADN double brin forme une structure enroulée qui permet la production d'un ARNm. Cet ARN sert de patron à la production d'ADN double brin par la transcriptase réverse et est également traduit.

I.B.2. Les procaryotes

I.B.2.α. Les archées

Les archées sont des organismes unicellulaires qui ont su coloniser des niches assez peu propices au développement d'autres formes de vie (milieux aqueux à haute salinité, eaux très acides, proche des cheminées de fumée noire). Elles ont d'ailleurs longtemps été considérées comme des bactéries extrêmophiles. On sait maintenant que les archées sont génétiquement et biochimiquement très différentes des bactéries et qu'elles seraient même plus proches des eucaryotes que de ces dernières. Jusqu'à aujourd'hui, un seul cas d'archée symbiotique (sans certitude qu'il s'agisse d'un parasite) a été rapporté (Waters et al., 2003).

I.B.2.β Les bactéries

Les bactéries représentent une partie importante des parasites déjà identifiés. Leur diversité est si importante que sa description exhaustive n'est pas envisageable ici. Ces organismes se multiplient de manière asexuée par fission. Leur faible temps de génération en fait des organismes avec un potentiel évolutif très important. De plus, elles sont à la fois capables d'échanger entre elles des parties de leur génome par conjugaison mais aussi de prélever du matériel génétique du milieu extérieur par transformation et de

l'exploiter. Il existe deux grands types de bactéries, les bactéries à Gram+ et les bactéries à Gram-. Les bactéries à Gram+ possèdent une membrane recouverte d'une couche de peptidoglycanes elle-même recouverte d'une membrane externe chez les bactéries à Gram-. La présence ou non de cette seconde membrane joue un rôle important dans la capacité de l'hôte à reconnaître la bactérie qui l'infecte et à déclencher la réponse appropriée (voir la deuxième partie de cette introduction et notamment la Figure 12).
Les bactéries possèdent des modes de vie très diverses allant de cellule isolée (*Escherichia coli*) à des organisations en biofilm (*Yersinia pestis* responsable de la peste). Certaines peuvent aussi produire des spores très résistantes, leur permettant de survivre dans un environnement défavorable durant des années (*Bacillus anthracis* causant la maladie du charbon). La découverte des antibiotiques a permis une lutte éfficace contre les infections bactériennes. Cependant, l'acquisition par certaines souches de bactéries pathogènes pour l'homme de gènes de résistance aux antibiotiques pose de nombreux problème sanitaire.

I.B.3. Les eucaryotes

I.B.3.α. Les protozoaires

Ce sont des eucaryotes unicellulaires hétérotrophes qui sont pour la plupart libres. Ils se reproduisent principalement par fission et la fréquence de reproduction sexuée, quand il y en a, n'est pas bien définie. Certains de ces protozoaires peuvent former des cystes capables de résister à de hautes températures et des conditions de sécheresse extrêmes pendant plusieurs semaines. Pour les espèces parasites, les cystes forment le principal stade de transmission vers un nouvel hôte. Quatre grands groupes de protozoaires sont définis: les flagellés (Mastigophorae), les amibes (Sarcodinae), les sporozoaires (Sporozoa et Apicomplexa) et les cilliés (Cilliophora). Tout ces groupes contiennent des parasites (Cheng, 1986).

I.B.3.β. Les champignons

Les champignons sont des eucaryotes pluricellulaires hétérotrophes. Les quatre différents phylums de champignons (les Chytridiomycètes, les Zygomycètes, les Ascomycètes et les Basidiomycètes) contiennent des parasites de plantes, d'animaux et parfois d'autres champignons. Un autre groupe, les microsporidies, dont l'appartenance aux champignons reste soumise à discussion, est uniquement composé de parasites intracellulaires obligatoires.

I.B.3.γ. Les métazoaires (animaux)

Les métazoaires sont des eucaryotes pluricellulaires hétérotrophes mobiles. Certains

groupes de métazoaires sont reconnus pour contenir de nombreuses espèces parasites. C'est le cas pour les nématodes (vers ronds non segmentés), les plathelminthes (vers plats) et les annélides.

I.C. Notion de spécificité d'hôte

Il n'existe à l'heure actuelle aucun « super parasite » connu capable d'infecter toute les espèces. Les parasites sont par conséquent spécialistes de certains hôtes. Cette spécialisation peut s'observer à différents niveaux phylogénétiques. Certains parasites sont spécifiques d'une espèce bien particulière (spécialistes) tandis que certains autres sont capables d'infecter un très grand nombre d'espèces (généralistes) (Edwards et Vidrine, 2006). Par exemple, certains acariens sont seulement capable d'infecter une espèce de moule d'eau douce bien précise (Edwards et Vidrine, 2006) tandis que *Pseudomonas aeruginosa*, une bactérie infectieuse opportuniste, est capable d'infecter un plusieurs espèces parfois très éloignées les unes des autres telles qu'*Arabidopsis thaliana*, *Caenorhabditis elegans* et reste encore un problème majeur en santé humaine (Cao et al., 2001; Kominos et al., 1972). Le degré de spécificité d'hôte est une caractéristique importante d'un parasite. Un généraliste aura un éventail d'hôtes potentiels beaucoup plus large qu'un spécialiste et par conséquent généralement une capacité de dispersion plus importante. A l'inverse, les spécialistes sont capables de mieux exploiter leur hôte de prédilection, et ainsi d'en tirer meilleur profit que n'importe quel généraliste sur un de ses nombreux hôtes (Poulin, 1998; Garamszegi, 2006). Il apparaît important de connaître le degré de spécificité d'un parasite pour ses hôtes afin de comprendre la dynamique écologique et évolutive des interactions hôte-parasite. Prévoir la capacité d'une espèce de parasite à sauter d'une espèce d'hôte à une autre peut être déterminant dans la gestion des épidémies ou de l'émergence de maladies nouvelles (Daszak et al., 2000).
Pour estimer la spécificité d'interaction, deux approches complémentaires sont possibles. Premièrement, lors de l'échantillonnage de l'espèce hôte, il suffit de relever les parasites présents dans l'hôte soit par observation, pour les parasites visibles, soit en utilisant des marqueurs moléculaires. Il est également possible d'isoler un parasite donné et d'infecter expérimentalement d'autres espèces à disposition, sans tenir compte de la probabilité de rencontre des deux organismes (Figure 4B). Ces deux approches permettent de distinguer les infections réalisables (biologiquement possibles) des infections réalisées (naturellement observées) (Perlman et al., 2003). Pour qu'une infection ait naturellement lieu, il faut à la fois que le parasite soit physiologiquement capable d'infecter l'hôte mais aussi bien évidemment que l'hôte soit présent dans le même environnement et la même zone géographique que le parasite (Figure 4A). Ainsi, on observe généralement une

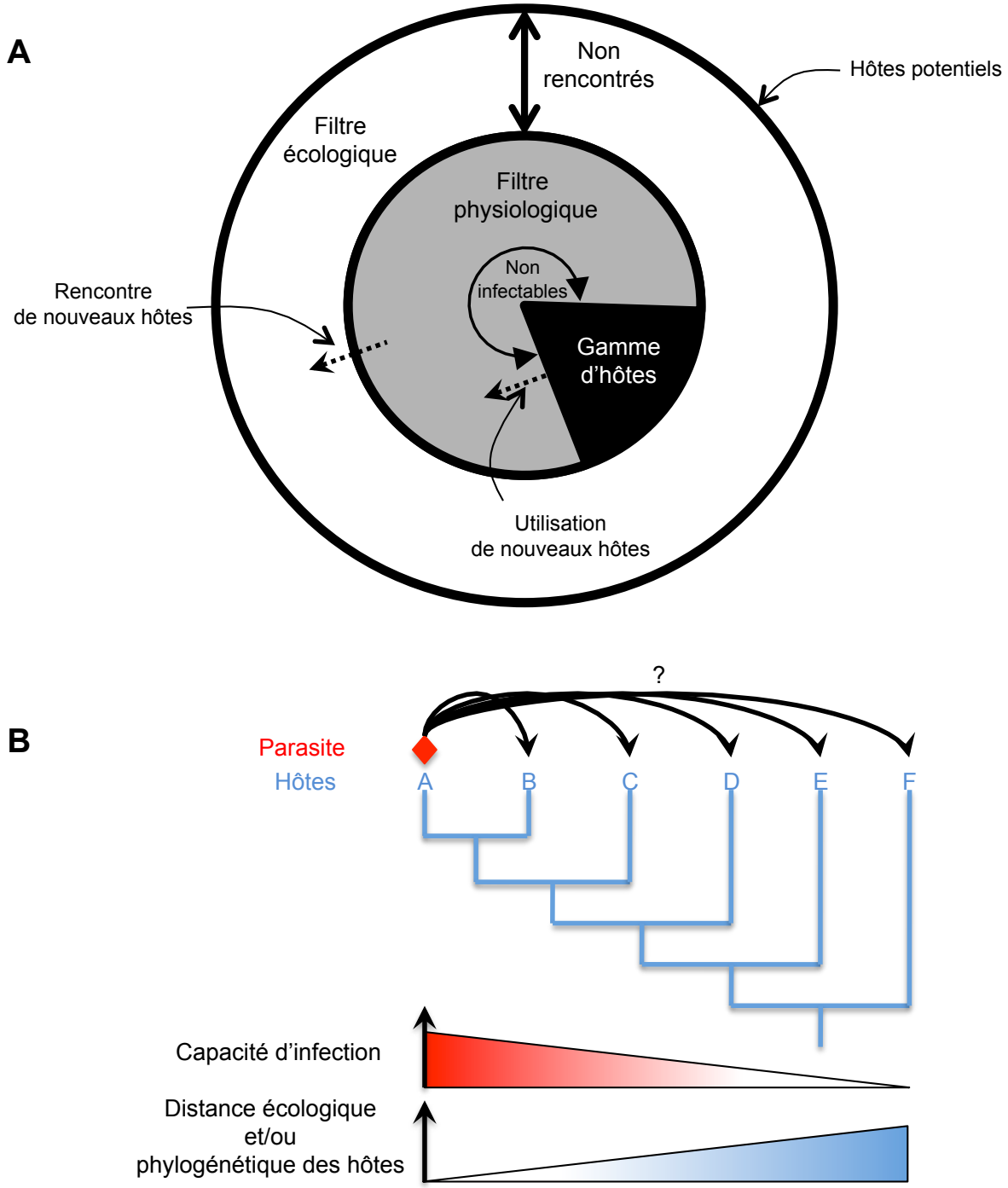


Figure 4. Facteurs affectant la gamme d'hôte d'un parasite. A: Les filtres écologiques et physiologiques déterminent la gamme d'hôtes. Le cercle externe représente le réservoir d'hôtes potentiels. L'espace entre le cercle externe et le cercle interne représente le filtre écologique, i.e. les hôtes jamais rencontrés par le parasite. Le disque interne contient les hôtes rencontrés par le parasite. Le parasite est physiologiquement apte à infecter certains de ces hôtes. Des changements évolutifs peuvent avoir lieu et ainsi étendre ou restreindre la gamme d'hôte (flèches en pointillés). Adapté de Combes (2001). **B:** La capacité d'infection d'un parasite décroît généralement avec la distance écologique ou génétique entre l'hôte testé et l'hôte original.

corrélation négative entre la capacité d'un parasite à infecter un organisme et la distance génétique ou écologique séparant cet organisme de l'hôte original (Combes, 2001) (Figure 4B). Bien évidemment, ces barrières ne sont pas fixes dans le temps et certains parasites peuvent acquérir ou perdre la capacité d'infecter un hôte donné au cours de leur évolution.

La spécificité émerge d'une longue coévolution entre l'hôte et le parasite. L'hôte représente l'environnement immédiat du parasite et le parasite représente l'environnement immédiat de l'hôte lorsque ce dernier est infecté. Les deux organismes exercent une pression de sélection si forte l'un sur l'autre que l'évolution de chacun des deux antagonistes est essentielle à sa survie. On parle alors de coévolution antagoniste entre l'hôte et le parasite. La coévolution peut s'observer à l'échelle phylogénétique lorsque l'on étudie les phénomènes de co-spéciation. Si co-spéciation il y a, on s'attend alors à ce que les ancêtres communs de deux parasites actuels soient eux même des parasites des ancêtres communs des hôtes actuels (Figure 5). Ainsi, la règle de Fahrenholz prévoit que la phylogénie des parasites est parfaitement congruente avec la phylogénie des hôtes (Klassen, 1992). Ce cas de figure est parfois rencontré mais cette règle est utilisée en générale comme une hypothèse nulle pour tester la coévolution entre un groupe de parasites et un groupe d'hôtes. Les sauts d'hôtes ou les non-spéciations (Figure 6) de parasites sont très fréquents et il est rare de trouver une correspondance parfaite entre les phylogénies.

I.D. Cycle de vie et traits d'histoire de vie des parasites

Comme tous les organismes, les parasites nécessitent des nutriments pour croître et se reproduire. Pour les parasites, ces conditions d'acquisition de ces nutriments sont particulières puisqu'elles nécessitent l'envahissement d'un autre organisme. Nous allons essayer de décortiquer ici les étapes clées d'un cycle de vie typique de parasite (Figure 7).

I.D.1. Dispersion

La première étape est bien évidemment la recherche d'un hôte potentiel qui peut se trouver à une distance importante comparée à la taille du parasite. Les parasites ont donc développé de nombreux mécanismes évolutifs permettant de pallier ce problème.

I.D.1.α. La production de nombreux descendants

Un de ces mécanismes est la production de nombreux descendants, augmentant ainsi la

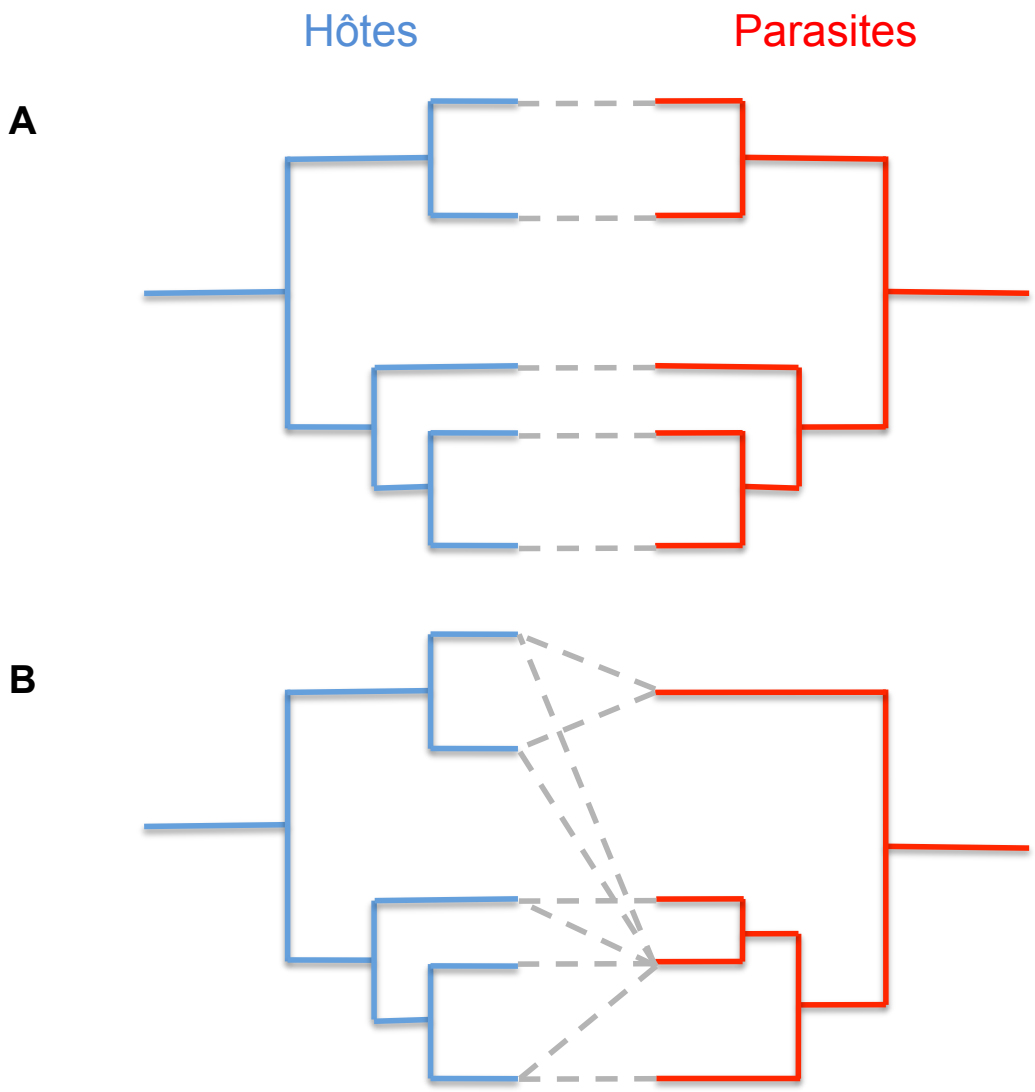


Figure 5. Patrons d'évolution des hôtes et des parasites associés à l'échelle macroévolutive.
A: Illustration de la règle de Farenholz. La phylogénie des hôtes et celle des parasites sont congruentes. Chaque évènement de spéciation des hôtes est suivi par une spéciation du parasite. **B:** La phylogénie des parasites n'est pas congruents avec celle des hôtes. Des spéciations indépendantes ou des changements d'hôtes peuvent avoir eu lieu.

probabilité d'une seule réinfection (la majorité des descendants ne fonderont pas de nouvelles infections). La production de nombreux descendants sur le lieu même de l'infection n'est généralement pas suffisante et des systèmes de dispersion variés ont été développés.

I.D.1.β. La dispersion passive

Les stades de transmission parasitaires sont le plus souvent des formes quiescentes et très résistantes à de nombreuses conditions environnementales. En conséquence, elles ne sont que très peu actives et le plus souvent immobiles. Certains parasites se transmettent par l'air ou l'eau, deux éléments généralement essentiels à leurs hôtes. Le Rhinovirus de l'homme, provoquant le rhume, se transmet tout simplement par l'air lors des éternuements (Goldmann, 2000). *Vibrio cholerae*, induisant le cholera chez son hôte, est transmise par l'eau par l'intermédiaire des matières fécales de son hôte (Sharma et al., 2003).

I.D.1.γ. La dispersion active

Certains parasites sont mobiles et capables de trouver un hôte de manière active. C'est en générale le cas des macroparasites comme les guêpes parasitoïdes, les tiques ou les puces. Certains parasites ont évolué la capacité de modifier le comportement de leurs hôtes, maximisant ainsi la probabilité de rencontre avec un nouvel hôte.

I.D.2. Infection et établissement dans l'hôte

Une fois l'hôte identifié, le parasite doit pouvoir entrer dans celui-ci, dans le cas d'endoparasites, ou s'ancrer à celui-ci dans le cas d'exoparasites. L'infection des animaux se fait très communément par des points d'ouverture de l'animal sur le milieu extérieur: l'appareil digestif, l'appareil respiratoire ou l'appareil reproducteur mais aussi par la peau et parfois les yeux.

I.D.3. Exploitation de l'hôte, croissance et reproduction (endoparasites)

Une fois le parasite à l'intérieur de l'hôte, l'exploitation de celui-ci se met en place, que ce soit par prélèvement de nutriments ou par utilisation de la machinerie cellulaire de l'hôte. Cette exploitation permet la croissance et la multiplication du parasite. Ces étapes peuvent avoir des effets nocifs sur la valeur sélective de l'hôte. La diminution de la fitness de l'hôte, la virulence, se mesure généralement par l'observation de l'effet d'une infection sur la survie et la fécondité de l'hôte infecté. La virulence est une mesure récurrente des études portant sur les interactions hôte-parasite. Elle permet de mesurer

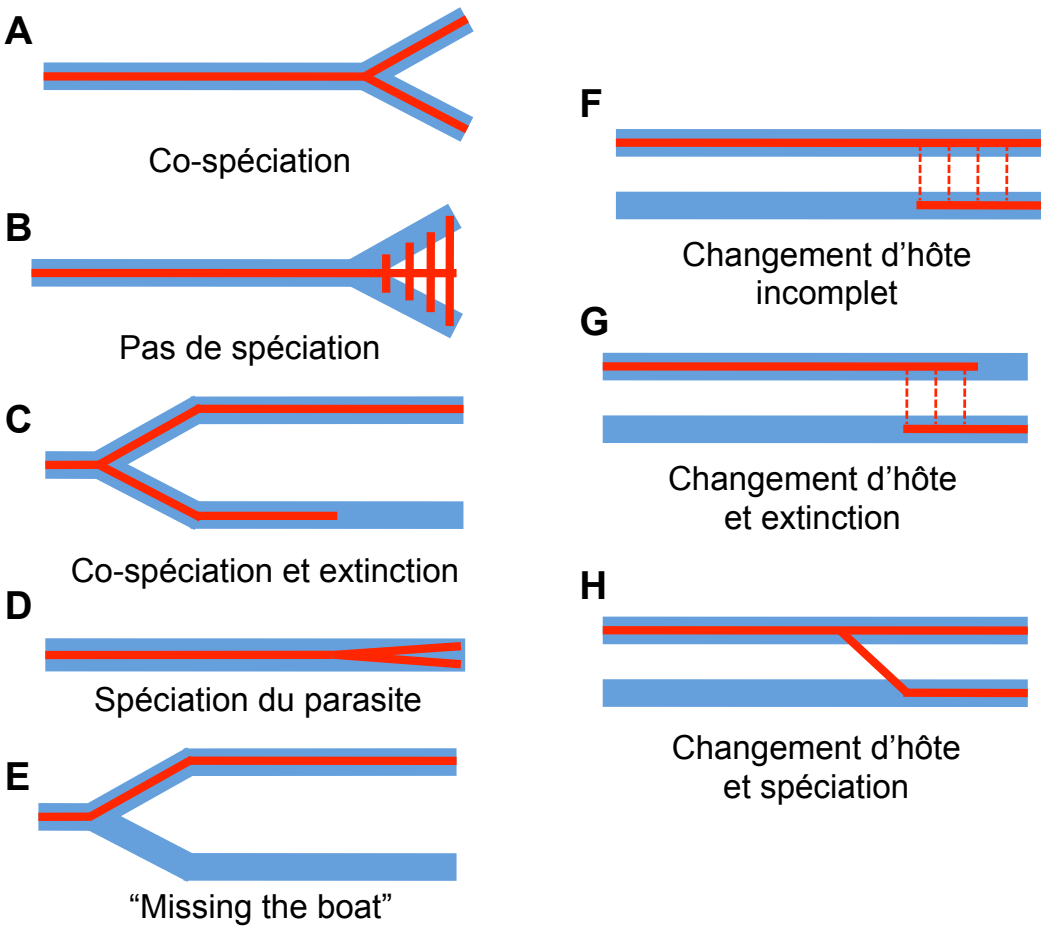


Figure 6. Différents scénarios permettent d’expliquer les écarts à la congruence des phylogénies. **A:** Cas de co-spéciation classique. La spéciation de l’hôte est suivie par une spéciation du parasite. **B:** Dans certains cas, la spéciation du parasite n’a pas lieu et une espèce de parasite peut infecter plusieurs espèces d’hôtes possédant un même ancêtre commun. **C:** La co-spéciation a lieu mais une des deux espèces de parasite s’éteint. **D:** Une spéciation parasitaire a lieu indépendamment d’une spéciation de l’hôte. **E:** Une spéciation au niveau des hôtes n’entraîne pas une spéciation parasitaire. **F:** Le parasite peut également acquérir de nouveaux hôtes et maintenir des échanges génétiques avec le parasite de l’hôte initial. **G:** Ce parasite peut également perdre la capacité d’infecter son hôte initiale et ne conserver uniquement que le nouvel hôte. **H:** L’acquisition d’un nouvel hôte par un parasite peut être suivie par une spéciation de ce dernier, donnant naissance à une nouvelle espèce de parasite.

très indirectement la capacité d'un parasite à exploiter un hôte donné. Suite à sa reproduction, la population parasitaire sort de l'hôte (sortie active ou passive) et doit trouver un nouvel hôte pour se reproduire bouclant ainsi la boucle.

Le succès d'une infection dépend en grande partie de la capacité du parasite à franchir une succession de mécanismes de défense de l'hôte.

I.E. Séquence de défense de l'hôte

La séquence de défense de l'hôte dépend du type de parasite étudié. Par exemple, alors que les endoparasites doivent impérativement pénétrer dans leurs hôtes, les ectoparasites ont moins de barrières à franchir. Dans les deux cas, on peut établir deux principales étapes. Une séquence de défense antérieure à l'infection et une séquence de défense postérieure à l'infection (i.e. postérieure à l'établissement du parasite sur/dans son substrat) (Figure 8). La plupart des étapes de la séquence antérieure à l'infection sont en général négligées par les études portant sur l'immunité des hôtes alors qu'elles jouent un rôle parfois plus important.

I.E.1. Défenses antérieures à l'infection

I.E.1.α. Evitement des parasites

Une infection n'aura jamais lieu si le parasite ne rencontre jamais d'hôte. Une diminution de la probabilité de rencontre du parasite est donc une défense efficace contre les infections.

i. Evitement spatial et temporel

Chez certains hôtes, on peut observer des modifications comportementales lorsque un parasite est présent dans le milieu. Chez l'épinoche, la présence du parasite *Argulus canadensis* dans le milieu provoque un évitement des bas-fonds et de la végétation par les poissons alors que ces milieux représentent des cachettes efficaces contre les prédateurs (Poulin et Fitzgerald, 1989). La fourmi *Solenopsis richteri* restreint son activité de prédation à des périodes de la journée où ses parasitoïdes (*Pseudacteon phoridae*) sont le moins actifs (Folgarait et Gilbert, 1999).

ii. Modification de la consommation de nourriture et du régime alimentaire

Certains organismes sont capables de modifier leur régime alimentaire en fonction de la présence de parasites dans leur nourriture habituelle. Des études ont démontré que

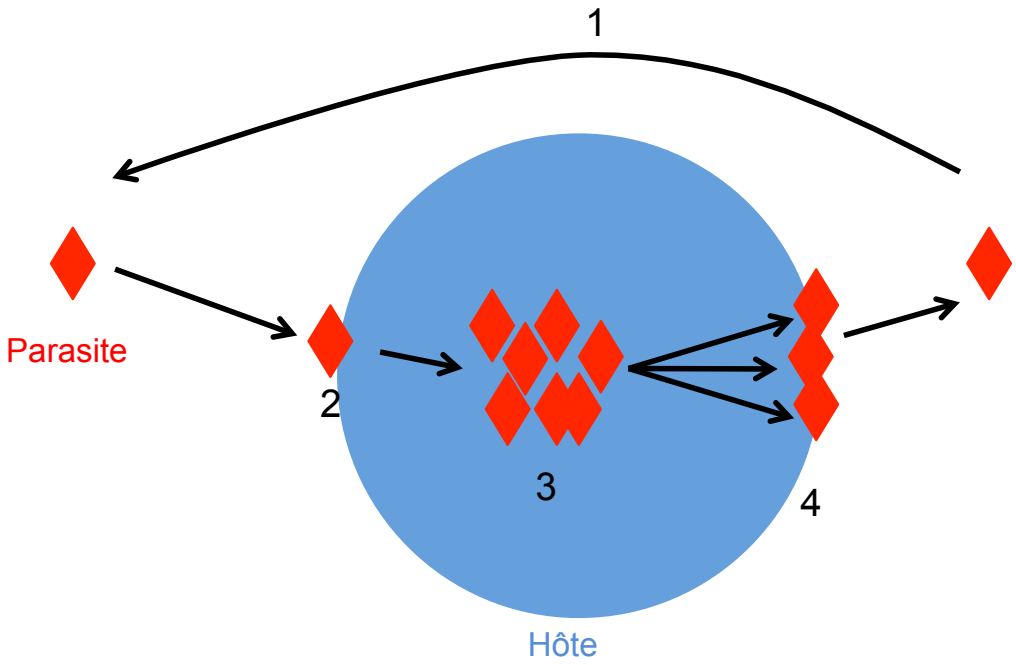


Figure 7. Cycle de vie typique d'un parasite. Dans ce schéma le parasite trouve un hôte (1) soit par recherche active ou par rencontre fortuite. Une fois l'hôte en contact avec le parasite, ce dernier doit pénétrer dans l'hôte (2) pour pouvoir l'exploiter et ainsi se reproduire (3). On considèrera ici que l'étape (3) implique à la fois reproduction clonale et sexuée. Enfin, le stade de transmission du parasite s'évade de son hôte (4) et recommencer le cycle.

certains moutons évitent de se nourrir dans les hautes herbes, pourtant plus nutritives, lorsqu'un parasite y est présent (Hutchings et al., 2001). Chez certains clones de *Daphnia dentifera*, la présence de parasites dans le milieu corrèle négativement avec son taux de prélèvement de nourriture (filtration) dans le milieu (Civitello et al., 2013). Chez le nématode *Caenorhabditis elegans*, la consommation de *Bacillus thuringiensis* toxiques, mesurée par la fréquence de préhension de nourriture, est réduite par rapport à la consommation de souches non toxiques (Schulte et al., 2012). Ce phénomène est accuentué après la une expérience de co-évolution entre ces deux organismes.

iii. Comportements sociaux

Les espèces sociales ont adopté des comportements visant à limiter la transmission des parasites au reste de la colonie. Les abeilles se débarrassent des larves mortes pour prévenir les infections bactériennes de la ruche (Rothenbuhler, 1964). De même, les "déchets" de certains insectes sociaux sont stockés à distance du nid (Weiss, 2006).

I.E.1.β. Réponses anticipées à l'infection

Certaines défenses physiologiques peuvent être déclenchées avant même un contact direct entre l'hôte et le parasite. Les femelles drosophiles sont capables de détecter spécifiquement la présence de guêpes parasitoïdes dans leur environnement. Elles vont alors pondre leurs œufs préférentiellement dans une source de nourriture riche en alcool qui assure une meilleure protection des larves contre ce type de parasite (Kacsoh et al., 2013). Toujours chez les arthropodes, plus particulièrement chez certains lépidoptères, une augmentation de la réponse ProPhenolOxydase (PPO) est observée lorsque l'on augmente la densité d'hôtes en présence de parasites, et ce avant que l'infection n'ait lieu (Wilson et al., 2003).

I.E.1.γ. Barrières physiques à l'infection

Une fois la rencontre entre l'hôte et le parasite effectuée, le parasite doit s'installer sur ou dans son hôte. Pour ce faire, il doit se frayer un chemin vers l'intérieur de l'organisme. De nombreux organismes possèdent des barrières physiques empêchant la pénétration des parasites. La plupart du temps, ces barrières ne sont pas seulement dédiées à la résistance aux parasites mais elles constituent une étape limitant l'infection. Les Nématodes et les Arthropodes possèdent une cuticule qui les protège contre l'infection de multiples parasites en empêchant une interaction directe avec les cellules de l'hôte. Chez l'Homme, la peau est une protection non-négligeable contre les parasites. La sécrétion de mucus, en général au niveau des points d'entrée de l'organisme, permet aussi de réduire la possibilité de pénétration de corps étrangers dans l'hôte.
Une fois ces barrières franchies le parasite va pouvoir s'installer sur/dans son hôte. Cette

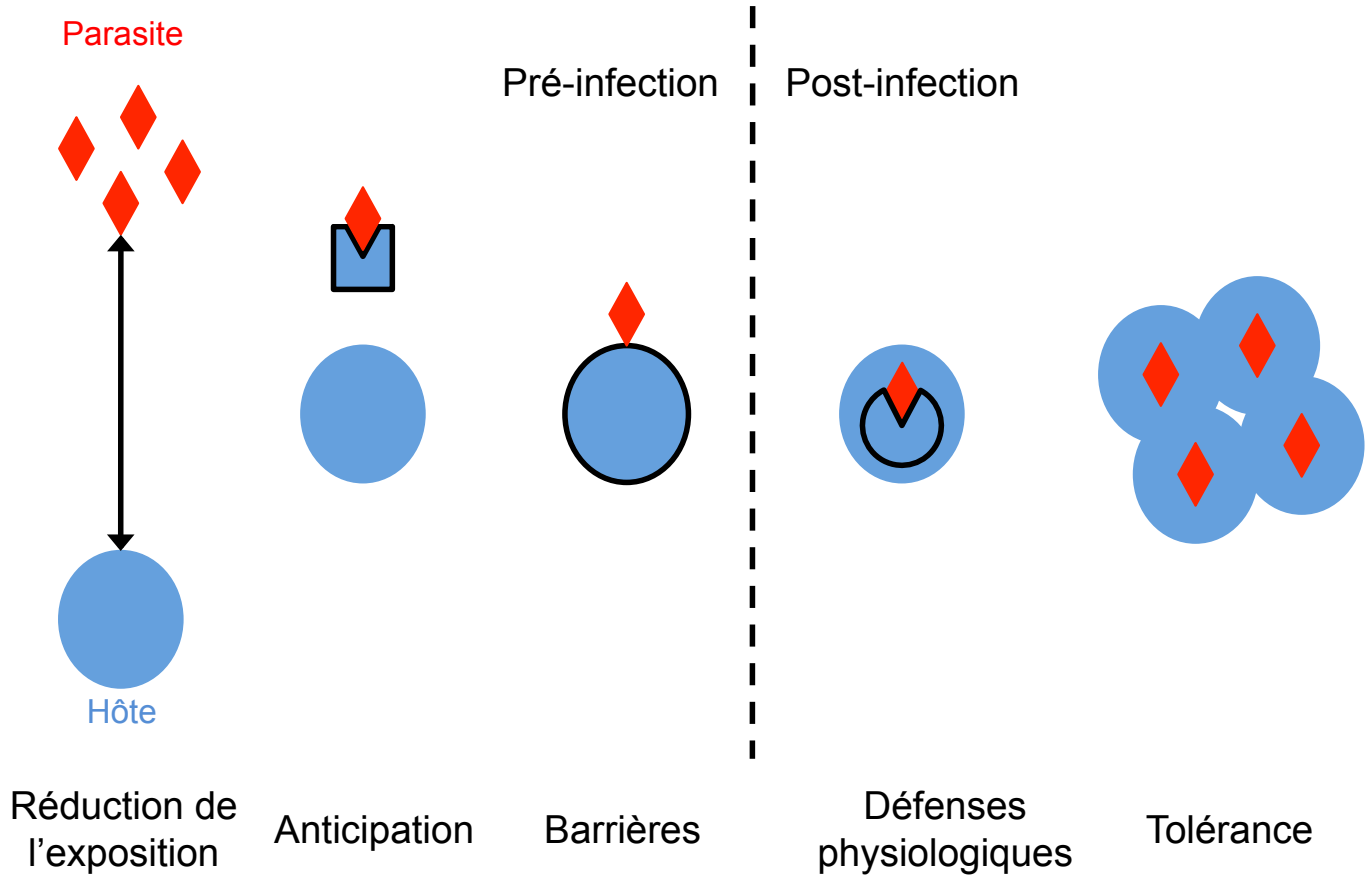


Figure 8. Diversité des mécanismes de défense de l'hôte face à un parasite. Les défenses antérieures à l'infection incluent une réduction de l'exposition au parasite par l'hôte (changements comportementaux, modifications du régime alimentaire), une anticipation de l'infection (surexpression de gènes de l'immunité, modification des conditions physico-chimiques du milieu) et la présence de barrières (peau, cuticule, mucus). Les réponses postérieures à l'infection incluent les défenses physiologiques (système immunitaire) et la réduction des effets nocifs de l'infection (tolérance).

installation est réversible et dépend de la capacité du parasite à faire face aux défenses de l'hôte postérieures à l'infection.

I.E.2. Défenses postérieures à l'infection

I.E.2.α. Réponse immunitaire ou défense physiologique

La réponse immunitaire est le type de défense antiparasitaire le plus étudié et par conséquent le plus documenté, plus particulièrement chez les vertébrés (souris, homme). Cette réponse est en général induite par la détection d'un corps étranger en surface où à l'intérieur de l'hôte. Les nombreux et diverses mécanismes de cette défense seront abordés plus loin dans ce manuscrit. Le système immunitaire, tout comme les autres systèmes biologiques (respiratoire, digestif, nerveux, reproducteur...) d'un organisme, implique la coordination de nombreux éléments de reconnaissance et de signalisation. Pour qu'une réponse immunitaire ait lieu, le système immunitaire doit dans un premier temps être capable de distinguer le soi du non-soi et ensuite d'induire une réponse appropriée à l'infection rencontrée. Certaines réponses ont un large éventail d'action et peuvent cibler différents types de parasite tandis que d'autres sont spécifiques de parasites particuliers. Tous les organismes vivants possèdent un système immunitaire, le plus rudimentaire soit-il. Ces réponses immunitaires sont le résultat d'une longue coévolution entre les parasites et leurs hôtes.

I.E.2.β. Tolérance

La sélection s'exerçant sur l'hôte se mesure principalement par la capacité de l'hôte à survivre et à se reproduire en présence du parasite. Une alternative à l'élimination du parasite est la diminution de l'effet du parasite sur l'hôte: la tolérance. Deux paramètres sont nécessaires pour mesurer la tolérance d'un hôte: la virulence de l'infection qui rend directement compte de l'effet du parasite sur l'hôte et la charge parasitaire qui sert à pondérer cet effet. Pour deux infections parasitaires distinctes, si la charge parasitaire est identique pour les deux infections l'hôte le moins affecté montre alors une meilleure tolérance (Kause et Ødegård, 2012). Par exemple, chez l'Homme, les individus possédant les allèles responsables de l'alpha-thalassémie meurent moins fréquemment lors d'une infection par la malaria. Cependant, il est maintenant clair que les individus hétérozygotes ou homozygotes pour ces allèles portent une charge parasitaire similaire aux individus non-atteints par l'alpha-thalassémie, suggérant une augmentation de la tolérance au parasite (Wambua et al., 2006). Les bases génétiques de la tolérance n'ont pour le moment pas été clairement identifiées. Certaines souris mutantes présentent une nette diminution de fitness comparativement aux individus non mutés mais avec une

charge parasitaire similaire (Råberg et al., 2007). Une des grandes interrogations concernant la tolérance est de savoir si elle correspond à un état stable (une évolution vers la tolérance de la part de l'hôte) ou si elle résulte du fait que le parasite, grâce à son potentiel adaptatif élevé, réussit à contourner en permanence les défenses de son hôte, tant et si bien que la charge parasitaire reste constante au cours du temps.

I. Conclusion

Un parasite est un organisme symbiotique total ou partiel qui vit aux dépens d'un autre organisme, l'hôte. Un parasite peut exercer une très forte pression de sélection sur son hôte, qui, en retour exerce une très forte pression de sélection sur le parasite.
On estime que la moitié des espèces identifiées jusqu'à ce jour ont adopté un mode de vie parasitaire. Ce mode de vie est apparu de nombreuses fois indépendamment dans plusieurs groupes. Cependant, aucun parasite n'est capable d'infecter tout type d'hôte et il existe une spécificité très claire de certains parasites pour des hôtes particuliers. Cette spécificité est dépendante d'une compatibilité physiologique (infection physiologiquement possible?) entre le parasite et l'hôte mais aussi d'une compatibilité écologique (le parasite a-t-il accès à cet hôte?).
Le succès d'une infection dépend à la fois des capacités du parasite à trouver un hôte, à pénétrer à l'intérieur et à l'exploiter mais aussi à la capacité de l'hôte à se défendre.
Dans ce chapitre, nous n'avons qu'effleuré la diversité des défenses mises en place par les hôtes durant leur coévolution avec les parasites. Dans le chapitre suivant, nous nous attarderons sur la diversité rencontrée au niveau d'une étape cruciale de l'infection: la réponse immunitaire.

II. Diversité et évolution des défenses immunitaires

Dans ce chapitre, nous nous focaliserons uniquement sur les infections par les microparasites et les défenses immunitaires correspondantes.
Le devenir d'une interaction entre un parasite et son hôte dépend des caractéristiques évolutives de chacun. Le parasite évolue généralement beaucoup plus vite que son hôte grâce à un temps de génération plus court et à la production plus importante de descendants. Ainsi, l'apparition d'une mutation conférant un gain de fitness est plus fréquente pour le parasite que pour son hôte. Pour être capable de faire face à la grande diversité de parasites, l'hôte possède un arsenal de moyens de défense regroupés dans ce qu'on appelle le système immunitaire. Son principe est simple: reconnaître les éléments étrangers, informer la cellule, le tissu ou l'organisme de cette présence et monter une réponse appropriée au parasite rencontré. Dans ce chapitre, nous essaierons d'avoir une approche macroévolutive du système immunitaire quand cela est possible. Nous commencerons par une description des différents effecteurs du système immunitaire, principaux acteurs de la destruction des parasites, pour ensuite nous intéresser aux voies de signalisation et aux récepteurs qui régulent la production de ces effecteurs.

II.A. Les effecteurs de l'immunité ou "comment neutraliser les parasites".

Deux types de réponse sont utilisés par les organismes pour neutraliser ou éliminer un parasite : une réponse cellulaire et une réponse humorale, basée sur la sécrétion de composants solubles dans les fluides extracellulaires.

II.A.1. Les différentes réponses cellulaires

On peut considérer deux types de réponse cellulaire: celles effectuées par des cellules spécialisées et celles effectuées par la plupart des cellules de l'organisme de l'hôte, indépendamment de son type cellulaire.

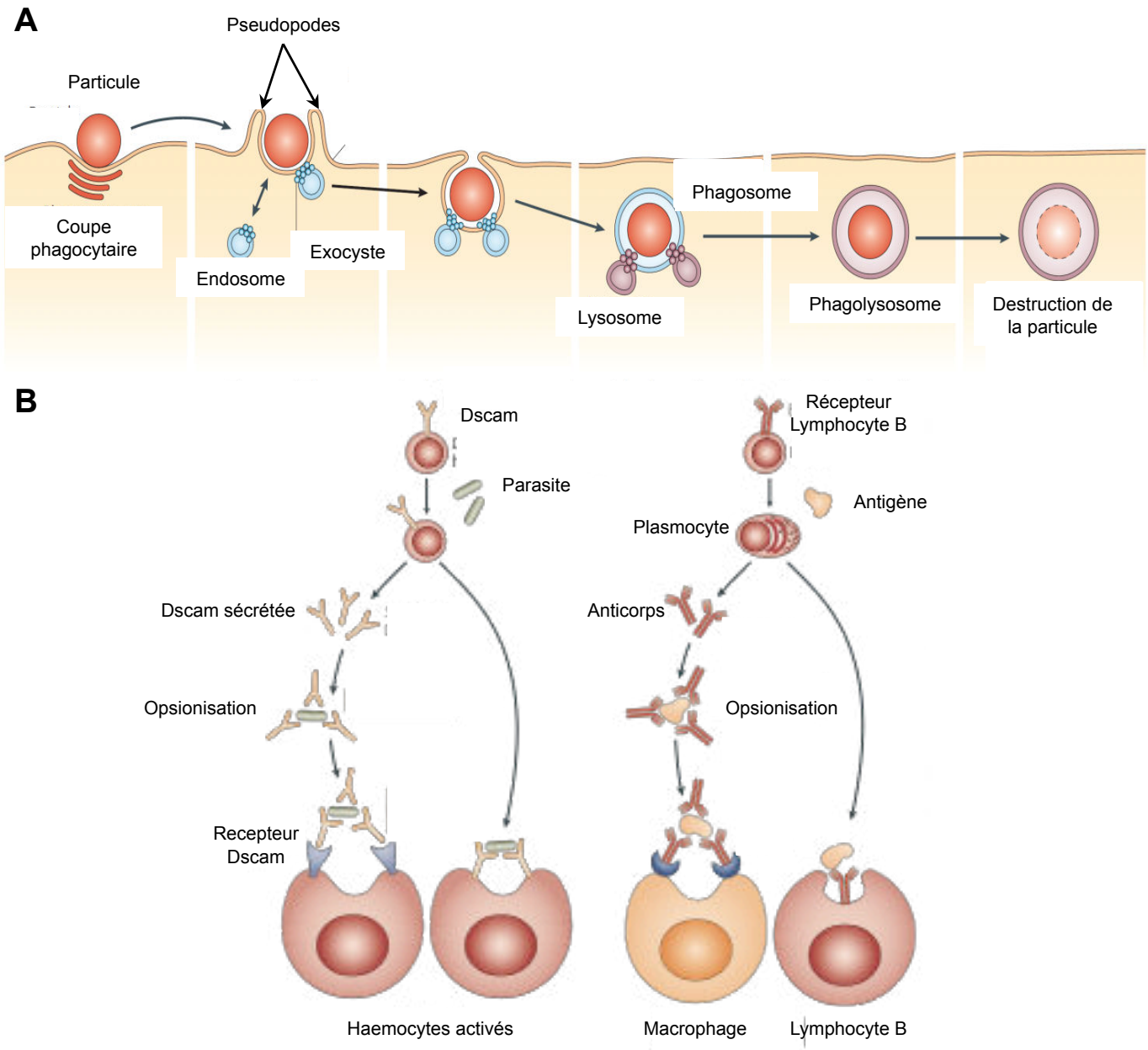


Figure 9. Processus de phagocytose dans l'immunité. A: La reconnaissance par un récepteur d'une particule d'origine étrangère induit la formation de pseudopodes par modification de la structure du cytosquelette au niveau de la membrane plasmique. La particule est internalisée dans une vésicule qui fusionne par la suite avec d'autres vésicules (endosomes et lysozomes) contenant des composants dénaturalisant, formant ainsi le phagolysosome. La particule est détruite dans ce phagolysosome. **B:** La phagocytose peut parfois être très spécifique. Chez les insectes (schéma de gauche), la reconnaissance par les hémocytes phagocytaires se fait par l'intermédiaire d'immunoglobulines solubles très diverses: les Dscam. Ces Dscam se fixent au parasite (opsionisation) et sont reconnues par des récepteurs membranaires de l'hémocyte phagocytaire. La reconnaissance par des Dscam membranaires spécifiques peut directement induire la phagocytose chez ces mêmes cellules. Chez les mammifères (schéma de droite), le même processus a lieu. En lieu et place des Dscam, ce sont les anticorps produits par les lymphocytes B qui servent de récepteurs spécifiques (solubles ou membranaires). La fixation de la partie non variable des anticorps à certains récepteurs de la membrane des macrophages déclenche la phagocytose. Les lymphoctes B possèdent des immunoglobulines membranaires directement capables de reconnaître le parasite. Figure issue de (Stuart and Ezekowitz, 2008).

II.A.1.α. Réponse immunitaire par les cellules spécialisées

Chez les animaux, certains types cellulaires sont spécialisés dans la défense immunitaire de l'organisme. Ces cellules ont été particulièrement bien étudiées chez les vertébrés et les arthropodes. Elles sont généralement mobiles dans les deux types d'organisme : libres chez les arthropodes, dans le système sanguin chez les vertébrés.

i. Mélanisation et encapsulation

La mélanisation est un système de défense spécifique des invertébrés, particulièrement chez les arthropodes, impliqué dans la défense contre les bactéries ou les champignons. La reconnaissance du parasite provoque la cascade de la ProPhenolOxydase (PPO) dans les cellules crystales. Cette cascade induit la production d'intermédaires toxiques comme des phénols, des quinones et des dérivés réactifs de l'oxygène (ROS) qui peuvent tuer directement le parasite (Nappi et Christensen, 2005; Nappi et Ottaviani, 2000; Nappi et Vass, 1993). Elle intervient également dans la synthèse de mélanine par les lamellocytes qui forment une capsule de mélanine autour de l'élément étranger, l'isolant ainsi du reste de l'organisme (Lemaitre et Hoffmann, 2007).

ii. Phagocytose

La phagocytose est une fonction cellulaire caractéristique des cellules animales. L'organisme ciblé par la phagocytose est absorbé par la cellule phagocytaire puis détruit par des organelles spécialisés. Après la reconnaissance du parasite, le cytosquelette des cellules phagocytaires est réarrangé, permettant à la fois une meilleure mobilité de la cellule mais participant également à l'internalisation du parasite par endocytose. Ces mécanismes sont conservés entre arthropodes et vertébrés (Greenberg et Grinstein, 2002; Stuart et Ezekowitz, 2008). La vésicule contenant le parasite, nommée phagosome, peut ensuite fusionner avec d'autres vésicules (endosomes et lysozomes) pour former un phagolysozome dont le contenu promeut la destruction et la dégradation du parasite (Figure 9A). Chez *Drosophila melanogaster,* plus de 600 protéines sont impliquées dans le fonctionnement des phagosomes et 70% possèdent des orthologues chez les mammifères. Chez les vertébrés, la phagocytose déclenche également des signaux stimulant la mise en place de la réponse acquise (paragraphe II.C).

iii. Les lymphocytes B et T

Chez les vertébrés, les cellules présentatrices d'antigènes absorbent les cellules infectées par un parasite intracellulaire, comme par exemple un virus, et présentent à leur surface des fragments de molécules issus d'agents pathogènes, appelées antigènes, via le CMH (Major Histocompatibility Complex) de classe II. Ces antigènes vont permettre la sélection des lymphocytes T (Thymus, lieu de leur production). Ces lymphocytes T

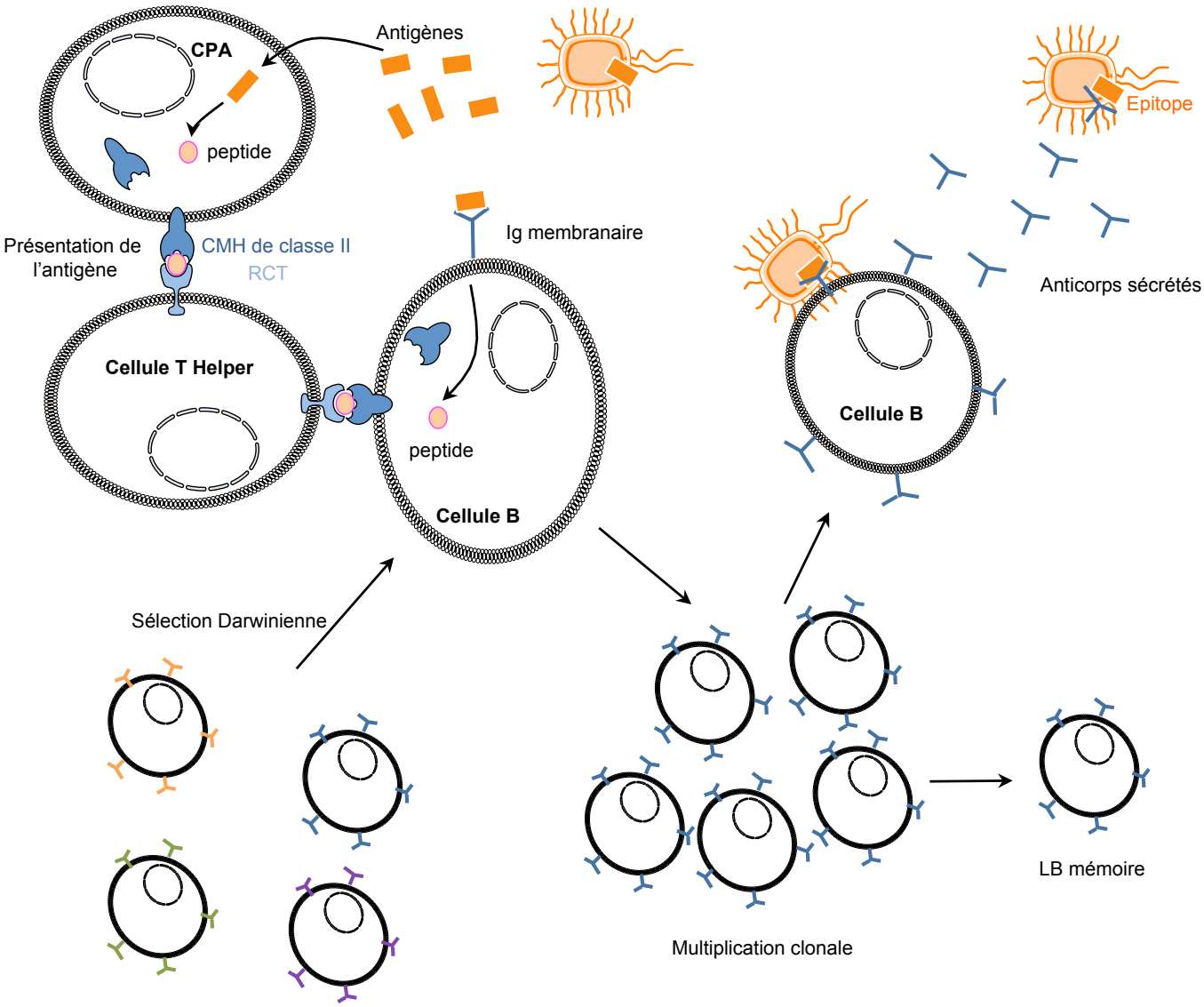


Figure 10. Mécanisme simplifié du système immunitaire acquis contre un parasite extra-cellulaire chez les vertébrés. Une molécule dérivée du parasite (antigène) est reconnue et internalisée par une cellule présentatrice d'antigène (CPA). La molécule est alors dégradée et un peptide est présenté à la surface de la CPA par le Complexe Majeur d'Histocompatibilité (CMH). Si ce complexe est reconnu par un Récepteur de Cellule T (RCT), la cellule T Helper se fixe à la CPA. Les cellules B sont aussi capables d'internaliser un antigène et de présenter un peptide à leur surface par l'intermédiaire du CMH. Si ce peptide est également reconnu par le RCT (s'il s'agit du même peptide que celui présenté par la CPA), la cellule B est activée et se multiplie clonalement. Une partie de ces cellules B est conservée en mémoire, tandis que l'autre partie peut intervenir directement dans la réponse immunitaire en sécrétant des anticorps qui vont neutraliser le parasite porteur de l'antigène.

présentent chacun à leur surface des TCR (T-Cell Receptor) différents dont la partie variable est générée par des réarrangements somatiques via les protéines RAG-1 et RAG-2. Lorsqu'un antigène porté par un CMH (Complexe Majeur d'Histocompatibilité) de classe II correspond à un TCR, alors ce lymphocyte T se divise pour former une population clonale de lymphocyte T spécifique de l'agent infectieux.
N'importe quelle cellule infectée est capable de présenter à sa surface un peptide provenant du parasite par l'intermédiaire d'un récepteur CMH de classe I. Si ce dernier est reconnu par un lymphocyte T, celui-ci se différencie en lymphocyte T cytotoxique et tue la cellule infectée.
La reconnaissance par un lymphocyte T d'un antigène présenté par une cellule présentatrice d'antigène et de ce même antigène présenté par un lymphocyte B provoque la multiplication clonale du lymphocyte B en question et la destruction de ceux qui ne correspondent pas (Rajewsky, 1996). Une partie des lymphocytes B (lymphocytes B mémoire) est mise en mémoire tandis que l'autre partie sécrète des immunoglobulines qui neutralisent l'agent infectieux (Figure 10). Ces immunoglobulines, lorsqu'elles se fixent à un antigène présenté par une cellule infectée par un parasite ou à un antigène présenté par le parasite lui-même, permettent aux lymphocytes NK (Natural Killer) et aux macrophages de reconnaître ces cellules et de les éliminer (Figure 9B) ainsi que le recrutement du système du complément (voie classique), aboutissant à une lyse de la cellule infectée ou du parasite lui-même.

II.A.1.β. Réponse immunitaire par les cellules non spécialisées

i. La défense par les petits ARN

Il semblerait que l'utilisation des petits ARN en tant que moyen de défense contre les virus soit utilisée par les trois domaines du vivant mais par des mécanismes différents.
Chez les eucaryotes, le mécanisme d'exploitation des petits ARN, en général utilisé contre les infections virales, est bien conservé (Ding et Voinnet, 2007) (Figure 11). Une fois que le virus a pénétré le cytoplasme de la cellule, une protéine dicer (DCL pour les végétaux, DCR pour les animaux) reconnaît les ARN double brin viraux et produit des ARN simple brin de 21 nt à 24 nt, appelés petits ARN interférents ou siARN. Chaque siARN primaire est recruté par une protéine argonaute (AGO) (Ding et Voinnet, 2007). Le complexe AGO-siARN se lie à l'ARN viral complémentaire qui est soit directement clivé (chez *D. melanogaster*) par le complexe RISC (RNA-induced silencing complex), soit utilisé comme patron pour la production de siARN secondaires (chez les plantes et chez *C. elegans*). Dans ce dernier cas, les siARN secondaires sont recrutés par d'autres AGO et le complexe RISC ainsi formé reconnaît les ARN viraux complémentaires et les

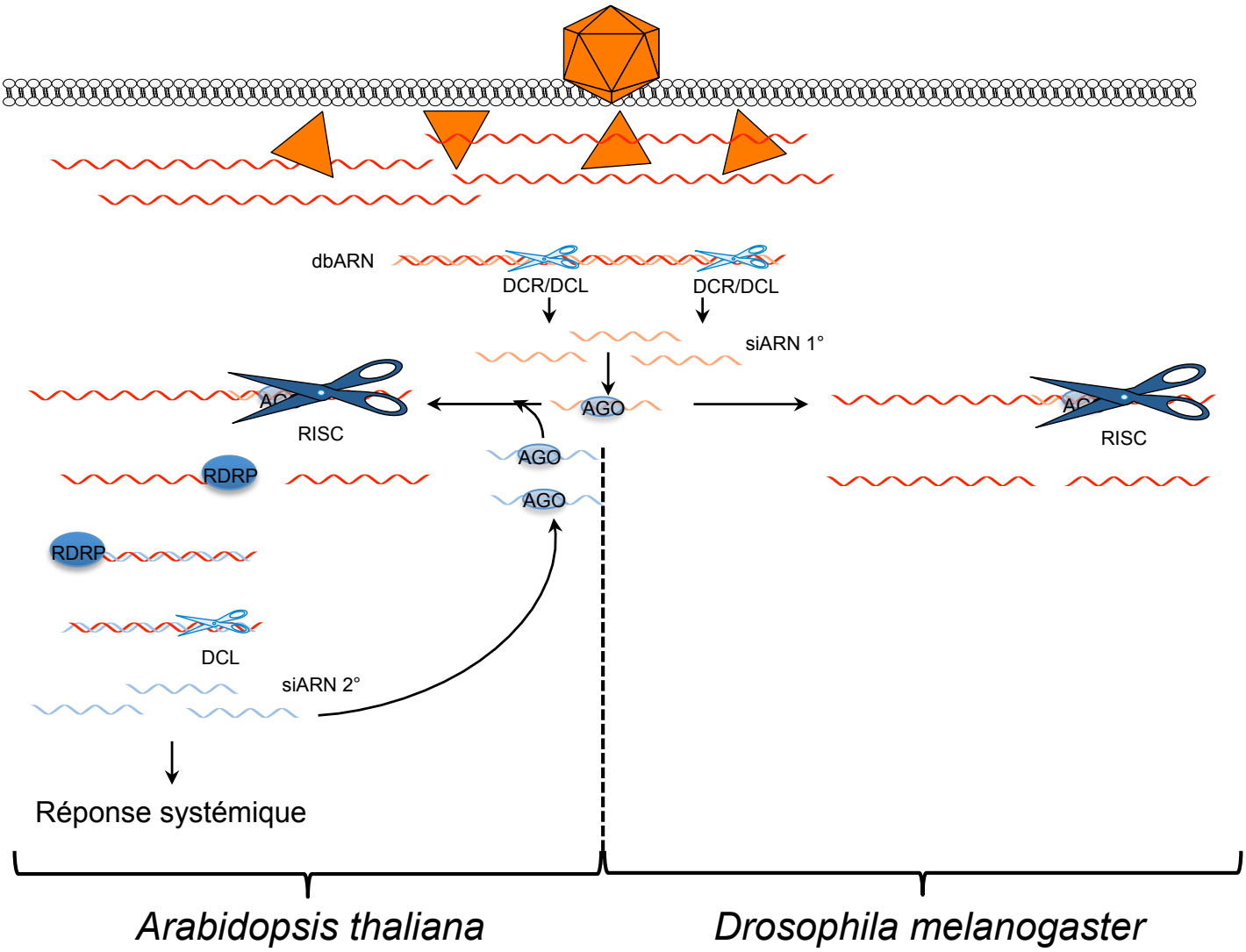


Figure 11. Réponse antivirale par les petits ARN chez *Arabidopsis thaliana* et *Drosophila melanogaster*. La présence d'un ARN double brin (dbARN) est détectée par un complexe protéique incluant une dicer (DCR-2 chez *D. melanogaster,* DCL-2,4 chez *A. thaliana*). Cet ARNdb est clivé par ce même complexe en petits ARN simple brin (siARN 1°) de taille comprise entre 21 et 24 nucléotides. Ces siARN 1° sont pris en charge par des protéines argonautes. Le complexe ainsi formé se fixe à un ARN viral complémentaire du siARN 1° et permet le recrutement du complexe RISC. RISC dégrade alors l'ARN viral. Le processus prend fin à cette étape chez *D. melanogaster.* Chez *A. thaliana*, une RDRP produit un ARN double brin à partir de l'ARN viral clivé. Cet ARNdb est ensuite reconnu par une dicer et permet la production de siARN 2°. Ce phénomène permet une amplification de la réponse des petits ARN. Les siARN 2° sont ensuite exportés vers les cellules voisines par l'intermédiaire des plasmodesmes.

clive (Ding et Voinnet, 2007). Les rares virus à ADN (génome circulaire) se répliquent dans le noyau de la cellule. Les siARN secondaires, en se liant à des AGO spécifiques, sont capables d'empêcher la transcription en recrutant des complexes qui vont méthyler le chromosome viral (mécanisme également impliqué dans la régulation des transposons). Qu'ils soient à ARN ou à ADN, certains virus sont capables de produire des répresseurs viraux de l'extinction (VSR) qui perturbent la voie des siARN comme par exemple la protéine B2 du Flock House Virus (Chao et al., 2005).

Un des mécanismes majeurs conférant une résistance à la présence d'acides nucléiques étrangers (virus ou plasmide bactériens) chez les bactéries et les archées est le système CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated). Ce système est retrouvé chez la moitié des espèces bactériennes connues et dans 90% des archées. Le système est basé sur la présence de clusters de courtes séquences répétées (de 28 nt à 37 nt) espacées par des séquences uniques (spacers) longues de 32 à 38 nt. Ces courtes séquences répétées, découvertes dès 1987 (Ishino et al., 1987), correspondent à des séquences étrangères au génome de son hôte qui ont été plus tardivement identifiées comme étant d'origine virale et plasmidique (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). Lorsqu'un virus infecte une colonie bactérienne, certaines de ces bactéries vont survivre à l'infection. Ces dernières ont intégré au niveau 5' de leur locus CRISPR une courte séquence du génome du phage (protospacer), sélectionnée car proche d'un PAM (Protospacer Adjacent Motif). Après détection d'un élément étranger, le locus CRIPSR est entièrement transcrit et conduit à la formation de crARN (crisper ARN) pour chaque spacer (Brouns et al., 2008) par l'intermédiaire des protéines Cas. Ces crARN sont ensuite pris en charge par un complexe effecteur qui clive l'ARN ou l'ADN d'un agent pathogène envahisseur dont la séquence correspond (Barrangou et al., 2007; Garneau et al., 2010; Brouns et al., 2008; Garneau et al., 2010). La détection d'un PAM est essentiel à la distinction entre séquence étrangère et spacer (soi vs non soi).

Même si le principe de la défense par les petits ARN semble être similaire entre les procaryotes et les eucaryotes, les protéines impliquées dans chacun des deux systèmes sont très différentes, infirmant l'hypothèse d'une origine commune entre les deux systèmes (Haft et al., 2005; Makarova et al., 2006). Il semble cependant que les protéines essentielles au processus de l'ARN interférence chez les eucaryotes aient des origines procaryotes, que ce soit la protéine dicer (MacRae et Doudna, 2007) et les protéines argonautes (Aravind et al., 2000), ou des origines virales pour l'ARN polymérase dépendante de l'ARN (Iyer et al., 2003) présentes chez les plantes et *C. elegans*. Les fonctions de ces protéines ne sont pas conservées entre procaryotes et eucaryotes.

ii. La mort cellulaire programmée

Un autre type de défense est communément adopté par les cellules infectées: la mort cellulaire. Elle consiste simplement au suicide de la cellule et participerait à diminuer la propagation du parasite. Malgré des mécanismes différents, la mort cellulaire programmée se retrouve à la fois chez les animaux (Kerr et al., 1972) et les végétaux (Morel et Dangl, 1997) mais aussi chez les procaryotes (Gross, 1954).

II.A.2. Les différentes réponses humorales

La réponse humorale est définie par l'action de molécules sécrétées dans le milieu extracellulaire. Ces molécules agissent généralement à distance de la cellule qui les a produites.

II.A.2.α. Les peptides antimicrobiens

Les peptides antimicrobiens (AMP) représentent un groupe hétérogène de peptides courts (<100 acides aminés) impliqués dans la défense contre de nombreux parasites (bactéries, champignons, virus, protozoaires). On les retrouve bien évidemment sur les sites des infections, mais également dans les muqueuses intestinales ou encore dans le tract urinaire. Plusieurs centaines d'AMP, appartenant à de nombreux organismes eucaryotes ont été identifiés. Ils sont peu spécifiques et sont donc capables de lutter contre des parasites jamais rencontrés auparavant (Casteels et al., 1993) par l'intermédiaire de mécanismes très variés (Brogden, 2005; Zasloff, 2002). Par exemple, certains sont capables de former des pores dans la membrane tandis que d'autres peuvent traverser cette dernière et cibler des composants à l'intérieur du parasite (Sahl et al., 2005). L'apparition de cette catégorie de molécules est très ancienne puisqu'on les retrouve dans les trois règnes du vivant (Yeaman et Yount, 2007). Malgré une longue co-évolution avec les parasites, ces molécules restent toujours aussi efficaces. Ceci est probablement dû au fait qu'elles ciblent préférentiellement des composants essentiels et inamovibles de la membrane des parasites.

II.A.2.β. Les cascades protéolytiques

Une cascade protéolytique est une dégradation de certaines protéines parasitaires, généralement présentes membranaires, par des protéases induisant la destruction de l'organisme ciblé. Elles sont déclenchées par la fixation d'une protéine à un motif particulier du parasite. La cascade protéolytique la plus étudiée est celle du complément, conservée chez la plupart des animaux (Nonaka et Kimura, 2006; Nonaka et Miyazawa, 2002). Les composants clés du complément se retrouvent chez les deutérostomiens ainsi

que chez les anémones de mer, suggérant une apparition du complément datant d'au moins un milliard d'année (Flajnik, Du Pasquier 2008). La voie du complément peut être également activée par les lectines de type C qui reconnaissent les sucres à la surface des parasites. Les lectines de type C, présentes dans de nombreux phylums, sont très diverses et polymorphes (Flajnik, du pasquier 2008). Certains de ces phylums possèdent des lectines de type C mais pas de compléments, suggérant un autre rôle, encore inconnu, de ces protéines.

II.A.2.γ. Les immunoglobulines sécrétées

Elles représentent un très grand nombre de molécules possédant diverse fonctions lorsqu'elles sont associées à d'autres éléments du système immunitaire. Structurellement, elles sont caractérisées par la présence de domaines immunoglobulines (110 acides aminés). Les familles d'immunoglobulines diffèrent parmi les eumétazoaires.

i. Les anticorps

Ces immunoglobulines, seulement présentes chez les vertébrés à mâchoires, sont des glycoprotéines libérées par les lymphocytes B après l'activation de ces derniers. Elles participent, par exemple, à la reconnaissance d'un parasite par le complément ou par les macrophages, permettant ainsi la destruction de la cellule infectée ou du parasite (Figure 9B). Les anticorps sont constitués de deux chaînes légères et de deux chaînes lourdes reliées entre elles par des ponts disulfures. Chaque chaîne possède une région variable qui détermine la spécificité de fixation à l'antigène (Putnam, 1969). La variabilité de la spécificité est due à plusieurs phénomènes dont les réarrangements somatiques des loci V(D)J par les protéines RAG (Recombinase-activating gene) et les hypermutations somatiques des loci V.

ii. Les Dscam sécrétées

Les Dscam (Down syndrom adhesion molecules) sont des immunoglobulines sécrétées et parfois membranaires retrouvées uniquement chez les arthropodes et possèdent trois domaines très variables dont les ARNm peuvent subir de nombreux épissages alternatifs différents, leur conférant ainsi plus de 18000 isoformes possibles (Watson et al., 2005). Leur rôle reste pour le moment à définir mais on suppose qu'elles facilitent l'absorption des parasites par les cellules phagocytaires (Figure 9B).

iii. Les FREP

De manière similaire aux arthropodes, les mollusques possèdent des hémocytes, pour certains phagocytaires, impliqués dans la réponse immunitaire. Ces cellules sont capables de produire une grande variété de molécules, nommées FREP (fibrinogen related proteins) qui possèdent un domaine terminal appartenant à la superfamille des immunoglobulines

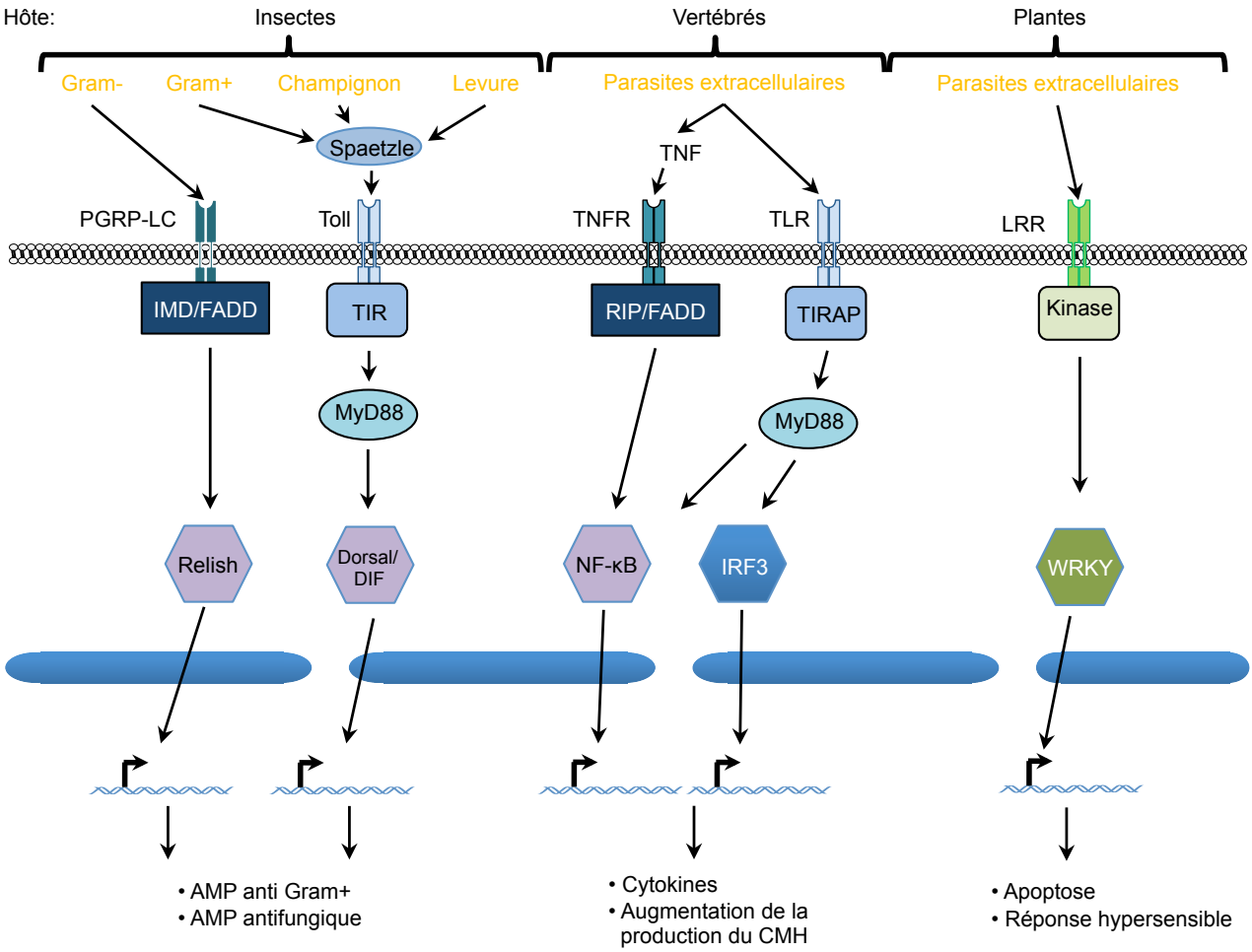


Figure 12. Récepteurs membranaires et voies de signalisations impliquées dans les défenses antiparasitaires. Malgré des ligands différents (Spaetzle pour Toll et molécule du parasite pour TLR), les voies Toll et TLR sont conservées entre les insectes et les mammifères. Le domaine intracellulaire TIR interagit avec la protéine MyD88. Cette dernière déclenche indirectement la voie de libération de facteurs de transcription appartenant à la famille des NF-κB (Dorsal/DIF chez *D. melanogaster*). Les voie Imd (insectes) et TNFR sont également très conservées entre les deux types d'organisme malgré des récepteurs différents. Seulement certains acteurs de la voie de signalisation sont représentés ici mais la plupart des protéines impliquées dans ces voies sont conservées. Les effecteurs produits par ces réponses diffèrent cependant: peptides (AMP) chez *D. melanogaster,* cytokines et production du MHC chez les mammifères. Comme pour les TLR et les Toll chez les mammifères et les insectes, les plantes possèdent aussi des récepteurs membranaires dont la partie exracellulaire est constituée de répétitions riches en leucines. Cependant, la partie intracellulaire est différente et déclenche une cascade de phosphorylation jusqu'à l'induction de la réponse hypersensible par l'intermédaire du facteur de transcription WRKY.
Les éléments de même forme et de même couleur traduisent une conservation des protéines.

(Adema et al., 1997). L'existence de phénomènes d'épissages alternatifs et de mutations ponctuelles somatiques engendre une grande diversité de molécules (Zhang et al., 2004). Même si le mode d'action de ces FREP reste encore incompris, leur expression est accrue lors d'une infection par un parasite, suggérant un rôle de défense contre ces derniers.

La production d'une réponse immunitaire est énergétiquement couteuse et parfois dangereuse pour l'hôte. C'est pourquoi la production de ces molécules doit être régulée par l'intermédiaire de voies de signalisations.

II.B. Reconnaissance du soi et réponses adaptées à l'infection

Comme dit précédemment, une réponse adaptée à une infection dépend bien évidemment de l'efficacité de l'identification du parasite. Les hôtes ont mis en place des récepteurs capables de reconnaître des domaines moléculaires conservés chez les parasites (PAMP pour Pathogen Associated Molecular Pattern) qu'on appelle plus communément les PRR (Pattern Recognition Receptor).

II.B.1. Les Récepteurs riches en répétitions de leucines

Les récepteurs à domaine riche en répétions de leucines (récepteur à LRR), retrouvés principalement chez les plantes et les métazoaires, sont regroupés en six familles selon leur taille et le nombre de motifs constituant la répétition (Ng et al., 2011). Les récepteurs LRR possèdent de 2 à 45 motifs, chacun constitué de 20 à 30 acides aminés formant un arc favorisant l'interaction avec d'autres protéines. Ils peuvent être ancrés dans la membrane ou libre dans le cytoplasme. Leur rôle principal dans le système immunitaire est de reconnaître les antigènes et les cytokines (Akira et Takeda, 2004) et d'induire un signal intracellulaire.

II.B.1.α. La voie Toll: une voie conservée entre les arthropodes et les vertébrés

Les récepteur Toll été initialement identifiés comme acteurs du développement des insectes mais sont aujourd'hui connus pour leur rôle dans la réponse immunitaire chez les invertébrés (Toll) et les vertébrés (TLR pour Toll-like Receptor) (Lemaitre et al., 1996). La plupart des protéines constituant les voies induites par Toll et les TLR sont conservées (Figure 12). Chez *D. melanogaster,* l'activation du récepteur Toll ne se fait pas directement par l'interaction avec un antigène du parasite. C'est le ligand Spaetzle, clivé suite à la détection d'un PAMP, qui se fixe au récepteur Toll (Levashina et al., 1999). Chez les vertébrés, les ligands interagissant directement avec les TLR n'ont pas encore

été identifiés précisément. Une fois activé, Toll et les TLR recrutent, par l'intermédiaire de leur domaine intracellulaire TIR, (Toll Interleukin Receptor), l'adaptateur MyD88 (Pelle chez *D. melanogaster*) (Medzhitov et al., 1998) qui interagit ensuite avec la protéine IRAK (Muzio et al., 1997). TRAFF-6 (Cao et al., 1996) sert ensuite d'adaptateur pour activer la MAPKKK TAK-1 (Ninomiya-Tsuji et al., 1999) qui permet la libération du facteur de transcription NF-κB (Dif/Relish chez *D. melanogaster*) (Irie et al., 2000) et le déclenchement de la production de peptides antimicrobiens chez *D. melanogaster* et celle de cytokines chez les vertébrés. Dix TLR ont été identifiés comme ayant un rôle dans l'immunité chez les vertébrés, tandis que plus de 200 ont été identifiés chez l'oursin (Gilmore et Wolenski, 2012).

II.B.1.β. Les récepteurs extracellulaires riches en leucine chez les plantes

Chez les plantes, la partie extracellulaire des PRR est également riche en répétitions de leucines mais est très différente des domaines LRR des PRR animaux, même si les PAMP reconnus sont des parties similaires du parasite (mais des domaines différents). Ces récepteurs ne possèdent pas de domaines intracellulaire TIR mais initie une voie MAPK aboutissant à la translocation du facteur WRKY (Eulgem et al., 1999) dans le noyau et à l'induction d'une réponse hypersensible (Figure 12).

II.B.1.γ. Les récepteurs cytosoliques riches en répétitions de leucines

Les animaux et les plantes possèdent également des PRR cytoplasmiques nommés récepteur CLR (CATERPILLAR) chez les mammifères et NBS-LRR (Nucleotide Binding Bite - Leuncin Rich Repeat) chez les plantes. Les CLR reconnaissent préférentiellement les produits de dégradation des peptidoglycanes tandis que les NBS-LRR peuvent reconnaître une multitude de PAMP ainsi qu'une modification de certaines protéines de l'hôte provoquée par l'infection (Ausubel, 2005). Ces récepteurs intracellulaires déclenchent des cascades de phosphorylations, la production de ROS, des flux calciques, des activations de facteurs de transcription et l'induction de nombreux gènes impliqués dans l'immunité. Cependant, il n'y a aucune preuve de la conservation des voies de signalisation initiées par ces PRR intracellulaires entre les animaux et les végétaux, suggérant une convergence évolutive de ces différents mécanismes (Ausubel, 2005).

II.B.2. Les voies de signalisations indépendantes de domaines LRR

Même si les domaines LRR semblent avoir été sélectionnés pour leur capacité à interagir avec de nombreuses molécules associées aux parasites, d'autres voies de signalisations parfois très conservées sont impliquées dans les réponses immunitaires.

II.B.2.α. Les voies Tumor Necrosis Factor / Imd

Les voies Tumor necrosis factor (TNF), chez les vertébrés, et Immuno deficiency (Imd), chez la drosophile, ont été découvertes indépendamment et la grande similitude entre les deux voies n'a été mise en évidence que plus tardivement (Leulier et al., 2002). Elles induisent la production de peptides antimicrobiens spécifiques des bactéries à Gram- chez *D. melanogaster* et celle de cytokines impliquées dans le recrutement de la réponse cellulaire chez les mammifères (Figure 12). Malgré une grande similitude des acteurs intracellulaires de ces voies de signalisation, les complexes ligand-récepteur initiateurs sont très différents. La voie TNF des vertébrés est activée par la fixation de la cytokine TNFα sur le récepteur TNF (TNFR) tandis que, chez les insectes, la voie Imd est directement activée par la fixation d'une molécule issue d'une bactérie à Gram- à une protéine de reconnaissance des peptidoglycanes (PGRP). Une fois activés, ces récepteurs recrutent des protéines impliquées dans la voie de signalisation de la mort cellulaire par l'intermédiaire de "Death domain" (Georgel et al., 2001) libérant le facteur de transcription RELISH (NF-κB).

II.B.2.β. La voie Jak-STAT

La voie Jak-STAT est depuis longtemps connue pour réguler l'expression des peptides antimicrobiens lors d'infections bactériennes et virales chez les animaux. Cette voie est activée par des récepteurs qui reconnaissent des molécules liées aux agents infectieux (PRR, Pattern Recognition Receptor). L'activation paracrine de cette voie de signalisation a lieu lorsqu'un PRR détecte la présence du ligand vago, caractéristique d'un stress cellulaire dû à une infection virale chez la drosophile (Deddouche et al., 2008) ou la présence de cytokines (Interférons) produites par une autre cellule du système immunitaire chez les mammifères (Darnell et al., 1994). La Janus Kinase (Jak) Hopscotch est transphosphorylée, menant à la dimérisation de STAT permettant son transport dans le noyau. STAT agit alors en tant que facteur de transcription et régule l'expression de gènes effecteurs dont la fonction est pour le moment mal connue.
La voie des siARN et celle de Jak-STAT sont liées puisque DCR-2 semble induire la production du ligand vago (Deddouche et al., 2008) lors d'une infection virale chez la drosophile.

II.B.2.γ. Les récepteurs membranaires des lymphocytes: les immunoglobulines membranaires

Chez les agnathes, les lymphocytes possèdent des récepteurs membranaires des TCR (T-Cell Receptor) et des BCR (B-Cell receptor) impliqués dans la reconnaissance et dans la neutralisation des parasites extracellulaires (Figure 9B et 10). Ils possèdent une structure variable synthétisée de la même manière que les anticorps, un domaine transmembranaire

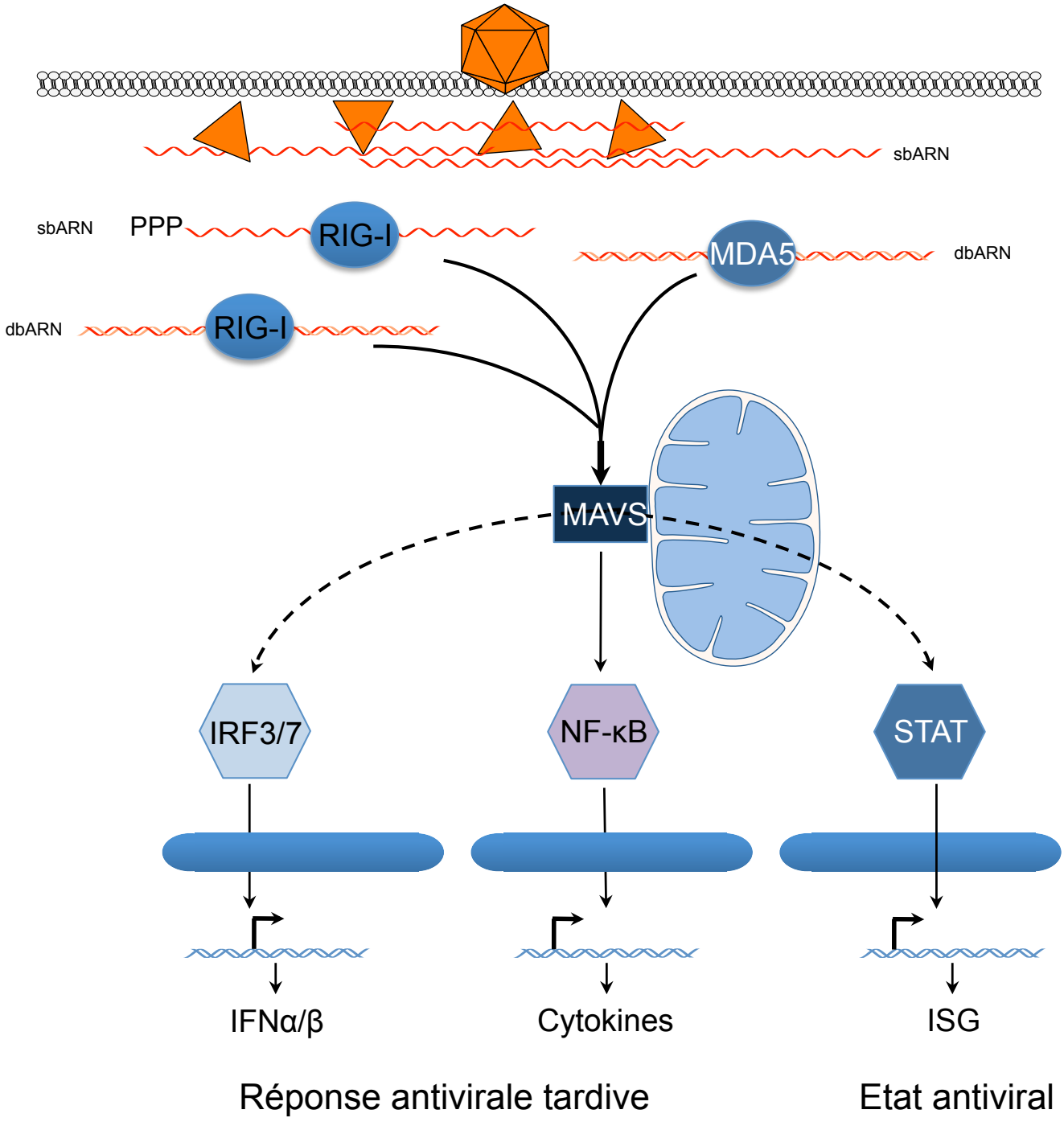


Figure 13. Récepteurs cytoplasmiques associée à une infection virale chez les mammifères. La présence d'ARN ou d'ADN viral active les RLR (RIG-I_Like Receptor) RIG-I et MDA5. Cette activation induit un changement de conformation du récepteur et révèle le domaine CARD (Caspase Activation and Recruitment Domain). Ce domaine peut alors se fixer au domaine CARD des protéines MAVS (Mitochondria AntiViral Signaling protein), qui sont des récepteurs transmembranaires mitochondriaux. Cette interaction induit des voies de signalisation qui activent l'expression d'interferons de type I (IFNα/β), de cytokines et de genes stimulés par interférons (ISG) par les facteurs de transcription INF3/7, NF-κB et STAT, respectivement.

et un domaine intracellulaire. Une fois activés par la reconnaissance d'un antigène, ces récepteurs induisent la transcription de gènes effecteurs par l'intermédiaire du facteur de transcription NF-κB.

II.B.2.δ. Des récepteurs cytosoliques d'acides nucléiques: les RIG-I-Like Receptors

Les RLR (Retinoic acid induced gene-Like Receptor), RIG-I et MDA-5 sont des récepteurs cytosoliques activés par la présence de double brin d'ARN viraux chez les agnathes. L'ensemble de ces récepteurs induisent plusieurs voies de signalisation (Jak-STAT; NF-κB; IRF) par l'intermédiaire de leur domaine CARD (Caspase activation et recruitment domain) (Schlee, 2013) aboutissant à la production d'interférons de type I et III et de cytokines (Figure 13). Une fois sécrétées, ces dernières recrutent sur le lieu de l'infection d'autres cellules impliquées dans la réponse immunitaire innée et activent la réponse immunitaire acquise (Lymphocytes B et T). Le domaine RIG-I, capable de fixer les ARN viraux est généralement conservé parmi les eumétazoaires mais est absent chez les plantes (Zou et al., 2009).

II.C. Une immunité innée et une immunité acquise pour tous

L'immunité est classiquement décomposée en deux types: l'immunité innée et l'immunité acquise (pour éviter les confusions avec la définition évolutive du terme adaptatif, nous n'emploierons pas ici le terme d'immunité adaptative). L'immunité innée se caractérise par une réponse rapide et non spécifique à l'infection. Les molécules effectrices produites sont efficaces contre un type donné de parasite (virus, bactérie à Gram+ et à Gram-, champignon ou protozoaire) mais pas contre une souche donnée. A l'inverse, la réponse acquise, est plus lente à mettre en place mais bien plus spécifique. Elle est basée sur le fait qu'une infection non létale par un parasite donné réduit l'impact des infections ultérieures par ce même parasite, suggérant également une mise en mémoire de la spécificité. Comme l'immunité acquise est considérée comme exclusif des vertébrés (lymphocytes B et T et pseudo-lymphocytes), il est admis aujourd'hui que le système immunitaire acquis est une innovation qui est venue "se greffer" sur le système immunitaire inné (Flajnik et Du Pasquier, 2004).
Cette vision "vertébré-centrée" de l'immunité suggère que seuls les vertébrés sont constamment sollicités par une grande diversité de parasites. Il est certain que les vertébrés ne possèdent pas les moyens de gagner la course aux armements contre les parasites simplement par les innovations qu'apportent la mutation, la recombinaison ou la reproduction sexuée tant leur cycle de vie est long comparé à ceux de leurs parasites.

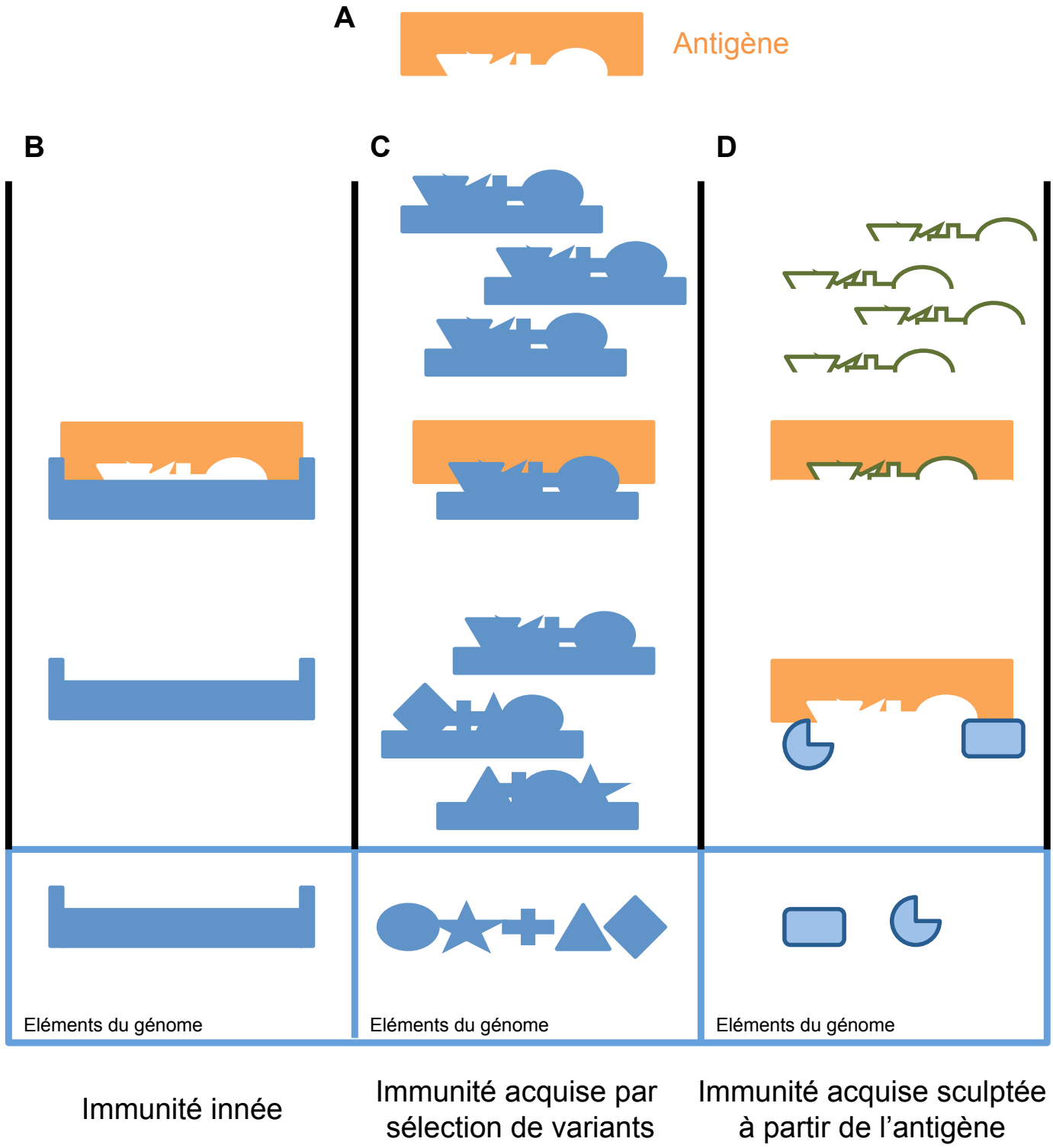


Figure 14. Les différents types d'immunités rencontrés dans le vivant. A: Antigène provenant d'un parasite. **B:** Immunité innée. Le génome code pour un récepteur invariable qui peut être capable de reconnaitre non spécifiquement l'antigène et de déclencher une réponse immunitaire (exemple des récepteurs TLR). **C:** Immunité acquise par sélection de variants à l'échelle des molécules. Le génome est capable de produire une grande variété de récepteurs à partir d'éléments du génome (par recombinaison somatique par exemple). Le récepteur correspondant le mieux à l'antigène est sélectionné et amplifié (exemples: Lymphocytes B et Dscam). **D:** Le génome de l'hôte produit uniquement des protéines accessoires. Lorsque ces protéines rencontrent un antigène, elles sculptent directement une molécule à partir de celui-ci exemple: CRISPR/Cas9 et siARN).

Cependant, presque par définition et quelque soit l'hôte, le parasite se reproduit et innove beaucoup plus rapidement que son hôte. Il semble donc quasiment inimaginable que la plupart des organismes en proie à des parasites ne possède pas la capacité de produire une réponse spécifique.
La définition de la réponse acquise se décompose en deux points. D'une part, l'acquisition d'une réponse spécifique non originellement présente dans l'hôte à partir du parasite. D'autre part, une immunité spécifique qui se maintient au cours de la vie de l'individu.

II.C.1. Acquisition d'une spécificité immunitaire

D'après les mécanismes connus à ce jour, on peut définir trois manières d'acquérir une spécificité immunitaire (Figure 14).

II.C.1.α. Spécificité innée

Elle résulte principalement de la présence ou de l'absence d'un récepteur spécifique dans le génome de l'hôte. Une augmentation de la spécificité ne peut se faire que par l'intermédiaire de la reproduction, soit par acquisition d'un nouveau récepteur par la reproduction sexuée, soit par des mutations dans la lignée germinale. La grande diversité des récepteurs riches en répétitions de leucines chez les plantes provient directement de l'expression de gènes correspondants (Figure 14B).

II.C.1.β. Spécificité acquise par sélection de variants à l'échelle des molécules

L'illustration parfaite de ce phénomène est la sélection des clones de lymphocyte B et T produisant des immunoglobulines spécifiques, mais aussi la création des molécules Dscam chez les arthropodes. Le soma est capable de produire une énorme diversité de molécules par l'intermédiaire de recombinaisons somatiques, de mutations somatiques ou d'épissage. Les molécules proposant la plus grande spécificité au parasite sont sélectionnées et leur production peut être amplifiée (Figure 14C).

II.C.1.γ. Spécificité acquise sculptée directement à partir d'un patron

C'est le cas des petits ARN des eucaryotes et du CRISPR des procaryotes. Certaines molécules du soma participent à l'élaboration de la molécule spécifique, mais la spécificité de la molécule provient directement d'un patron provenant du parasite. Dans le cas des petits ARN, les protéines dicer, les argonautes, les RDRP et les protéines du complexe RISC sont guidés par des molécules correspondant parfaitement au parasite. La diversité de molécules engendrée par ce phénomène est immense puisque le nombre de siARN différents s'élèvent est compris entre 4^{20} et 4^{25} pour des ARN mesurant de 20 à

25 nucléotides (Rimer et al., 2014) (Figure 14D) permettant ainsi une grande spécificité de la réponse.

La découverte des systèmes de défense par les petits ARN remet en cause la vision classique que nous avions de l'immunité. Ces systèmes produisent des molécules bien plus spécifiques que celles produites par les modifications somatiques des loci V(D)J. De plus, la découverte des Dscam et des FREP remet également en cause l'exclusivité d'une réponse immunitaire spécifique exclusive aux vertébrés même s'il n'est pas certains que cette réponse spécifique puisse être mise en mémoire comme c'est le cas pour les lymphocytes B.

II.C.2. Maintien/mise en mémoire d'une spécificité immunitaire

Les vertébrés possèdent un temps de génération suffisant pour tester le maintien ou la mise en mémoire de la réponse spécifique lors d'une réinfection. En effet, il est beaucoup plus complexe d'acquérir de vérifier si une mémoire immunitaire s'est mise en place pour des organismes avec une faible durée de vie. A durée de vie similaire, il semblerait que les plantes soient capables de mémoriser la réponse immunitaire (ROSS, 1961). La réponse à une infection locale se propage généralement dans toute la plante et prévient de ce fait une infection par le parasite concerné (SAR pour Systemic Aquired Resistance). Ce SAR est potentiellement transmis à la descendance, par l'intermédiaire de petits ARN ou de marques déposées sur l'ADN (Luna et Ton, 2012). De nombreux arthropodes, *Bombus terrestris* (Sadd et Schmid-Hempel, 2006), *Tribolium castaneum* (Roth et al., 2009), *Drosophila melanogaster* (Pham et al., 2007) montrent aussi une capacité à répondre rapidement à une seconde infection, voire même à transmettre cette résistance à leur progéniture (Sadd et Schmid-Hempel, 2007). Les mécanismes impliqués restent encore à élucider même s'il semble très peu probable qu'ils soient conservés parmi tous les animaux.

II. Conclusion

Le système immunitaire est basé sur un principe simple: identifier l'intrus, signaler sa présence et le neutraliser. La nécessité d'une protection constante contre des parasites omniprésents a favorisé la mise en place de nombreux mécanismes de protection au cours de l'évolution (Tableau 1). L'étude de l'évolution de ces mécanismes est conditionnée par la diversité des organismes modèles utilisés. On peut cependant constater de nombreuses similitudes entre certains systèmes immunitaires. Les voies Toll/TLR et Imm/TNF ont été largement conservées entre les arthropodes et les vertébrés

					Organismes étudiés	Récepteurs	Voies de signalisation	Effecteurs
Eucaryotes	Métazoaires	Bilatériens	Deutérostomiens	Vertébrés à mâchoire	Des requins aux mammifères	LRR; TLR; RLR; TNF; BCR; TCR; MHC	TIR;NF-κB ; STAT; cytokines, interférons, p38 MAPK	C'; phagocytose; AMP; lectines, IgSF (Anticorps); MCP; siARN?,NK
				Vertébrés agnathe	Lamproie	LRR; TLR	TIR; NF-κB; STAT	C'; phagocytose; AMP; lectines, IgSF (VLR)
				Céphalochordés	Amphoxius	LRR; TLR; TNF	TIR; NF-κB; STAT; PPO	C'; lectines, IgSF
				Urochordés	Ascidie	LRR; TLR; TNF	TIR; NF-κB; STAT; PPO	C'; phagocytose; lectines, IgSF, NK-like
				Echinodermes	Oursin	LRR; TLR; PGRP	TIR; NF-κB; STAT	C'; phagocytose; lectines, IgSF
			Protostomiens	Annélides	Lombric	TNF?	PPO	AMP; NK-like
				Mollusques	Escargot	IgSF (FREP)	PPO; TIR; NF-κB	AMP; lectines; IgSF (FREP)
				Nématodes	*C. elegans*	?	TIR; p38 MAPK	AMP; lectines; NLP
				Arthropodes	Drosophile; Moustique; Crevettes Papillos	LRR; Toll; PGRP; IgSF (Dscam); Dicer;	TIR; NF-κB; STAT	C'; phagocytose; AMP; IgSF (Dscam); siARN
		Cnidaires			Corail	?	PPO	C'
		Porifères			Eponge	Toll	TIR; NF-κB; PPO	Lectines; IgSF
	Plantes				*Arabidopsis thaliana*	LRR; LRR-NBS; Dicer	TIR; WRKY	MCP; ROS; AMP; siARN
Procaryotes					Bacterie; Archée			CRISPR; System R-M; MCP-like

Tableau 1. Evolution du système immunitaire. Récapitulatif des éléments de la réponse immunitaire identifiés dans différents groupes. En noir, les éléments du système immunitaire inné. En vert, les éléments dont la diversité est issue de modifications somatiques par sélection de variants. En rouge, les éléments dont la diversité est directement sculptée à partir du parasite. AMP: Peptides AntiMicrobiens; BCR: Récepteurs des Cellules B; C': Complément; CRISPR: Clustered Regularly Interspread Short palyndromics Repeats; IgSR: Super Famille des Immunoglobulines; LRR: Répétitions Riches en Leucine; MCP: Mort Cellulaire Programmée; MHC: Complexe Majeur d'Histocompatibilité; NBS-LRR: LRR avec un site de fixation aux nucléotides; NK: Cellule Tueuse Naturelle; PGRP: Protéine de Reconnaissance des Péptidoglycanes; PPO: ProPhenolOxydase; RLR: Récepteur de type RIG-I; TCR: Récepteur des Cellules T; TIR: Récepteur Toll-Interleukine; TLR: Récepteur de type Toll; TNF: Facteur de Nécrose Tumorale.

(Figure 12). L'obtention de nouveaux génomes suggère même une conservation de ces voies parmi les Eumétazoaires. D'autres mécanismes semblent encore plus anciens, comme l'ARN interférence, réponse antivirale, dont l'origine remonterait à l'ancêtre commun des eucaryotes. Certains domaines protéiques ont été sélectionnés indépendamment pour leur grande capacité d'interaction avec les molécules associées aux parasites ; c'est le cas, par exemple, des récepteurs riches en leucines présents chez les animaux mais aussi chez les plantes (Figure 12).
La découverte récente du système CRISPR/Cas, des Dscam, des FREP et le fait que l'ARN interférence semble être un système partagé par beaucoup d'organismes remet en cause la dichotomie classique entre l'immunité innée et l'immunité acquise telle que l'entendent les immunologistes.

La prochaine partie de cette introduction va se focaliser sur la compréhension de la dynamique des interactions hôte-parasite à l'échelle des populations. Elle décrira dans un premier temps les nouveaux outils actuellement disponibles pour l'identification des bases génétiques de ces interactions dans les populations naturelles pour ensuite se focaliser sur l'adéquation entre les modèles d'interaction et les études empiriques effectuées sur des populations naturelles.

III. Bases génétiques et microévolution des interactions hôtes-parasites

La plupart des mécanismes décrits précédemment sont issus d'études de génétique inverse effectuées sur des organismes modèles. Ces études ont beaucoup apporté quant à la compréhension des mécanismes de l'immunité mais ne rendent pas compte de l'évolution des interactions hôte-parasite dans la nature. Les organismes étudiés en écologie évolutive sont généralement différents de ceux étudiés en biologie cellulaire et moléculaire. Par conséquent les techniques de génétique inverse ne sont pas encore développées dans ces organismes. Cependant, la nouvelle ère de la génomique permet d'identifier les bases génétiques des interactions hôte-parasite par des approches pangénomiques.

III.A. Bases génétiques des interactions hôte-parasite

III.A.1. Architecture génétique de la résistance aux parasites chez l'hôte

L'amélioration et la diminution du coût des techniques de séquençage permettent l'acquisition de la séquence de nombreux génomes pour l'étude de la variabilité génétique des populations naturelles ou artificielles ainsi que la comparaison des génomes entre espèces proches. Les différentes approches qui vont être décrites apportent des informations essentielles sur les bases génétiques (nombre de loci impliqués et leur localisation) de la résistance de l'hôte à un parasite donné.

II.A.1.α. Cartographie de Loci de Caractère Quantitatif

Dans le cadre de l'étude des interactions hôtes parasites, de nombreux caractères sont quantifiables: le degré de résistance à une infection, la charge parasitaire, l'intensité de la réponse immunitaire...
La cartographie de Loci de Caractère Quantitatif (LCQ, ou QTL en anglais) a pour but de localiser les gènes potentiels (locus) contrôlant un caractère quantitatif et de déterminer leur interaction. On utilise pour cela deux génotypes parentaux généralement différents pour le caractère étudié. Ces génotypes sont croisés afin d'obtenir des descendants hybrides en génération F1. Ces individus F1 sont croisés entre eux afin d'obtenir différentes combinaisons des deux génomes parentaux dans la génération F2. Selon l'organisme étudié (système de reproduction et dépression de consanguinité), les

croisements frère-sœur ou les cycles d'autofécondation peuvent être poursuivis afin d'obtenir des individus possédant des génotypes "mosaïques" à l'état homozygote. L'apport des nouvelles techniques de séquençage réside dans la facilité d'obtention de marqueurs génétiques le long des chromosomes. Le séquençage des lignées recombinantes permet d'identifier l'origine parentale des différentes régions de leur génome. Le caractère étudié est quantifié pour ces lignées et est associé à leur génotype. Le résultat de l'étude ne permet généralement pas de localiser précisément les loci responsables du caractère étudié mais une région proche ou contenant ces loci. Une cartographie plus précise peut être effectuée à l'aide de nouveaux croisements informatifs si d'autres marqueurs sont à disposition dans la région d'intérêt.

II.A.1.β. Etude d'Association Pangénomique

La technique précédente permet d'identifier les loci responsables de la variation d'un caractère entre deux génotypes particuliers. La diversité naturelle n'est donc pas entièrement représentée. De plus, les espèces étudiées ne sont pas toutes compatibles avec des croisements contrôlés. L'étude d'association pangénomique (EAP, ou GWAS en anglais) permet de prendre en compte la totalité de la diversité connue pour l'espèce étudiée. Elle fonctionne sur le même principe que l'analyse LCQ mais repose sur l'étude d'une population naturelle. Alors que les analyses LCQ sont fréquemment retrouvées dans la littérature, les EPA sont encore assez peu utilisées pour l'étude des interactions hôtes-parasites (Magwire et al., 2012) alors qu'elles sont utilisées régulièrement pour l'identification des bases génétiques des maladie héréditaires chez l'homme.

II.A.1.γ. Bases génétiques de la résistance

Il est fréquent de trouver plus d'un seul locus responsable de la sensibilité à un parasite chez un hôte. La variance d'un phénotype (V_P) se partitionne entre plusieurs paramètres:

$$V_P=V_G+V_E+V_{GxE}+V_{Résiduelle}$$

Avec V_G, la variance associée au génotype, V_E représentant la part de l'environnement et enfin V_{GxE} la variance associée à une interaction entre le génotype et l'environnement. Dans le cas des analyses LCQ ou des EPA, le facteur environnemental est considéré comme constant.
Ainsi, on a:

$$V_P=V_G+V_{Résiduelle}$$

Cependant, la variance génétique peut également se partitionner en:

$$V_G=V_A+V_D+V_I+...+ \text{(autres termes)}$$

Avec V_A représentant la part des effets additifs entre les gènes (contribution indépendante des gènes), V_D représentant la dominance d'un allèle sur un autre (dans le cas des organismes polyploïdes) et enfin V_I qui représente l'interaction de type épistatique entre deux allèles à différents loci.

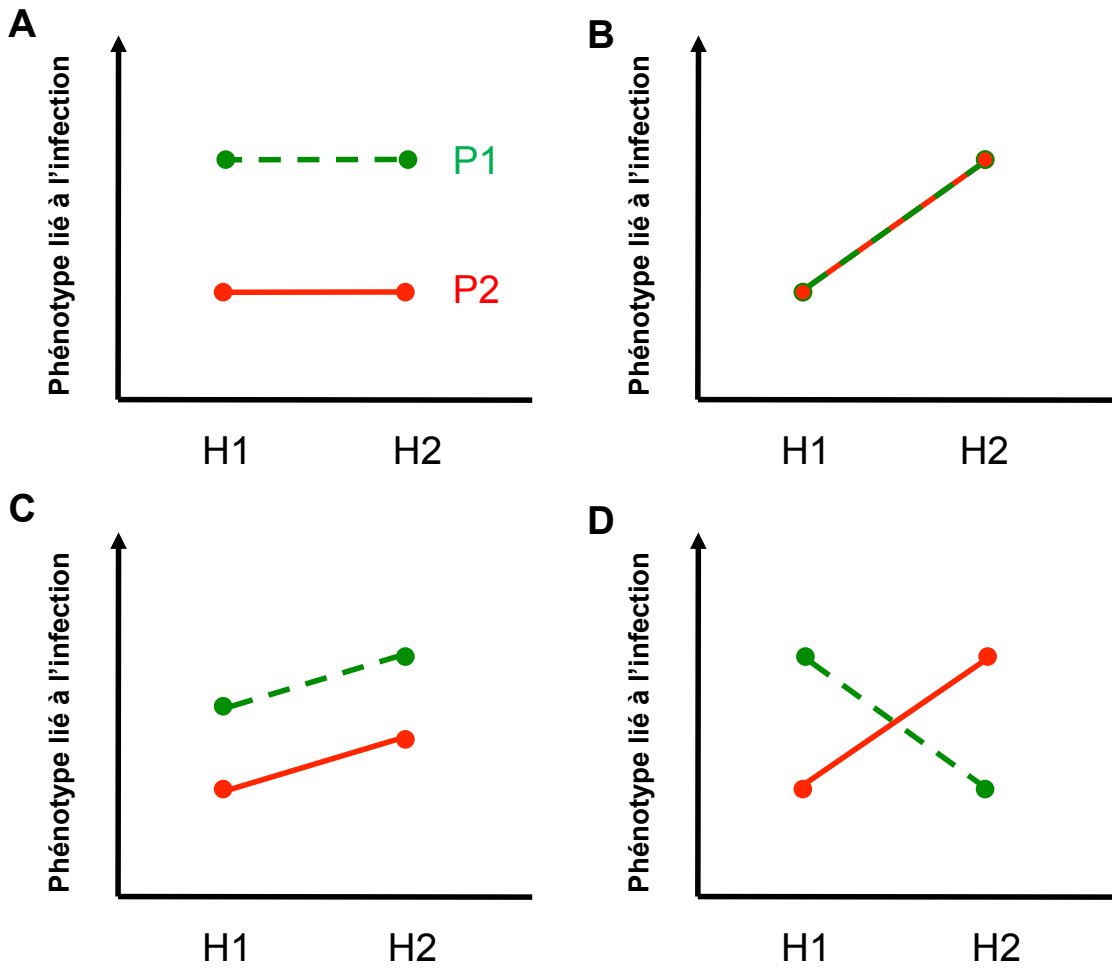
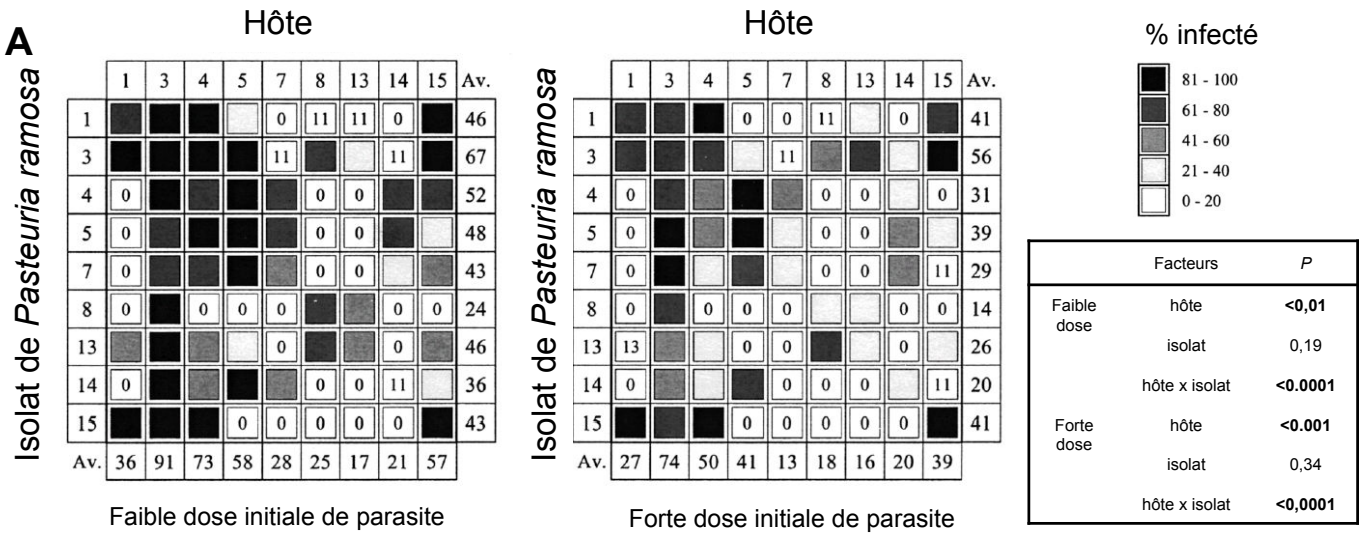


Figure 15. Infection hypothétique de deux génotypes d'hôte (H1 et H2) par deux génotypes de parasite (P1 et P2). A: La différence de phénotype lié à l'infection ne s'explique que par une différence de génotype des parasites: P1 a un effet plus important sur les deux hôtes. **B:** La différence de phénotype lié à l'infection ne s'explique que par une différence de génotype des hôtes: H1 est probablement plus résistant que H2. **C:** La différence de phénotype lié à l'infection s'explique à la fois par une différence de génotype des parasites et des hôtes: P1 a une effet plus important sur les deux hôtes et H2 est globalement plus sensible aux deux parasites. **D: :** La différence de phénotype lié à l'infection s'explique par la combinaison des génotypes de l'hôte et du parasite: P1 est capable d'infecter H1 plus facilement tandis que P2 est capable d'infecter H2 plus facilement.



	Facteurs	P
Faible dose	hôte	**<0,01**
	isolat	0,19
	hôte x isolat	**<0.0001**
Forte dose	hôte	**<0.001**
	isolat	0,34
	hôte x isolat	**<0,0001**

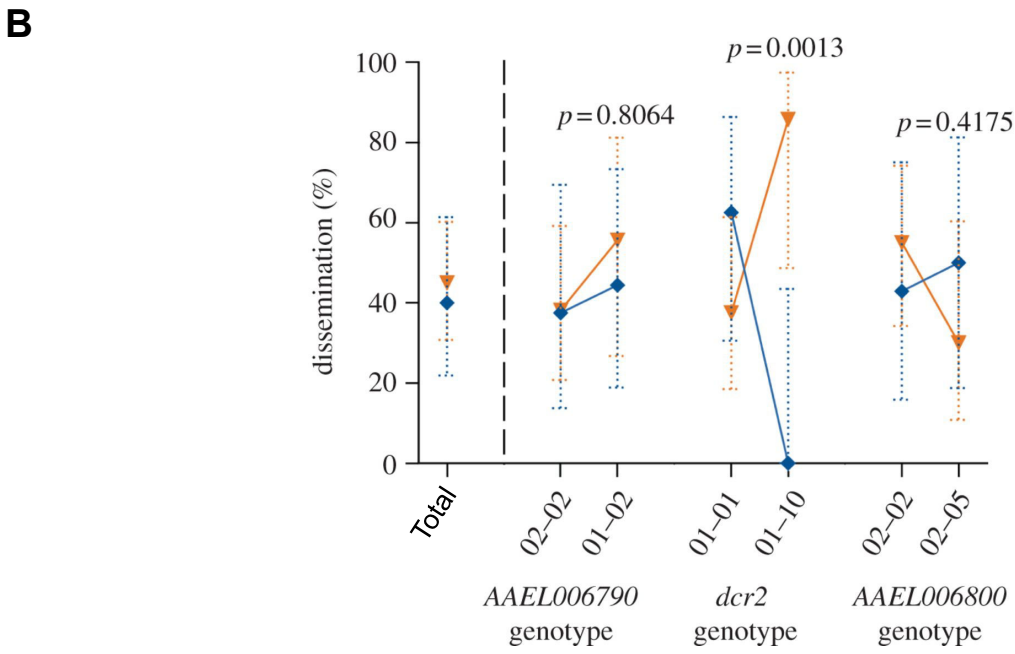


Figure 16. Mise en évidence d'interactions entre génotypes de l'hôte et du parasite. A: Test d'infections croisées de 9 clones naturels de *Daphnia magna* par 9 isolats de *Pasteuria ramosa* retrouvés sur ces hôtes pour deux concentrations initiales de parasite différentes. Le pourcentage d'animaux infecté dépend du génotype de l'hôte et de l'interaction entre le génotype de l'hôte et l'isolat de parasite. Adapté de Carius et al. (2001). **B:** Pourcentage de dissémination de deux souches du virus de la dengue (orange et bleue) en fonction de l'allèle au locus *drh-2* chez *Aedes aegypti*. Les locus *AAEL006790* et *AAEL006800* sont physiquement proches de *dcr-2*. L'interaction entre le génotype au locus *drh-2* et le souche de virus explique de manière significative le pourcentage de dissémination observé. Adapté de Lambrechts et al. (2013).

Des études théoriques prédisent que lors du processus d'adaptation à une forte pression de sélection, les premières mutations sélectionnées sont généralement celles avec le plus grand effet (Orr, 2002). Ces prédictions, observées par quelques études expérimentales (Bull et al., 2000; Wichman et al., 1999), suggèrent que seuls quelques loci sont suffisant pour expliquer une grande partie des variations phénotypiques soumises à une forte pression de sélection, comme par exemple la résistance à un parasite. Cette prédiction est vérifiée empiriquement par l'étude de 500 tentatives d'identification de QTL responsables de la sensibilité à un parasite. Ainsi, le nombre moyen de loci à effets principaux (V_A) identifiés est compris entre 3 et 5 expliquant entre 30% et 50% de la variance de sensibilité au parasite (Wilfert et Schmid-Hempel, 2008). Toujours d'après cette étude, il apparaît que les gènes à effets majeurs sont spécifiques des parasites rencontrés. Il n'en reste pas moins que de nombreux loci à effet mineur, sans doute épistatiques, participent aussi à la variation de sensibilité. Ces loci sont cependant plus difficiles à identifiés.

III.A.2. Bases génétiques de l'infection: interactions entre le génotype du parasite et le génotype de l'hôte

Dans la nature, il est rare d'observer une prévalence d'infection par un même parasite de 100%. Ceci peut être, entre autres, dû au fait que certains individus de la population sont moins sensibles au parasite que certains autres. Par exemple: considérons une population d'hôtes composés de deux génotypes différents H1 et H2 et une population de parasite composé de deux génotypes différents P1 et P2. L'inoculation de chacun des génotypes hôtes avec chacun des génotypes parasites (infections croisées) peut aboutir à différents types de patron d'infection: i) l'infection par P1, quelque soit l'hôte, est plus marquée que l'infection par P2 , ii) l'infection de H2, quelque soit le parasite, est plus marquée que l'infection de H1 ou iii) l'infection de H1 par P1 est plus marquée que l'infection de H2 par P1 et l'infection de H2 par P2 est plus marquée que l'infection de H1 par P2. Dans i), le succès de l'infection est simplement dû au génotype du parasite. Dans ii), le succès de l'infection est uniquement dû au génotype de l'hôte. Dans iii), il existe une interaction entre les génotypes: le génotype P1 infecte plus facilement le génotype H1 et le génotype P2 infecte plus facilement le génotype H2 (Figure 15).
De nombreuses expériences d'infections croisées entre différentes lignées d'hôtes et différentes lignées de parasites ont mis en évidence un rôle majeur des interactions spécifiques entre le génotype des hôtes et les génotypes des parasites dans différentes espèces (Figure 16): *Bimphalaria glabrata*/*Schistosoma mansoni* (Hoffman et al., 1998); *Linum sp./Melampsora lini* (Ellis et al., 1999); *Bombus terrestris*/*Crithidia bombi* (Schmid-Hempel 1999); *Daphnia magna*/*Pasteuria ramosa* (Carius et al., 2001);

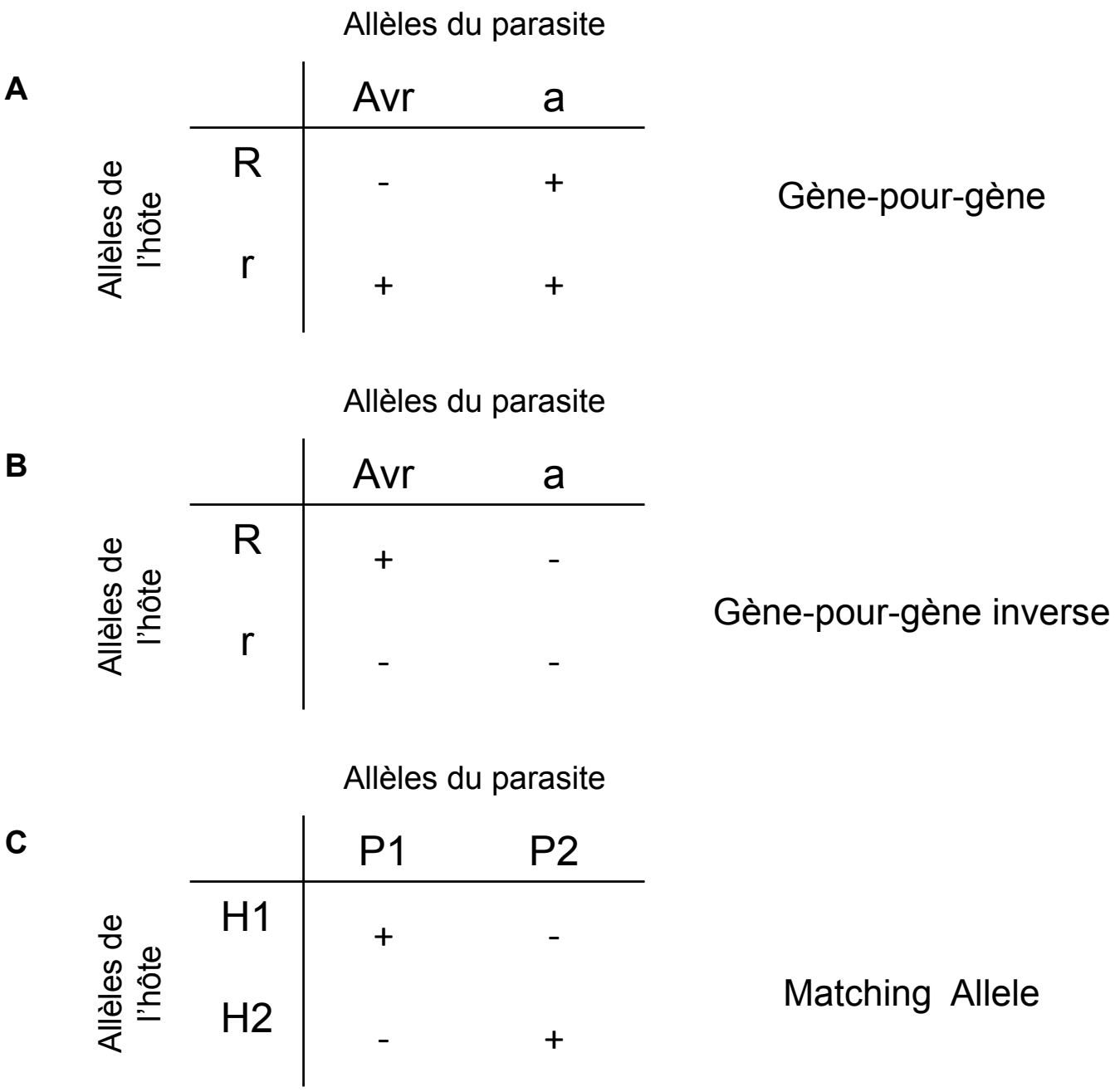


Figure 17. Modèles d'interaction entre les génotypes de l'hôte et du parasite. A: Modèle gène-pour-gène. La résistance (-) a lieu lorsqu'un gène de résistance R de l'hôte correspond à un gène d'« avirulence » Avr du parasite. Lorsque cela n'est pas le cas, l'infection est réalisée (+). **B:** Modèle gène-pour-gène inverse. Le parasite doit reconnaitre l'hôte pour que l'infection ait lieu. Si l'hôte perd ce récepteur, l'infection ne se fait pas. **C:** Modèle « Matching allele ». La résistance a lieu si le génotype de l'hôte correspond au génotype du parasite, autrement dit, si l'hôte possède les allèles capables de reconnaitre ceux du parasite.

Drosophila melanogaster/Serratia marcescens (Lazzaro et al., 2004); *Aedes aegypti/Dengue* (Lambrechts et al., 2013). Ce type d'interaction a également été révélé entre des souches naturelles de *Caenorhabditis elegans* et différentes souches de *Serratia marcescens* (Schulenburg et Ewbank, 2004).
Différents modèles tentent d'expliquer ce type d'interaction entre génotypes. Deux modèles principaux ont émergé: le modèle d'interaction gène-pour-gène (Flor, 1956), historiquement admis chez les plantes et le modèle d'interaction "matching-allele" traditionnellement utilisé chez les animaux.

III.A.2.α. Modèle gène-pour-gène

Le modèle gène-pour-gène est basé sur le fait qu'un facteur (gène) provenant du parasite interagit directement avec un facteur (gène) de l'hôte. La résistance n'a lieu que si le facteur de l'hôte correspond au facteur du parasite, suggérant une interaction spécifique entre un génotype d'hôte et un génotype de parasite (Figure 17A). Ce modèle implique qu'il existe des "super-parasites" capables d'infecter tous types d'hôtes et des hôtes sensibles à tous les parasites.
Ce modèle est en grande partie validé par le fonctionnement du système immunitaire chez les végétaux. La plupart des bactéries et des champignons, à défaut de pouvoir franchir directement la paroi végétale, injectent dans la cellule des éliciteurs, généralement codés par des gènes dits d'avirulence (Avr). Les plantes possèdent une grande diversité de récepteurs LRR cytosoliques (gène R) capables de reconnaître une grande diversité d'éliciteurs. Cette reconnaissance initie une réponse hypersensible détruisant les tissus infectés et protégeant ainsi le reste de la plante. Si la plante ne possède pas le récepteur correspondant ou si le parasite produit un nouvel éliciteur, l'infection a lieu. La molécule produite par le parasite doit être une molécule importante pour le fonctionnement du parasite et qui, par conséquent, doit être sous sélection stabilisante.
Ce système n'est pas exclusif aux plantes. Les animaux présentent des protéines capables de reconnaître les peptidoglycanes, molécules spécifiques et indispensables des membranes bactériennes et une relation gène-pour-gène a été mise en évidence entre la Drosophile et le sigma virus (Wilfert et Jiggins, 2010).
Bien évidemment, ce modèle montre des limites. Il implique notamment l'existence de super-parasites ou d'hôtes universellement sensibles. Ce genre d'organisme n'a pour le moment jamais été découvert, infirmant ainsi la généralisation du modèle gène-pour-gène à toutes les interactions hôte-parasite. Il existe cependant des parasites généralistes (*Bacillus thuringiensis*, *Pseudomonas aeruginosa*) qui possèdent une gamme d'hôtes très importante.

III.A.2.β. Modèle gène-pour-gène inverse

Dans ce modèle, une infection n'est possible que si le parasite est capable de reconnaître un récepteur de l'hôte (permettant son entrée par exemple) (Figure 17B). Ce modèle diffère du modèle gène-pour-gène dans le fait qu'il n'y a plus d'hôte sensible à tous les parasites ni de super parasites (Fenton et al., 2009). Ce modèle est plus réaliste pour les interactions bactérie/phage puisque l'entrée des phages est généralement conditionnée par la reconnaissance de molécules de surfaces de la cellule hôte par les virions.

III.A.2.γ. Modèle "Matching allele"

Ce modèle est basé sur la reconnaissance spécifique d'un allèle (et non d'un facteur) du parasite par un allèle de l'hôte. Si l'hôte reconnaît l'allèle du parasite, l'infection n'a pas lieu (Figure 17C). Malgré le fait que ce modèle semble bien s'accorder avec le fonctionnement global du système immunitaire des animaux, peu d'études empiriques confirment ce modèle. Une des seules études provient du modèle *Daphnia magna*/*Pasteuria ramosa* (Luijckx et al., 2013): par ségrégation des loci et des allèles, les auteurs parviennent à identifier le déterminisme génétique responsable de la résistance ou la sensibilité de l'hôte en fonction du génotype du parasite.

Ces différents modèles restent bien évidemment théoriques. La réalité montre que le nombre de loci impliqués dans les interactions hôte-parasite est assez important (3 à 5 loci à effet majeur en moyenne). Chaque locus, que ce soit chez l'hôte ou le parasite, interagit différemment avec son locus correspondant, suggérant que les modèles gène-pour-gène et matching-allèles ne sont que des cas extrêmes d'un continuum qui pourrait combiner chacun des modèles (Agrawal et Lively, 2003).

III.B. Impact de la coévolution sur la diversité génétique des hôtes et des parasites

III.B.1. Le modèle gène-pour-gène: une succession de balayages sélectifs

Le problème réside dans l'adéquation entre les modèles et les résultats observés empiriquement. Le modèle gène-pour-gène semble correspondre au type d'interactions qu'ont les plantes avec leurs parasites. Si on laisse évoluer ce modèle, on s'attend à une succession de fixation de nouveaux allèles tant chez le parasite que chez l'hôte (Figure 18B). Lorsqu'un nouveau gène Avr apparaît dans une population de parasites, on s'attend à ce qu'il se fixe rapidement. De même, le gène *R* correspondant, procurant la résistance

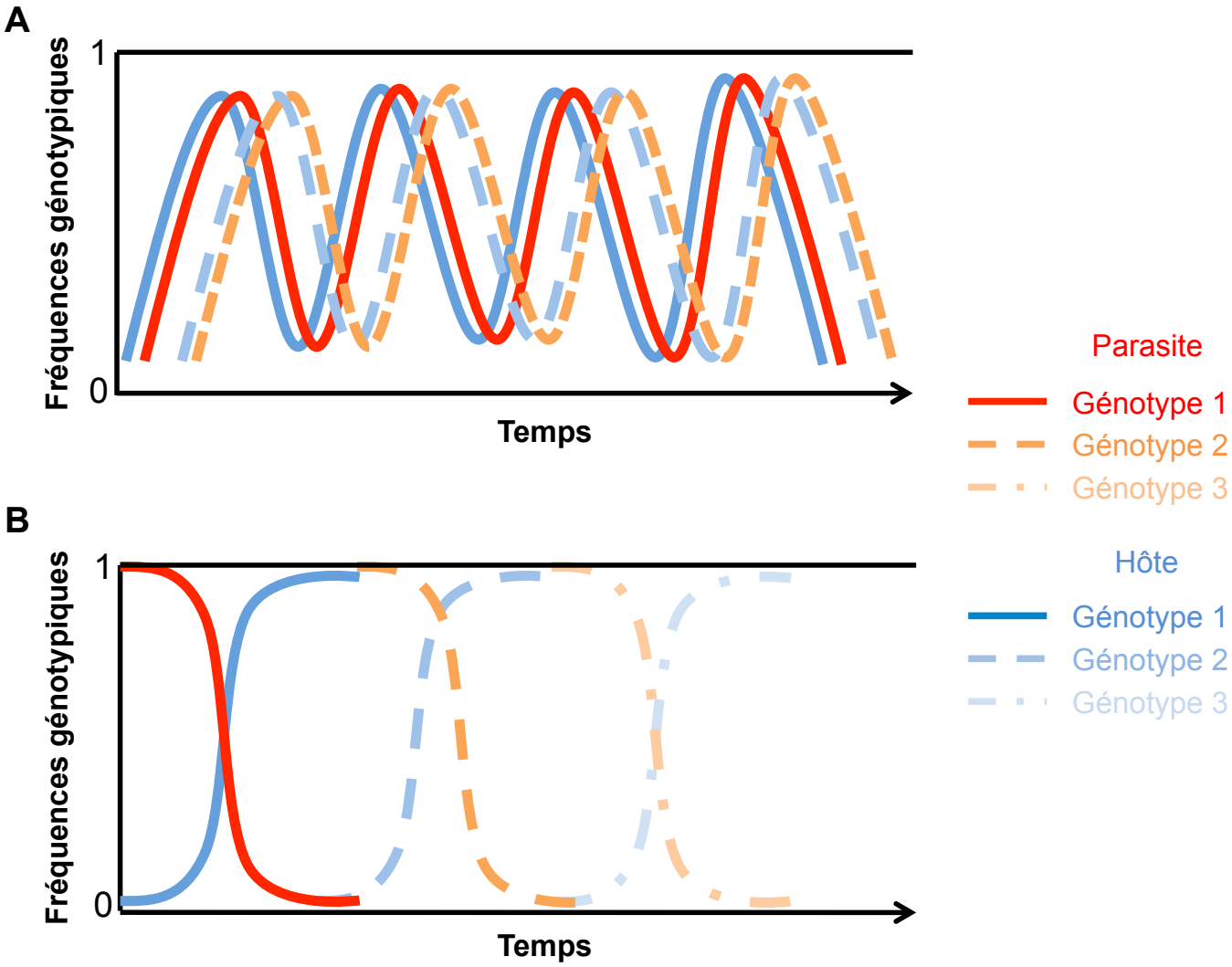


Figure 18. Evolution des fréquences allèliques au cours du temps dans un système d'interactions antagonistes. A: Sélection négative dépendante de la fréquence ou dynamique de la Reine Rouge. L'allèle le plus commun est contre-sélectionné dans la population d'hôte puis dans celle de parasites. Le polymorphisme est conservé sur le long terme dans les populations. **B:** Succession de balayages sélectifs. La présence d'un nouveau parasite induit une fixation rapide dans la population d'un allèle conférant une résistance chez l'hôte. De manière similaire, un allèle avantageux pour le parasite va se fixer rapidement dans la population. Le polymorphisme initial ne se maintient pas au cours du temps.

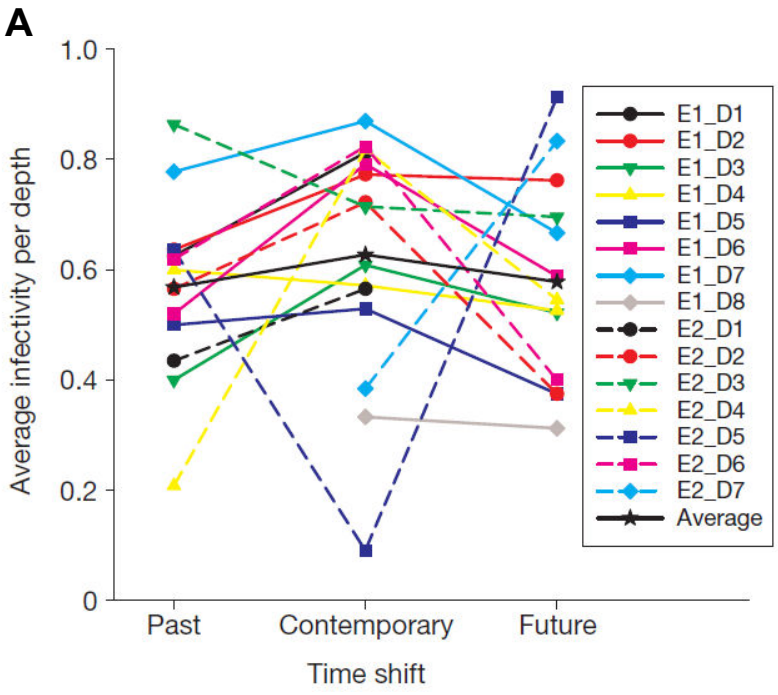


Table 1 | Results of logistic regression on infectivity

Parameter	Degrees of freedom	Residuals of deviance	Residuals of degrees of freedom	Deviance	*P*
Null	–	–	844	1,143.56	–
Time shift	2	5.89	842	1,137.67	0.05
Experiment	1	0.20	841	1,137.47	0.65
Clone depth	7	11.61	834	1,125.86	0.11
Clone depth × time shift	12	57.44	820	1,063.11	<0.0001

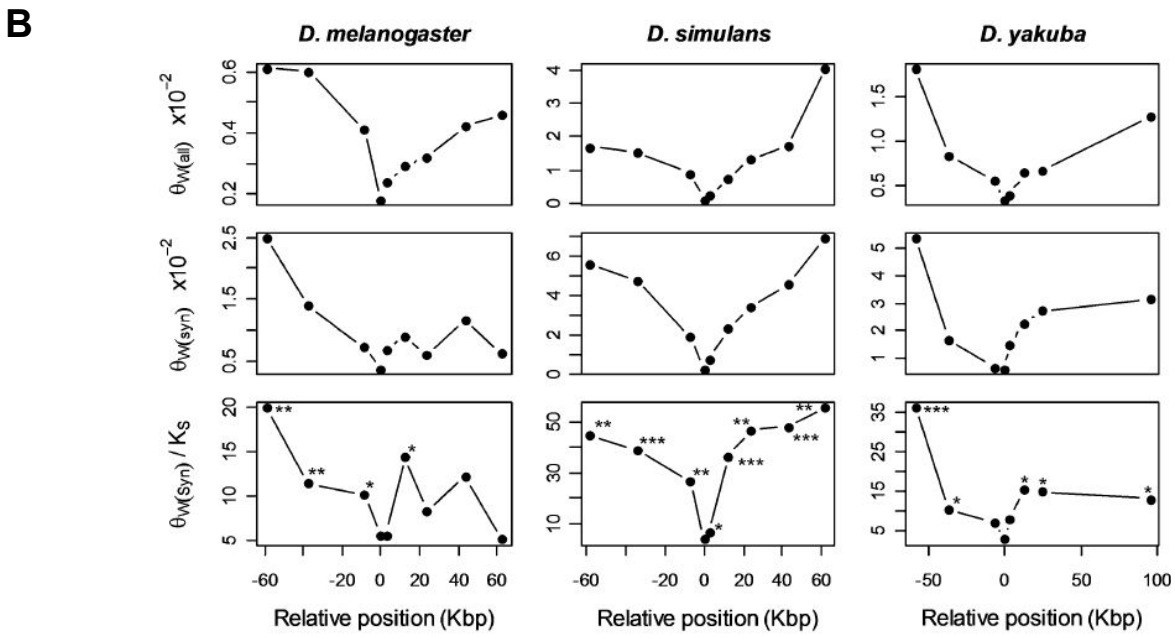


Figure 19. Exemple de sélection négative dépendante de la fréquence et de balayage sélectif. A: Infections croisées de *Daphnia magna* provenant de différentes couches de sédiments d'un même lac par des isolats de *Pasteuria ramosa* provenant de la même couche (Contemporary), de couches inférieures (Past) ou de couches supérieures (Future). Les parasites semblent plus efficaces sur leur hôte contemporain, suggérant une dynamique de type reine Rouge. D'après Decaestacker et al. (2007). **B:** Diversité génétique autour du gène AGO-2 (argonaute impliqué dans la réponse antivirale des drosophiles) pour *Drosophila melanogaster, D. simulans* et *D. yakuba*. L'absence totale de diversité suggère une sélection directionnelle pour ce gène dans les trois espèces. θ_w représente la diversité génétique à tout les sites (all) et pour les sites synonymes (syn). K_S représente la divergence entre les espèces. D'après Obbard et al. (2010).

aux nouveaux parasites, va se fixer rapidement dans la population d'hôte. Ce type de résultat ne correspond pas à ce qui est observé dans la nature: un maintien du polymorphisme chez l'hôte et chez le parasite (Brown et Tellier, 2011). Toutefois, la prise en compte de nombreux facteurs non intégrés originellement dans le modèle peut aboutir au maintien du polymorphisme (Brown et Tellier, 2011) par l'obtention d'une dynamique de sélection négative dépendante de la fréquence (NFDS).

III.B.2. Le modèle "matching allele": sélection négative dépendante de la fréquence

Un maintien du polymorphisme par NFDS est directement obtenu lorsque le modèle de matching-allele est utilisé (Frank, 1993). Cette dynamique, aussi nommée dynamique de la Reine Rouge en référence au roman de Lewis Carroll "*De l'autre côté du miroir*" (Van Valen, 1973), prédit une variation cyclique de la fréquence des allèles en fonction de leur prépodérance dans la population (Figure 18A). Plus un allèle devient fréquent dans la population, plus celui-ci est désavantagé tandis qu'un allèle rare sera favorisé. En effet, il est avantageux pour le parasite de pouvoir infecter l'hôte le plus fréquent tandis qu'il est plus avantageux pour l'hôte que l'allèle le moins susceptible à l'infection soit majoritaire. Ce modèle est soutenu par des expériences d'infections croisées entre des lignées d'hôtes et de parasites provenant de populations actuelles et des lignées d'hôtes et de parasites provenant de carottes sédimentaires (Decaestecker et al., 2007) (Figure 19A). Les résultats suggèrent que les parasites contemporains de leurs hôtes possèdent une fitness plus élevée que les parasites infectant les autres hôtes. Malgré une recrudescence d'études visant à prouver empiriquement ce modèle (Brockhurst et Koskella, 2013; Eizaguirre et al., 2012; Jokela et al., 2009; Lively et Dybdahl, 2000), le modèle est cependant limité par une de ses conditions: une population infinie ou constante au cours du temps. L'introduction d'une fluctuation dans la taille des populations, particulièrement évidente dans le contexte d'interactions antagonistes, tend à une fixation beaucoup plus rapide des allèles dans les populations (Gokhale et al., 2013). La notion de taille de population nécessite également l'introduction de la stochasticité démographique, qui ne fait que promouvoir la fixation des allèles dans la population (Gokhale et al., 2013).
L'utilisation de modèles pour comprendre la dynamique des interactions hôtes-parasites à l'échelle des populations se justifie par les difficultés pratiques qui peuvent être rencontrées lors de l'utilisation d'animaux ou de végétaux et de leurs parasites. Cependant, le nombre de paramètres (génétiques, écologiques, démographiques etc.) à prendre en compte pour une modélisation précise du phénomène semble illimité.

III.B.3. Adaptation locale

La mosaïque géographique, définie par la variation spatiale des conditions environnementales, joue un rôle majeur dans l'évolution des interactions hôte-parasite et plus particulièrement dans le maintien de la diversité intra-spécifique. De nombreux facteurs environnementaux influencent le devenir de la coévolution à l'échelle locale tant et si bien que les sous-populations évoluent indépendamment les unes des autres (Buckling et Rainey, 2002; Kaltz et Shykoff, 1998). Une méta-analyse effectuée sur 57 systèmes hôte-parasite montre que l'adaptation locale d'un parasite (parasites plus adaptés à leur propre population) est associée à un taux de migration supérieur de la part du parasite (Greischar et Koskella, 2007; Laine, 2009). De manière surprenante, cette adaptation locale peut être observée quelle que soit la distance entre les populations, suggérant l'importance fondamentale des microenvironnements dans le maintien de la diversité génétique à l'échelle des populations.

III.C. Signatures de la sélection par les parasites

La forte pression de sélection exercée par les parasites sur le système immunitaire de leurs hôtes laisse des signaux dans le génome de ces derniers. De manière similaire, le système immunitaire exerce une forte pression de sélection sur les parasites. De manière générale, les études pangénomiques tendent toutes vers la même conclusion : chez les hôtes, les gènes impliqués dans l'immunité évoluent plus rapidement que les autres catégories de gènes (Hurst et Smith, 1999); (Obbard et al., 2009); (Downing et al., 2009).
L'observation des signaux de sélection présents dans les génomes des hôtes et des parasites permet de se libérer des modèles précédemment établis et de reconstituer l'histoire évolutive des interactions hôte-parasite au cas par cas. L'accumulation de données empiriques sur les génomes pourrait ensuite permettre une compréhension globale de l'évolution de ces interactions.
Une fois les polymorphismes expliquant la résistance ou la sensibilité au parasite identifiés, il est possible d'étudier le type de sélection exercé par le parasite sur les gènes impliqués.
Trois types de sélection peuvent être identifiés: la sélection stabilisatrice, la sélection directionnelle et la sélection équilibrante. La sélection stabilisatrice ne présente pas ici grand intérêt puisqu'elle n'induit pas de changements majeurs dans le processus de coévolution. Nous nous intéresserons particulièrement aux deux autres types de sélection : la sélection directionnelle et la sélection équilibrante.

III.C.1. Comment détecter la sélection dans les génomes ?

La sélection directionnelle fixe rapidement les allèles avantageux dans une population. Si la sélection est rapide, l'environnement génomique immédiat de l'allèle va être très peu variable (peu de recombinaisons et peu de mutations). Par conséquent, une perte de diversité dans une région génomique est signe d'une potentielle sélection directionnelle rapide. En revanche, on observe une sélection équilibrante à un locus lorsque plusieurs allèles sont maintenus à des fréquences intermédiaires dans la population. La sélection équilibrante maintient la diversité dans une population plus longtemps qu'en absence de sélection. Sa signature réside dans l'excès de sites polymorphes à une fréquence intermédiaire.

III.C.1.α. Méthode basée sur l'évolution des gènes

Les méthodes basées sur l'évolution des gènes utilisent généralement le ratio d_N/d_S (ou encore K_A/K_S) et s'utilise à la fois intra et inter espèces. Il s'agit de comparer le taux de substitutions non-synonymes par site d_N au taux de substitutions synonymes par site d_S. Les substitutions synonymes sont supposées être neutre (silencieuse) tandis que les substitutions non-synonymes sont soumises à sélection. Un $d_N/d_S>1$ suggère une sélection positive du gène étudié tandis qu'un $d_N/d_S<1$ témoigne d'une sélection stabilisante s'applique sur ce gène (Hurst, 2002). Le calcul de ce ratio peut être corrigé par la probabilité d'observer un changement donné à cette position (taux de transitions ou de transversion) ainsi que par le biais d'usage préférentiel de codons particuliers par certains organismes. Le test de McDonald-Kreitman (MKT) compare les d_N/d_S inter et intra spécifiques (McDonald et Kreitman, 1991). Sous hypothèse de neutralité, les rapports sont similaires. En revanche, un rapport inter espèces significativement supérieur au ratio intra espèces suggère une sélection positive pour une des espèces considérées. Inversement, un d_N/d_S inter espèces significativement inférieur au d_N/d_S intra espèces suggère que ce gène est soumis à une sélection équilibrante ou un relâchement de la sélection à ce locus pour une ou plusieurs des espèces considérées. De manière similaire, le test de Hudson-Kreitman-Aguadé (HKA) compare les différences observées aux échelles intra et inter espèces. Ce test compare les différences de loci fixés entre espèces (D) au polymorphisme intra espèces à ces loci (P) (Hudson et al., 1987), deux paramètres dépendants du taux de mutation des organismes étudiés. Un rapport D/P important signifie soit que les changements aboutissant à la spéciation entre espèces se sont faits rapidement, soit que la diversité intra espèce est faible (signature d'une sélection directionnelle au sein de l'espèce). Des valeurs faibles pour ce rapport suggèrent un maintien des différents allèles au sein de l'espèce caractéristique d'une sélection équilibrante.

III.C.1.β. Tests basés sur la distribution de fréquences alléliques

Lorsqu'un allèle est positivement sélectionné dans une population, sa fréquence augmente généralement rapidement, jusqu'à ce qu'il soit éventuellement fixé (balayage sélectif). Cette fixation rapide provoque une diminution de la diversité autour du locus sélectionné. Une fois cette région fixée, de nouvelles mutations, de faible fréquence, apparaissent dans la région et persistent pendant de nombreuses générations. Ainsi, une région qui a subi un balayage sélectif se caractérise par un excès d'allèles rares comparé au reste du génome. Le D de Tajima fut le premier test utilisant cette propriété et reste aujourd'hui communément utilisé (Tajima, 1989). Le D de Tajima compare le nombre de différences entre individus pris deux-à-deux avec le nombre de sites polymorphes dans une population. permet de tester la fréquence de distribution d'un allèle dans la population dans un échantillon donné par rapport à la situation attendue en absence de sélection (neutralité, (Kimura, 1983)). Ce test utilise le paramètre π, qui représente la diversité nucléotidique pour une séquence donnée dans une population ainsi que le paramètre *S* qui représente le nombre de sites polymorphes dans la séquence. Ainsi un estimateur θ_S (Watterson, 1975) du nombre observé de sites polymorphes pour un échantillon de n séquences est :

$$\theta_S = \frac{S}{\sum_{i=1}^{n-1} \frac{1}{i}}$$

Sous réserve de neutralité, l'estimateur θ_π de la diversité attendue est:

$$E(\pi) = \theta_\pi = 4N_e\mu$$

avec N_e la taille effective de la population et μ le taux de mutation sous neutralité. θ_S est comparé à θ_π pour le calcul du D :

$$D = \frac{\theta\pi - \theta_S}{\sqrt{V(\theta\pi - \theta_S)}}$$

Si la valeur D est négative, l'hypothèse de neutralité est rejetée et l'hypothèse de sélection directionnelle est envisageable. Lorsque le D de Tajima est positif et que l'indice de fixation (F_{ST}) entre les sous-populations à ce locus est faible, on considère alors que ce locus est sous un régime de sélection équilibrante. Ce type de sélection est attendue lors d'un avantage aux hétérozygotes ou d'une NFDS (Tajima, 1989). De nombreux autres

tests sont basés sur des tests similaires au D de Tajima. Plusieurs variations de cette méthode ont été utilisées, notamment en prenant en compte l'état ancestral ou dérivé de chacun des allèles en utilisant un groupe externe et en mesurant la proportion d'allèles rares de différentes manières (Fu et Li, 1993; Fu, 1997). Les balayages sélectifs induisent également une augmentation de la fréquence des allèles physiquement proches de l'allèle sélectionné (processus d'auto-stop évolutif). Le H de Fay et Wu (Fay et Wu, 2000) compare deux-à-deux le nombre de différences entre individus au nombre d'individus homozygotes pour l'allèle dérivé. Un faible H indique non pas un excès d'allèles rares mais un excès d'allèles dérivés rares, suggérant une sélection positive dans la région observée.

III.C.1.γ. Tests basés sur le déséquilibre de liaison

Il y a déséquilibre de liaison (DL) entre deux allèles de loci différents lorsque qu'il y a une surreprésentation d'une combinaison de ces deux allèles entre ces loci. Lors d'un balayage sélectif, les allèles en fort déséquilibre de liaison avec l'allèle sélectionné sont également présents à une fréquence très importante dans la population. La combinaison de l'allèle sélectionné et des allèles liés représente un haplotype. Le principe des tests décrits ici se base sur la longueur des haplotypes. Plus un allèle est sélectionné rapidement dans une population, moins il y aura de recombinaison autour de cet allèle et plus l'haplotype associé à l'allèle sélectionné sera large. Ainsi, un fort DL entre l'allèle sélectionné et les allèles l'entourant témoigne d'une forte sélection pour cet allèle. A la différence des tests précédents, les tests basés sur le DL permettent d'étudier des régions du génome qui ne sont pas totalement fixées dans la population, ce qui est souvent le cas pour des mutations récentes. La statistique EHH (Extended Haplotype Homozygosity) calcule la probabilité que deux chromosomes portant le locus observé choisis aléatoirement dans une population soient identiques par descendance autour du locus étudié (Sabeti et al., 2002). Ainsi, plus on s'écarte du locus étudié, plus cet EHH diminue, reflétant l'action de la recombinaison sur l'haplotype. Le test LRH (Long Range Haplotype) compare cet EHH à la fréquence de l'haplotype observé pour dénicher les haplotypes très fréquents dans la populations, signifiant une forte pression de sélection sur le locus étudié (Sabeti et al., 2002). L'iHS (integrated Haplotype Score, (Voight et al., 2006)) et le XP-EHH (Cross-Population EHH, (Sabeti et al., 2007)) sont basés sur le même principe que le EHH pour des applications particulières.

III.C.1.δ. Méthodes basées sur la différentiation des populations

Le principe de ces méthodes est d'utiliser l'indice de différentiation entre populations (F_{ST}, ou encore l'index de fixation de Wright) à l'échelle du génome (Lewontin et Krakauer, 1973; Excoffier et al., 2009; Bonhomme et al., 2010). Cet indice compare la variance des fréquences alléliques intra et inter populationnelles. Il est élevé lorsque les

populations sont différentes génétiquement au locus considéré et faible lorsque les deux populations sont similaires. Si on considère que deux populations séparées sont soumises à des pressions de sélections différentes, les fréquences allèliques observées dans ces populations sont censées refléter ces pressions de sélection. Par exemple, si un parasite est présent dans une population, on suppose que l'allèle conférant la résistance sera plus fréquent que dans une population non parasitée. Ainsi, une différence de fréquence allélique dans une population se traduit par un F_{ST} plus important au locus sélectionné dans une des populations par rapport aux autres loci du génome.

III.C.3. Quel type de sélection sur le système immunitaire?

Seuls trois organismes ont fait l'objet d'études approfondies sur le type de pression de sélection exercée spécifiquement sur les gènes de l'immunité: *Homo sapiens*, *Arabidopsis thaliana* et *Drosophila melanogaster*. Il a été montré chez *Drosophila melanogaster* que les gènes impliqués dans l'immunité font partie des gènes évoluant le plus rapidement (Obbard et al., 2006; Wayne et al., 2011). Des patrons de sélection équilibrante, synonyme du maintien de la diversité sur le long terme, ont été détectés pour les gènes du CMH de l'humain dont le rôle est de reconnaître les antigènes parasitaires et de les présenter à la surface de la cellule (Hedrick et Miller, 1994; Luo et al., 2012; Meng et al., 2012; Yasukochi et Satta, 2013). Cette sélection équilibrante serait principalement le reflet d'un avantage des hétérozygotes plutôt que d'une NFSD. De manière étonnante, de nombreux gènes impliqués dans les voies de signalisation (Interleukines, TNF), qu'on pourrait supposer comme très contraints, montrent également des patrons de sélection équilibrante (Casals et al., 2011). La sélection équilibrante pourrait être due au fait que ces gènes sont une des cibles principales de la contre-défense mise en place par les parasites. Ces gènes sont également impliqués dans les maladies auto-immunes, suggérant un compromis entre une défense efficace contre les parasites et une diminution de l'effet auto-immun. Chez les plantes, les loci associés à la production des gènes R chez les plantes semblent eux-aussi soumis à une sélection équilibrante (Bakker et al., 2006; Shen et al., 2006). Cette observation est surprenante puisqu'on s'attendrait à ce que les gènes R qui ont inspiré le modèle gène-pour-gène, soient plutôt la cible de la sélection directionnelle.

A l'inverse certains gènes sont soumis à une forte sélection directionnelle. C'est le cas de l'allèle *ref(3)D* conférant la résistance au Sigma virus chez *Drosophila melanogaster* et qui a rapidement envahi les populations naturelles (Magwire et al., 2011) .De même, chez le moustique *Anopheles gambiae*, le locus APL1 et le gène TEP1, impliqués dans la défense contre *Plasmodium falciparum*, montrent des patrons de sélection positive (Obbard et al., 2008; Rottschaefer et al., 2011).

Aucune corrélation entre un type de sélection et un type de facteurs (récepteur, protéine

de signalisation ou effecteur) impliqués dans la réponse immunitaire n'a été observée. Un des paramètres importants qui n'apparaît pas dans ces études est la diversité présente chez les parasites concernés par la réponse immunitaire étudié.

III. Conclusion

L'apport des NGS permet l'identification de l'architecture génétique de la résistance aux parasites. La base génétique de cette résistance implique généralement des loci à effet fort, suggérant un déterminisme génétique assez simple de ce caractère.
Les théories évolutives proposent deux modèles principaux permettant de décrire la coévolution entre hôtes et parasites. Le modèle gène-pour-gène propose une succession de balayages sélectifs avec, pour chacun des antagonistes, la fixation d'un allèle de virulence pour le parasite suivie par la fixation d'un allèle de résistance pour l'hôte. Le modèle "matching allele" prédit une conservation de la diversité initiale au cours du temps avec une sélection négative dépendante de la fréquence: un allèle rare sera favorisé dans la population d'hôte car moins sensible au parasite le plus commun et un allèle rare sera alors favorisé dans la population du parasite car il confère la capacité d'infecter le l'hôte le plus commun dans la population.
Ces deux modèles impliquent des marques de sélection différentes: une sélection directionnelle (disparition rapide de polymorphisme) pour le modèle gène-pour-gène et une sélection équilibrante (maintien de polymorphisme à fréquence intermédiaire) pour le modèle "matching-allele".
Aucun de ces deux modèles extrêmes n'est prépondérant à la vue des études effectuées sur les populations naturelles. Dans tous les cas, ces différents modèles prédisent une évolution de l'hôte vers la résistance à son parasite contemporain.

IV. Le nématode *Caenorhabditis elegans*

Caeorhabditis elegans fut isolé à partir du sol et décrit pour la première fois par Emile Maupas. Aujourd'hui, *C. elegans* est devenu un organisme modèle en biologie. Il a notamment permis de mettre en évidence de nombreux processus cellulaires et moléculaires généralement conservés chez les animaux. Durant les quinze dernières années, trois prix Nobel témoignent de l'intérêt majeur de l'utilisation de cet animal en biologie: en 2002 pour les travaux sur la génétique du développement et de la mort cellulaire programmée effectués par Sydney Brenner, Howard Robert Horvitz et John Sulston, en 2006 pour la découverte de l'ARN interférence à partir de doubles brins d'ARN par Craig Mello et Andrew Fire et enfin en 2008 pour la première expression de la protéine à fluorescence verte (GFP) dans un organisme pluricellulaire par Martin Chalfie. Comme chez tout organisme modèle, de nombreux outils ont été développés et son génome, le premier organisme multicellulaire à avoir été séquencé, est très bien annoté. Cependant, *C. elegans* reste avant tout un animal, qui interagit avec un environnement naturel très différent des conditions de laboratoire, et, par conséquent, la connaissance de son écologie et de son évolution est importante pour une compréhension intégrative de l'organisme. Le but de ce chapitre est de décrire l'organisme mais également de faire un bilan sur ce qui est connu de son écologie et son évolution. Nous nous attarderons plus particulièrement sur ses parasites et ses moyens de défense.

IV.A. Place de *C. elegans* dans le vivant

IV.A.1. Les nématodes dans l'arbre du vivant

C. elegans appartient au groupe des nématodes, l'un des embranchements des métazoaires. La caractéristique principale des nématodes est leur forme: ce sont tous des vers ronds non segmentés. Cependant, mis à part ce caractère, on observe une grande diversité dans leurs tailles, leurs formes ou leurs habitats. Les premières phylogénies étaient basées sur un nombre limité de caractères morphologiques observés par microscopie optique. Selon ces critères, les nématodes ont longtemps été considérés comme étant branchés avec les coelomates**.** L'utilisation de critères moléculaires permet dorénavant de fournir un grand nombre de caractères. C'est ainsi que la première étude moléculaire, basée sur la séquence de la petite sous-unité de l'ARN ribosomal (ARNr 18S) de *Trichinella spiralis*, place les nématodes parmi les ecdysozoaires (changement de stade larvaire par mue cuticulaire) (Aguinaldo et al., 1997). Pour cette étude,

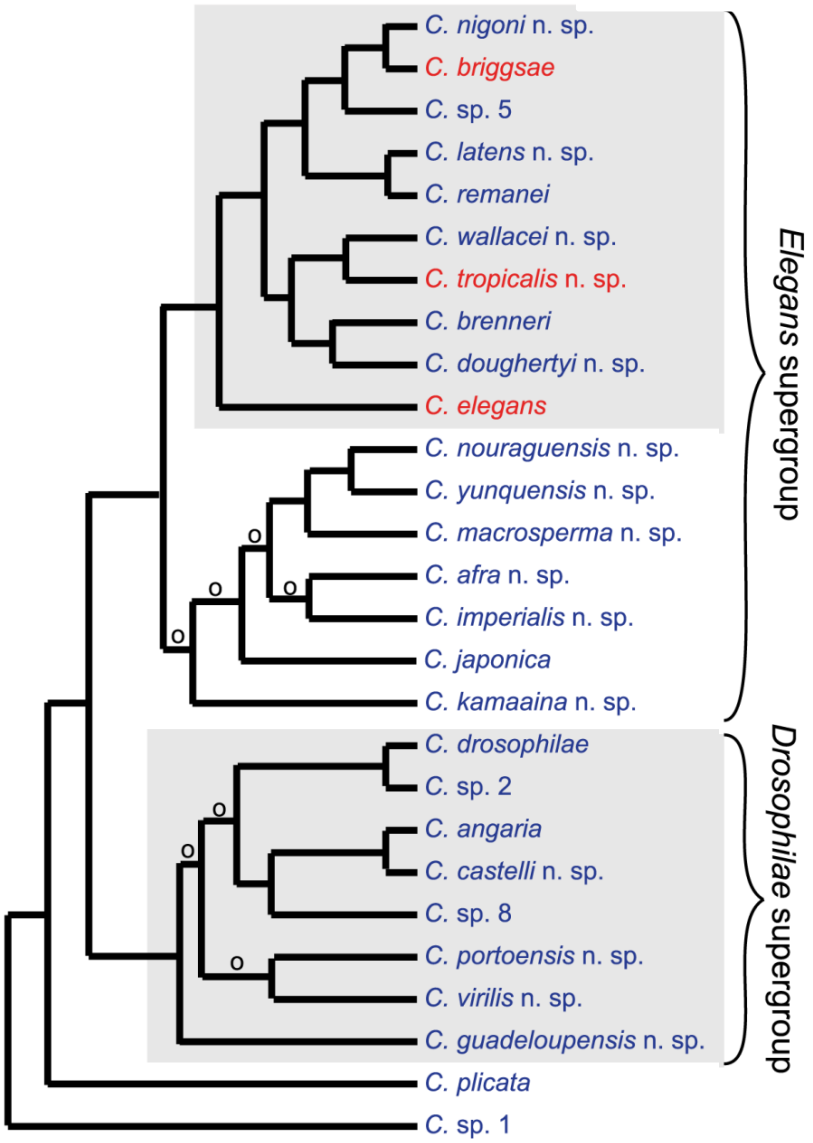


Figure 20. Phylogénie des espèces de *Caenorhabditis* cultivées en laboratoire.
En rouge, les espèces hermaphrodites et en bleu, les espèces gonochoriques. Sur fond gris, les espèces du groupe *Elegans* et du supergroupe *Drosophilae*. 'o' indique les branches peu supportées par l'analyse. Figure issue de Félix et al. (2014).

Trichinella spiralis fut choisit pour la faible vitesse d'évolution de la séquence de son ARNr 18S relative à celles des autres nématodes, évitant ainsi le biais d'attraction des longues branches (artefact d'analyse qui a pour conséquence de regrouper entre eux les taxons évoluant rapidement). L'addition d'autres nématodes, comme par exemple *C. elegans*, à ce type d'étude valida plutôt l'hypothèse d'un embranchement avec les coelomates. Finalement, l'ajout de nouvelles séquences et de taxons supplémentaires en tant que groupe externe tend à valider l'appartenance des nématodes aux ecdysozoaires (Philippe et al., 2005, 2005). Même si la place des nématodes parmi les eumétazoaires n'est pas encore définitivement déterminée, leur appartenance aux ecdysozoaires est l'hypothèse la plus largement acceptée.

IV.A.2. Phylogénie des nématodes

Environs 25.000 espèces de nématode, dont la plupart sont des parasites, ont dorénavant été décrites (Blaxter, 2003). Cependant, on estime entre 100 000 et 100 000 000 le nombre d'espèces existantes (Lambshead, 1993).
La phylogénie des nématodes a été établie à partir de la séquence des ARNr 18S de 53 espèces (Blaxter, 1998). En 2005, les séquences ARNr 18S de 600 nématodes étaient disponibles. La prise en compte de ces nouvelles données n'induit pas de changement majeur dans la topologie de l'arbre obtenu à partir des 53 espèces (Meldal et al., 2007). On peut regrouper ces espèces en cinq clades: Dorylaimia, Enoplia, Spirurina, Tylenchina et Rhabditinia. *C. elegans* appartient à ce dernier, à l'ordre des Rhabditida, au sous-ordre des Rhabditina, à la famille des Rhabditida et au genre *Caenorhabditis*.
Le genre *Caenorhabditis* regroupe à l'heure actuelle une quarantaine d'espèces réparties pour la plupart en deux super-groupes: *Elegans* et *Drosophilae* (nommés à partir d'une espèce emblématique de chacun des super-groupes) (Figure 20) (Kiontke et al., 2011).

IV.B. Anatomie et cycle de vie de *C. elegans*

IV.B.1. Anatomie des adultes

Deux types d'adultes coexistent chez *C. elegans*: les hermaphrodites (XX), s'autofécondant, sont prépondérants dans les populations naturelles tandis que les mâles (X0), ne pouvant féconder que les hermaphrodites, sont retrouvés à une fréquence très faible (Barrière et Félix, 2005; Barrière et Félix, 2007) et se maintiennent difficilement. Leur anatomie diffère principalement au niveau des caractères sexuels (qui seront décrits par la suite).

IV.B.1.α. Plan d'organisation

C. elegans possède un plan d'organisation caractéristique des nématodes: il s'agit d'un ver non segmenté. L'animal peut être séparé en deux parties, séparées par un pseudocoelome: la partie externe constituée de la cuticule, de l'hypoderme, du système excréteur, de la vulve, du système nerveux et des muscles; et la partie interne constitué du système digestif et des gonades. Un animal mesure environ 1 mm de longueur, pour un diamètre maximum de 80 μm (Figure 21).

IV.B.1.β. Partie entourant le pseudocoelome

i. La cuticule

Une cuticule de collagène de 0,5 μm d'épaisseur, sécrétée par les cellules constituant l'épiderme, recouvre entièrement l'animal. Certains tissus sont en contact direct avec l'extérieur et traversent la cuticule. C'est le cas de la bouche, du pore excréteur, de l'anus, de certains sensilla (neurones sensoriels), ainsi que de la vulve. Le pharynx et le rectum possèdent également des cellules sécrétrices de cuticule (moins épaisse et plus flexible). Cette cuticule joue un rôle de protection contre les agressions extérieures, un rôle de maintien de la forme de l'animal (exosquelette) et permet aussi sa mobilité. Elle est d'une importance capitale dans la protection contre les parasites puisqu'elle représente la première et principale barrière physique aux infections.

ii. L'épiderme

L'épiderme est responsable de la sécrétion de la cuticule. Il est composé d'un syncytium principal, appelé hyp7, parcourant l'animal dans sa longueur et de cellules hypodermiques plus petites au niveau de la tête et de la queue. Sur les parties latérales de la cuticule, on observe des lignes appelées alae. Ces alae sont sécrétées par les seam cells.

iii. Système nerveux

Les cellules du système nerveux s'organisent en ganglions au niveau de la tête et de la queue. Un animal hermaphrodite possède 302 neurones (contre 473 pour les mâles) répartis en deux systèmes nerveux indépendants: le système nerveux somatique (282 neurones) et un système nerveux pharyngé (20 neurones). Les neurones interviennent dans les comportements basiques de l'animal: le déplacement, la recherche de nourriture, la nutrition et la défécation. Ils interviennent aussi dans la perception de produits chimiques, d'odeurs, de la température et de la densité d'individus présents. Ces neurones sont donc importants pour la perception de son environnement par l'animal.

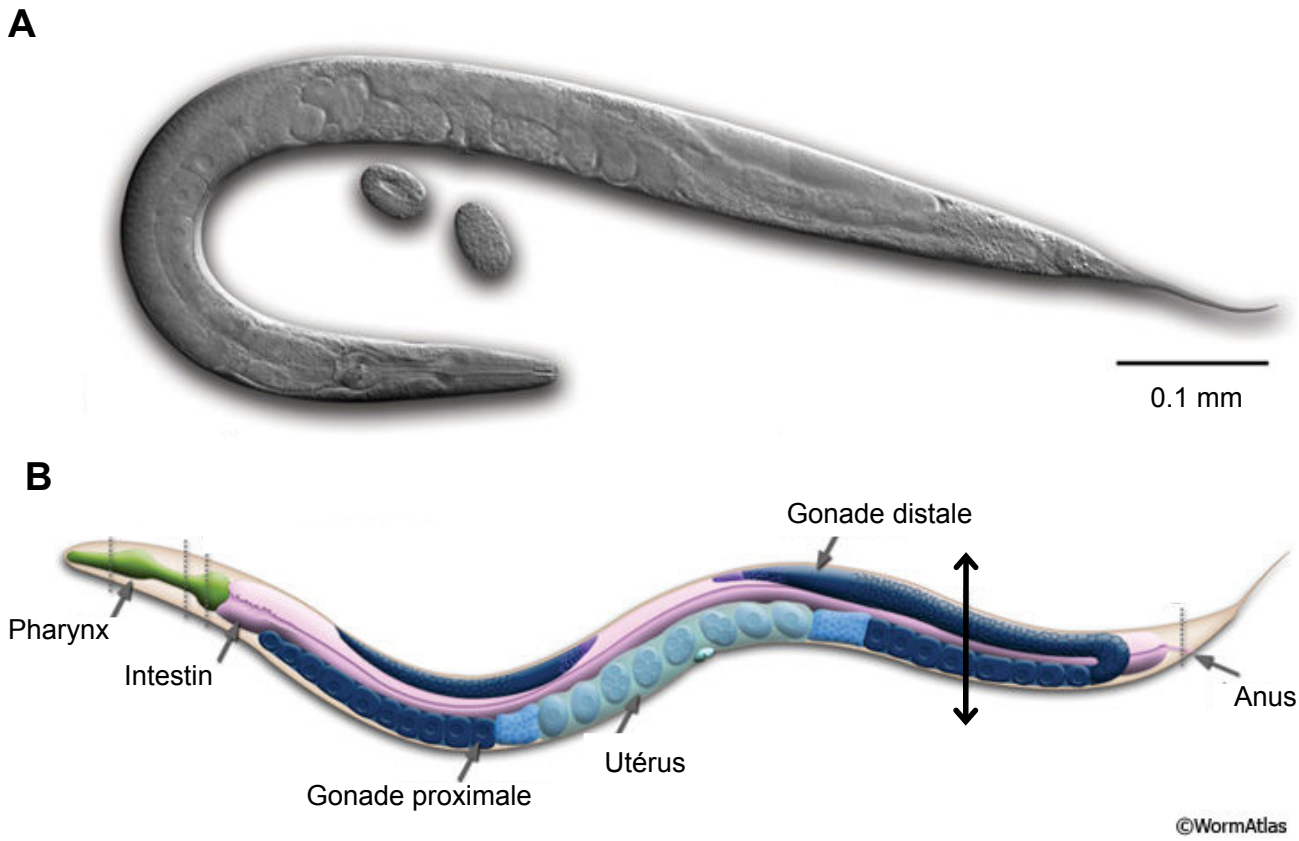


Figure 21. Anatomie de *C. elegans.* A: Image d'un hermaphrodite en microscopie optique, côté latéral gauche. **B:** Représentation schématique d'un hermaphrodite, côté latéral gauche. Chaque organe est coloré différemment. **C:** Coupe transversale d'un hermaphrodite, lieu de la coupe indiqué en **B**. D'après wormatlas (http://www.wormatlas.org).

iv. Le système excréteur

Quatre cellules situées sur la partie ventrale postérieures à la tête constituent le système excréteur. Ce système est impliqué dans l'osmorégulation et l'élimination des déchets.

IV.B.1.β. Pseudocoelome

Le pseudocoelome est une cavité séparant le tube externe du tube interne. Trois paires de coelomocytes y sont logées, adjacentes à la musculature somatique. Chez les plus gros nématodes, comme *Ascaris suum*, ces cellules sont capables de phagocyter des organismes étrangers de manière similaire aux macrophages présents chez les vertébrés ou aux hémocytes chez les insectes (Bolla et al., 1972). Chez *C. elegans*, malgré une activité d'endocytose avérée, aucune activité phagocytaire n'a été observée (Ewbank, 2002). Ces cellules sont, *a priori*, incapables de se mouvoir d'elles-mêmes. Pourtant, *C. elegans* supporte très mal la présence de seulement quelques bactéries pathogènes dans ce pseudocoelome (Ewbank, 2002). L'inaccessibilité de parasites au pseudocoelome pourrait avoir entraîné la perte de la fonction immunitaire de ces coelomocytes.

IV.B.1.γ. Partie entourée par le pseudocoelome

i. Le système alimentaire

Le système alimentaire est composé d'un pharynx, d'un intestin, d'un rectum et d'un anus (Figure 22).
Le pharynx est une pompe musculaire qui aspire, concentre et broie les bactéries en suspension. Il s'agit d'un organe d'origine épithéliale qui possède ses propres muscles, son propre système nerveux, des cellules glandulaires et des cellules structurales. Il mesure un dixième de la longueur totale de l'animal pour un diamètre de 20 μm dans sa partie la plus large. Sa partie antérieure est en contact avec la cavité buccale et sa partie postérieure est séparée de l'intestin par la valve pharyngo-intestinale. Son fonctionnement mécanique est indépendant du reste de l'animal avec lequel il communiquerait par l'intermédiaire d'hormones circulant dans le pseudocoelome. Le pharynx possède sa propre cuticule qui forme des structures spécialisées à certains endroits, comme pseudo-dents à l'entrée de la cavité buccale, des extensions cuticulaires capables de piéger les bactéries ou encore un broyeur au niveau du bulbe postérieur servant à les détruire (Avery et Thomas, 1997).
Le comportement de nutrition est le suivant: les particules nutritives, suspendues dans du liquide, sont piégées dans le pharynx. Le liquide est expulsé par la contraction du corpus et de l'isthme antérieur. Les particules sont ensuite transportées vers le bulbe terminal où elles sont broyées et transférées vers la lumière de l'intestin.

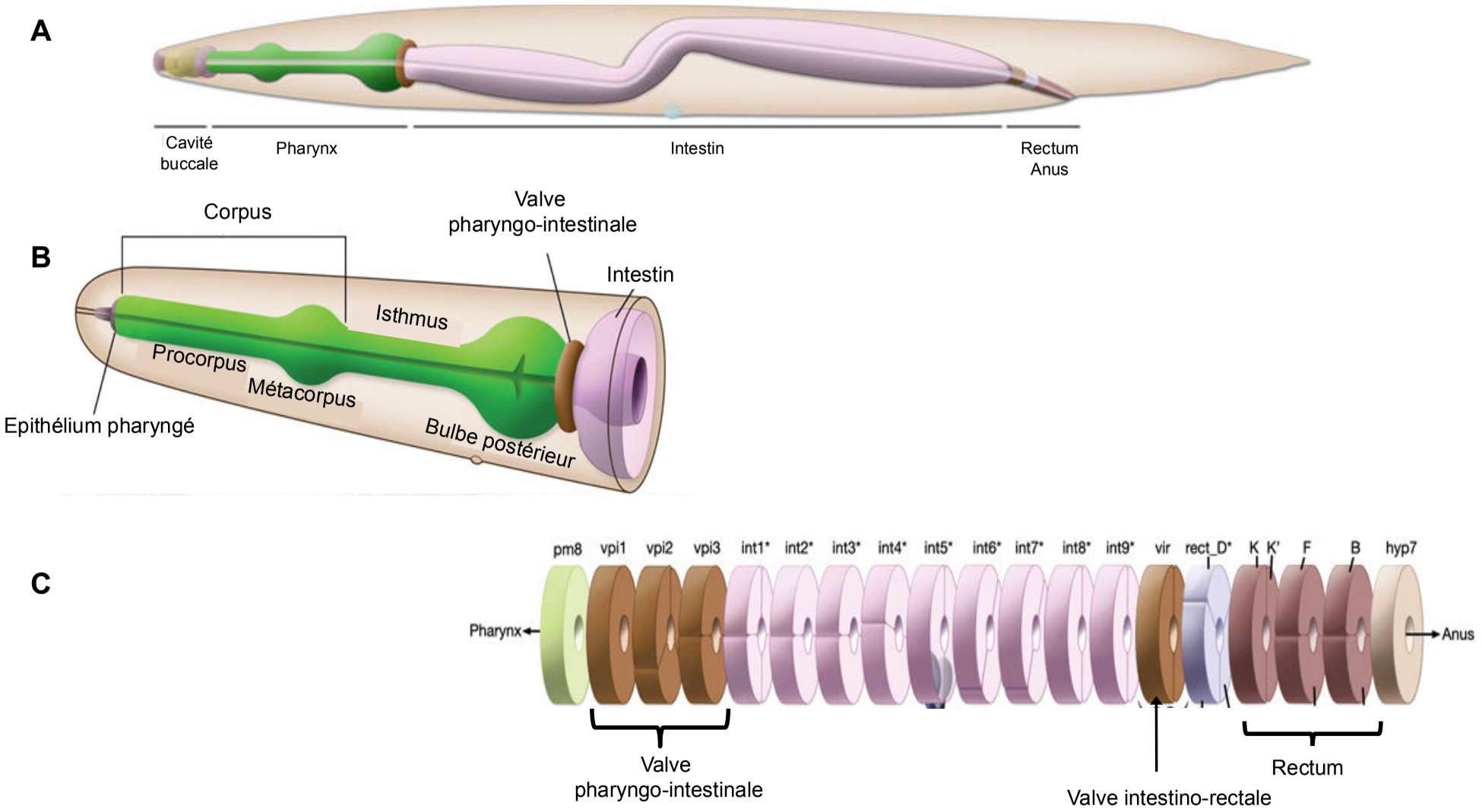


Figure 22. Appareil digestif chez *C. elegans*. A: les différentes parties du système digestif. **B:** schéma du pharynx. Le métacorpus est également appelé bulbe antérieur et le bulbe postérieur est parfois appelé bulbe terminal. La partie antérieure du pharynx communique avec la cavité buccale alors que la partie postérieure est reliée à l'intestin par la valve pharyngo-intestinale. **C:** schéma représentant l'organisation des vingt cellules intestinales (rose) autour du lumen pour un embryon de 480 heures. L'organisation reste sensiblement la même jusqu'à la mort de l'animal. D'après wormatlas (http://www.wormatlas.org).

L'intestin exerce de multiples fonctions chez *C. elegans*. Outre son rôle dans la digestion et l'absorption des nutriments, il synthétise et stocke des macromolécules, a un rôle majeur dans la réponse immunitaire aux parasites et transmet des nutriments aux cellules germinales (Kimble et Sharrock, 1983; McGhee, 2007; Schulenburg et al., 2004). L'intestin est composé de 20 cellules épithéliales. Ces cellules sont associées par paires (exception faite du premier anneau constitué de quatre cellules), formant un anneau autour de la lumière de l'intestin. Elles sont de grande taille, cuboïdes, et possèdent un pôle basal (sécrétion de lame basale) et un pôle apical, en contact avec la lumière de l'intestin, dont la membrane forme de nombreuses microvilli. Ces cellules sont reliées entre elles latéralement par des protéines de jonction proches du pôle apical (Labouesse, 2006) et communiquent entre elles par l'intermédiaire de jonctions gap.
Les cellules intestinales contiennent jusqu'à deux gros noyaux (30 à 34 noyaux pour 20 cellules), de nombreuses mitochondries, un réticulum endoplasmique granuleux imposant, de nombreux ribosomes et un nombre important de vésicules et de vacuoles. Les activités digestives et métaboliques sont assurées par de nombreux organelles tels que les endosomes, les vacuoles autophagiques ou encore les granules intestinaux. Les anneaux antérieurs (I et II) possèdent de grandes quantités de vésicules fusionnées à la membrane ainsi que de très nombreuses et diverses vacuoles, suggérant une sécrétion d'enzymes digestives plus importante que dans les anneaux postérieurs (Borgonie et al., 1995). En revanche, les cellules postérieures possèdent plus de vacuoles lipidiques que les anneaux antérieurs, suggérant un rôle important de celle-ci dans le stockage.
Le contenu intestinal non exploité est ensuite excrété vers l'extérieur de l'animal via la valve rectale qui relie l'intestin au rectum et à l'anus.

ii. Le système reproducteur des hermaphrodites

Le système reproducteur des hermaphrodites est composé de la gonade somatique, de la lignée germinale et de l'appareil de ponte. La gonade est constituée de deux bras symétriques formant un « U ». Elle se termine dans l'utérus en traversant la spermathèque. En position distale de la gonade, la lignée germinale forme un syncytium. Les cellules germinales entrent séquentiellement en étape de mitose, de prophase et de diacinèse de méiose au fur et à mesure qu'elles progressent vers la partie proximale de la gonade. Ensuite, les ovocytes s'élargissent, se détachent du syncytium et "mûrissent" tout en progressant vers l'utérus. Ils passent au travers de la spermathèque dans laquelle ils sont fécondés. Le zygote diploïde passe ensuite dans l'utérus et est pondu par la vulve.

iii. Le système reproducteur des mâles

La morphologie des mâles diffère de celle des hermaphrodites par la présence d'un appareil copulatoire à l'extrémité de la queue et l'absence de vulve. Cette structure en éventail permet l'injection des spermatozoïdes à l'intérieur de l'utérus de l'hermaphrodite,

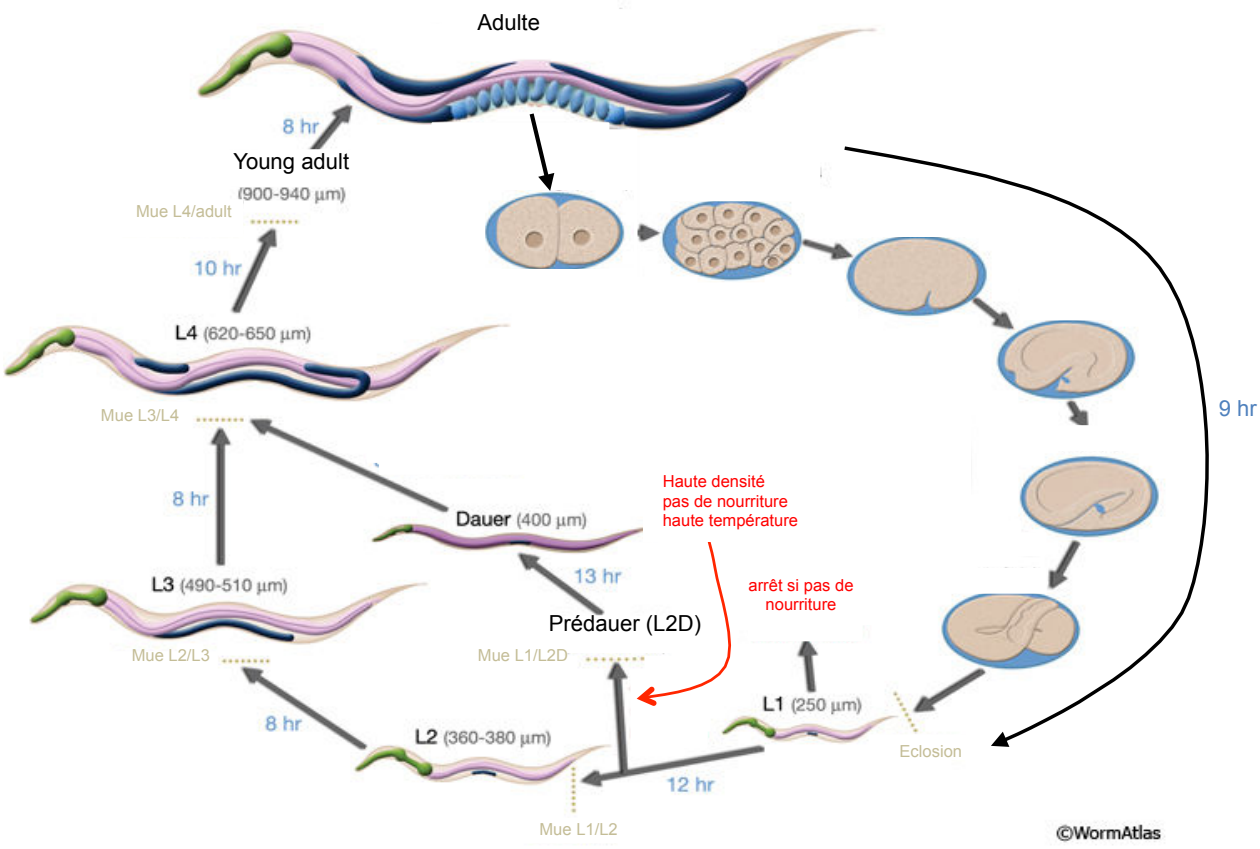


Figure 23. Cycle de vie d'un hermaphrodite de *C. elegans* au laboratoire à 20°C.
Voir texte. D'après wormatlas (http://www.wormatlas.org).

par la vulve. Le mâle ne possède qu'un seul bras de gonade et il possède des cellules musculaires et nerveuses supplémentaires, nécessaires à l'accouplement.

IV.B.2. Cycle de vie

Le cycle de vie de *C. elegans* est caractéristique de celui d'un nématode: un stade embryonnaire, quatre stades larvaires (L1 à L4) et le stade adulte. Entre chaque stade larvaire, l'animal entre dans une courte léthargie durant laquelle il mue. Chaque stade possède une cuticule spécifique. A 20°C, le temps de génération de *C. elegans* est de 3 jours (Figure 23).

IV.B.2.α. Développement embryonnaire

Il est constitué de deux étapes. La phase de prolifération cellulaire (0-350 minutes) consiste en l'obtention de 550 cellules à partir d'une seule. A la fin de cette étape, l'embryon est une masse sphéroïde de cellules organisées en trois couches: l'ectoderme, dont sont originaires l'hypoderme et les neurones; le mésoderme précurseur du pharynx et les muscles, et enfin l'endoderme, précurseur de l'intestin. L'embryon entre alors en phase d'organogénèse, exempte de division cellulaire, durant laquelle le plan d'organisation de l'animal va se mettre en place. L'éclosion de l'œuf a lieu 800 minutes après la fécondation. Le plan d'organisation est définitif et ne changera pas lors du développement post-embryonnaire.

IV.B.2.β. Développement post-embryonnaire

En présence de nourriture, le développement post-embryonnaire débute 3 heures après l'éclosion donnant naissance à une larve L1. En absence de nourriture, les L1 peuvent survivre 6 à 10 jours. Si la nourriture redevient disponible dans cet intervalle de temps, le développement post-embryonnaire peut reprendre normalement et l'animal atteint le stade L2. Si les conditions environnementales ne sont pas remplies (densité trop importante d'animaux, absence de nourriture, haute température) pour passer en L3, l'animal entre dans un stade L2dauer. Un maintien de ces mauvaises conditions induit l'entrée de l'animal en stade dauer au lieu de L3 (Cassada et Russell, 1975). Les dauers sont plus fines, possèdent une cuticule plus épaisse qui scelle la cavité buccale. Ce stade se termine lorsque les conditions favorables reviennent (dans une limite d'environ 6 mois). Les dauers accèdent directement au stade L4 durant lequel la gamétogenèse débute. Une dernière mue laisse place à un adulte fertile (hermaphrodite ou mâle) (Figure 23).
Environ 2,5 jours après son éclosion, à 20°C, l'animal est prêt à pondre son premier œuf. Durant une période de ponte longue de quatre jours, un hermaphrodite seul peut pondre

jusqu'à 300 œufs tandis qu'un hermaphrodite fécondé par un mâle peut en pondre jusqu'à 1400. Chez *C. elegans*, l'ovogénèse débute lorsque la spermatogénèse se termine. Ainsi, l'hermaphrodite produit des ovocytes jusqu'à sa mort alors que le nombre de spermatozoïdes est limité, l'injection de spermatozoïdes par le mâle augmentant le nombre de descendants.

IV.C. Ecologie et génétique des populations de *C. elegans*

C. elegans est reconnu aujourd'hui comme étant un animal modèle en biologie moléculaire et cellulaire, en développement, en neurologie et en physiologie. Facilement manipulable et observable, il constitue un modèle de choix pour des études d'évolution expérimentale (Gray et Cutter, 2014). Un des points noirs majeurs dans la compréhension de l'évolution de cet organisme est la connaissance de son écologie.

IV.C.1. Où trouver C. elegans*?*

La principale difficulté dans l'isolation de *C. elegans* réside dans la petite taille des individus car il est impossible de détecter directement les nématodes *in natura*. Il est donc nécessaire de prélever des substrats susceptibles de contenir des *C. elegans* et de les observer en laboratoire.
C. elegans a été initialement isolé à partir d'échantillons d'humus et a été catégorisé comme un nématode libre du sol. Dès lors, les premières tentatives d'isolation de *Caenorhabditis* sauvages ont été effectuées à partir d'échantillons d'humus et de tas de compost (Barrière et Félix, 2005; Cutter et al., 2006; Haber et al., 2005; Hodgkin et Doniach, 1997; Sivasundar et Hey, 2005; Petersen et al., 2014). Cependant, presque uniquement des dauers ont été isolées à partir de ce type de substrat. Les campagnes d'échantillonnages plus récentes ont révélé la présence de populations en prolifération (population composées de plusieurs stades) dans des fruits, des fleurs et des tiges en état de décomposition avancée ainsi que sporadiquement en association avec des invertébrés (Félix et Braendle, 2010; Félix et Duveau, 2012; Félix et al., 2013; Kiontke et al., 2011; Petersen et al., 2014) (Figure 24).
Depuis le début des années 2000, la collecte de *C. elegans* et plus généralement de *Caenorhabditis* s'est intensifiée (Haber et al., 2005; Hodgkin et Doniach, 1997; Barrière et Félix, 2005a; Sivasundar et Hey, 2005; Cutter, 2006; Félix et Braendle, 2010; Kiontke et al., 2011; Félix et Duveau, 2012; Félix et al., 2013; Petersen et al., 2014). On sait désormais que *C. elegans* se retrouve majoritairement en zone tempérée. Les quelques rares isolats provenant de zones tropicales sont issus de zones d'altitudes où la

température est plus basse.

IV.C.2. Quelles sont les meilleures conditions pour trouver C. elegans*?*

Au laboratoire, *C. elegans* est cultivé à des températures comprises entre 15°C et 25°C, le plus généralement à 20°C. Son développement s'arrête en deçà de 8°C et les animaux deviennent stériles au delà de 27°C. Des suivis temporels de la présence de *C. elegans* dans le verger semi-naturel de l'université d'Orsay et dans trois localisations au Nord de l'Allemagne (Kiel, Münster et Roxel) montrent que *C. elegans* est présent de Septembre à Décembre (Félix et Duveau, 2012; Petersen et al., 2014) lorsque l'humidité est assez importante. Ces résultats coïncident avec la présence de fruits en décomposition sur le sol du verger et une température correspondante à celles utilisées en laboratoire (Félix et Duveau, 2012). En revanche, *Caenorhabditis briggsae*, une des espèces les plus proches de *C. elegans*, se retrouve de Juin à Septembre dans le même type de substrat. On ignore cependant où et sous quelle forme subsiste *C. elegans* lorsque les températures ne sont pas optimales. *C. elegans* est parfois retrouvé associé à des escargots, connus pour hiberner dans le sol. On peut imaginer qu'à cet endroit, la température permet la subsistance de quelques animaux. L'existence de microenvironnements peut aussi favoriser la survie de quelques individus lors de la saison défavorable. Différents stades de développement montrent des capacités de résistance aux conditions défavorables. Le stade L1 est le stade utilisé pour conserver les souches à -80°C ou dans l'azote liquide au laboratoire et on peut très bien imaginer un arrêt développemental des individus en L1 dû au manque de nourriture lors de la saison froide et par conséquent une meilleure résistance au gel (Félix et Braendle, 2010). A l'inverse, le stade dauer est plus résistant aux températures élevées bien qu'également mis en place lors de l'absence de nourriture. Les conditions nécessaires à la mise en place de ces deux formes de résistance sont bien documentées au laboratoire mais restent cependant inconnues dans la nature.

IV.C.3. Génétique des populations

IV.C.3.γ. Diversité génétique

Outre une meilleure connaissance de son écologie, la collecte de populations naturelles nous donne accès au polymorphisme naturel de *C. elegans*. La diversité génétique naturelle chez *C. elegans* a donné lieu à plusieurs études employant des ressources génétiques différentes. Le séquençage de régions aléatoires du génome a permis d'obtenir un premier aperçu de la distance génétique entre N2, la souche standard de laboratoire, et d'autres souches sauvages d'origines géographiques diverses (AB1 d'Australie, CB4856 de Hawaï, CB4857 de Californie, RC301 d'Allemagne et TR403 du Wisconsin) (Koch et

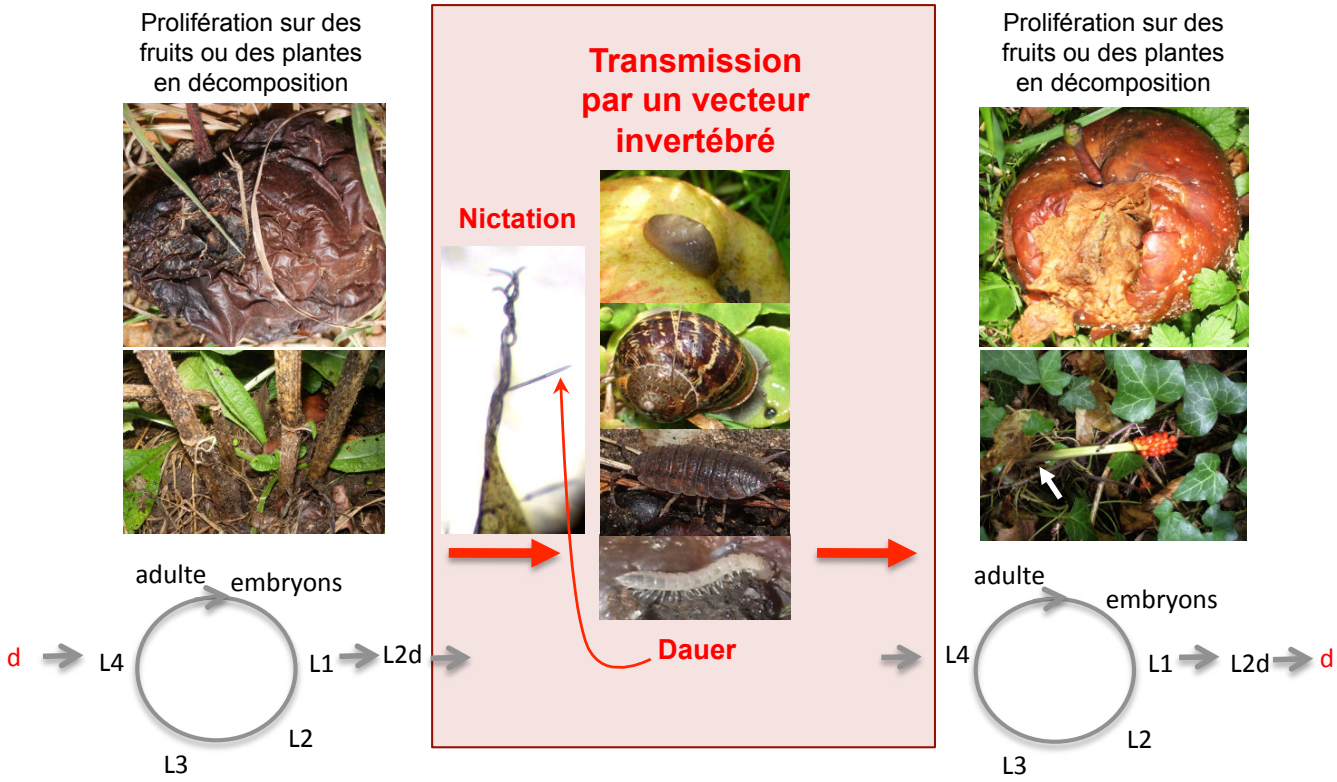


Figure 24. Habitat naturel et comportement migratoire chez *C. elegans*.
Les populations en prolifération sont trouvées principalement dans les fruits et les tiges en décomposition. Une fois le substrat épuisé en nourriture, le stade dauer devient prépondérant dans la population. Les dauers sont ensuite transportées par des vecteurs vers un autre substrat sur lequel la prolifération peut reprendre. Adapté de Braendle et Félix. (2010).

al., 2000; Swan et al., 2002; Wicks et al., 2001). La distance génétique révélée varie de $0,07x10^{-3}$ pour TR403 à $2,7x10^{-3}$ polymorphismes par paire de bases, soulignant une faible diversité génétique chez *C. elegans*. Cette information est confortée par des études ultérieures basées sur le polymorphisme de régions codantes pour un nombre d'isolats supérieur (11 à 20 isolats) (Graustein et al., 2002; Jovelin et al., 2003; Thomas et Wilson, 1991) ainsi que sur l'étude du polymorphisme de longueur de fragments amplifiés (AFLP) (Barrière et Félix, 2005) et de microsatellites (Haber et al., 2005; Sivasundar et Hey, 2003). Plus récemment, 97 isolats naturels de *C. elegans* ont fait l'objet d'une étude de diversité à partir de données obtenues par séquençage de l'ADN autour du site de restriction EcoRI (RAD sequencing) (Andersen et al., 2012). La diversité génétique observée varie selon les régions du génome, confirmant les précédentes études (Cutter et Payseur, 2003; Koch et al., 2000). Le taux de polymorphisme corrèle avec le taux estimé de recombinaison le long de chaque chromosome (Andersen et al., 2012; Rockman et Kruglyak, 2009). Ainsi, la diversité est plus importante au niveau des bras chromosomiques qu'au centre des chromosomes. Il n'en reste pas moins que toutes ces études s'accordent sur le fait que la diversité génétique est très faible chez *C. elegans*.

IV.C.3.β. Effectif efficace

La connaissance de la diversité génétique de différentes populations de *C. elegans* permet d'estimer l'effectif efficace de la population N_e qui représente le nombre d'individus d'une population idéale se reproduisant aléatoirement sous un régime de dérive génétique. Il est calculé à partir de π, diversité moléculaire, et μ le taux de mutation par la relation suivante (Wright, 1931):

$$\pi=4\ N_e\ \mu$$

Ainsi, à μ constant, N_e peut être directement estimé à partir de π. Les études d'AFLP proposent un N_e de 9600 pour quatre échantillons provenant de France (Barrière et Félix, 2005) alors que les études basées sur les microsatellites propose un N_e compris entre 200 et 44 000 (Sivasundar et Hey, 2003). Ces nombres sont relativement faibles comparés à la quantité d'individus retrouvés dans les substrats prélevés: jusqu'à 10 individus/g de compost (Barrière et Félix, 2005) et de 1 à 10 000 individus par fruits en décomposition (Félix et Duveau, 2012). Ce faible N_e peut être dû au fait que l'espèce soit quasi exclusivement hermaphrodite, à un assortiment non-aléatoire des gamètes dû à une structure géographique de la diversité ou à une très faible diversité génétique engendrée par de nombreux épisodes de balayages sélectifs.

IV.C.3.δ Structure des populations

Que ce soit par RAD sequencing (Andersen et al., 2012), par l'étude d'AFLP (Barrière et

Félix, 2005), par l'étude de microsatellites (Sivasundar et Hey, 2005) ou tout simplement par le fait que certains génotypes de *C. elegans* sont retrouvés sur différents continents très éloignés (Andersen et al., 2012; Koch et al., 2000), les données s'accordent pour infirmer l'existence d'une structuration géographique mondiale forte des populations de *C. elegans*. Ainsi, il apparaît que les populations locales, dont la structure spatiale est très forte, sont presque aussi diverses que la population mondiale prise dans son intégralité. Cette absence de structure peut s'expliquer soit par un fort taux de migration soit par une grande abondance du nombre de génotypes présents lors des évènements fondateurs de chacune des populations. Une faible corrélation entre la distance géographiques et la distance génétique des isolats est absente à l'échelle mondiale mais existe, même si très faible, à l'échelle locale (Andersen et al., 2012). Une grande capacité migratoire, expliquée par une grande capacité de résistance de certains stades, pourrait être responsable de l'absence de structure mondiale de *C. elegans*. Par conséquent, le faible effectif efficace de la population mondiale serait principalement dû à l'hermaphrodisme de *C. elegans*.

IV.C.3.δ. Sélection à l'échelle du génome

L'étude des 97 isolats révèle que de larges haplotypes sont communs entre la plupart des souches. Plus particulièrement, 70 à 90% des isolats partagent des régions génomiques de quelques Mb au centre des chromosomes I, IV, V et X. Ces régions à forte hétérozygotie haplotypique sont également les régions présentant un excès de SNP rares (Andersen et al., 2012) (Figure 25). Ce type d'observation est caractéristique d'un balayage sélectif dans une population avec un fort taux de déséquilibre de liaison. Cela signifie que ces grandes régions ont été positivement sélectionnées récemment. Une des hypothèses voudrait que ces balayages sélectifs soient dus à l'adaptation de *C. elegans* aux activités humaines.

L'utilisation des tests de sélection dépend bien évidemment de l'architecture du génome de l'espèce étudiée. Chez *C. elegans*, dont le mode de reproduction quasiment exclusivement hermaphrodite, le déséquilibre de liaison est très important dans le génome et notamment au centre des chromosomes (Andersen et al., 2012). Associé à cela, de nombreux balayages sélectifs récents sont décelables dans le génome et notamment au centre des chromosomes. Ces propriétés rendent assez délicates les tests basés sur la longueur des haplotypes, surtout si le locus étudié se trouve au centre des chromosomes. Il s'avère également que *C. elegans* est particulièrement isolé dans la phylogénie des *Caenorhabditis* (l'espèce la plus proche est très éloignée et il est par conséquent difficile d'évaluer le nombre de substitutions ayant eu lieu à un locus neutre) (Kiontke et al., 2011), rendant ainsi obsolète les tests comparant le polymorphisme inter et intra spécifique comme le test Hudson-Kreitman-Aguadé (Hudson et al., 1987).

Les études précédentes ont beaucoup apporté à la compréhension de l'évolution moléculaire des populations naturelles de *C. elegans*. Il reste cependant nécessaire d'identifier quelles pourraient être les différentes pressions de sélections auxquelles peuvent être soumis les populations naturelles pour comprendre les marques de sélection visibles à l'échelle du génome. Les pressions de sélection les plus fortes proviennent entre autres des interactions avec les autres organismes.

IV.C.4. Interactions avec les autres organismes

IV.C.4.α. Source de nourriture

Au laboratoire, *C. elegans* se nourrit d'*Escherichia coli*, suggérant que la nourriture principale est constituée de bactéries présentes dans le substrat duquel sont extraits les animaux. L'identité de ces bactéries est inconnue et il reste techniquement délicat de les isoler à partir de l'intestin des animaux afin de les identifier.

IV.C.4.β. Prédateurs

Aucune étude particulière n'a recherché précisément les prédateurs potentiels de *C. elegans*. Néanmoins, la collecte des nématodes a révélé certains d'entre eux. Parmi les espèces nématophages, ont retrouve des acariens, des champignons ainsi que d'autres nématodes(Félix et Braendle, 2010). Les champignons, comme par exemple ceux du genre *Harposporium*, peuvent présenter des structures telles que des lassos ou des hyphes adhésifs capables de piéger le nématode.

IV.C.4.γ. Vecteurs

Comme précisé plus haut, *C. elegans,* plus particulièrement au stade dauer, est parfois associé à des invertébrés tels que des limaces, des escargots, des isopodes, des myriapodes et probablement des insectes. Le stade dauer est le seul stade faisant preuve d'un comportement associé à la migration. Une fois au sommet d'une structure filamenteuse, elles se redressent et se balancent (Lee et al., 2012a). De plus, elles sont capables de s'enrouler les unes autour des autres tout en conservant ce mouvement de balancement. Ce phénomène pourrait maximiser les chances d'être "pris en stop" par un animal, permettant ainsi une dispersion sur une longue distance (Félix et Braendle, 2010; Félix et Duveau, 2012).

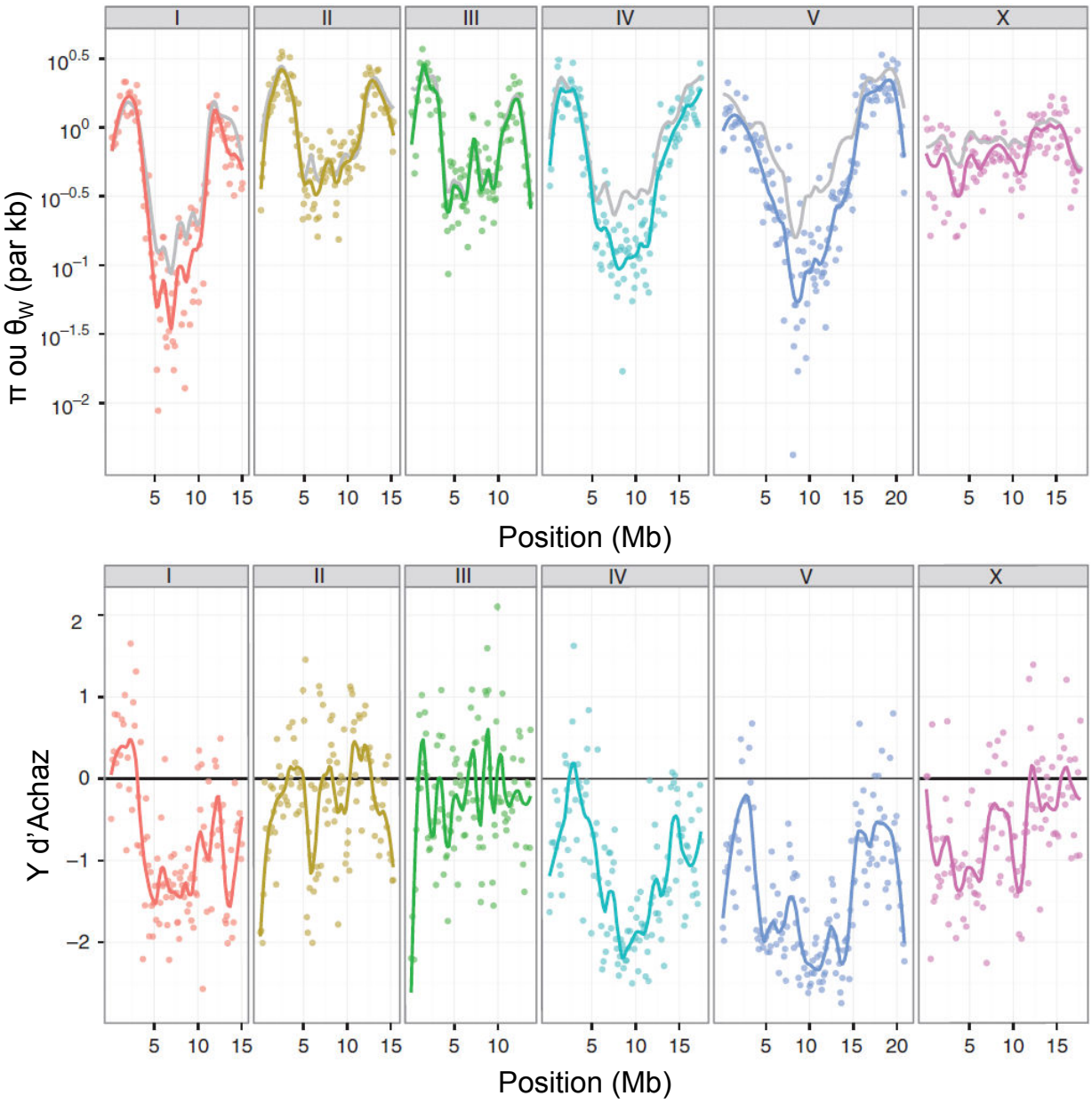


Figure 25. Polymorphisme le long des chromosomes chez *C. elegans*. A: Deux estimateurs du taux de polymorphisme dans les populations: π (points et lignes colorés) et θ_w. Chaque point représente une fenêtre de 110 marqueurs de RAD. La ligne représente la régression. **B:** Y d'Achaz, une mesure de la déviation par rapport à la neutralité. Les valeurs négatives indiquent un excès d'allèles rares. Issu de Andersen et al. (2012).

IV.C.4.δ. Parasites naturels

Peu de parasites infectant naturellement *C. elegans* ont été décrits dans la littérature.

i. Microbacterium nematophilum

Cette bactérie est issue d'une contamination en laboratoire. Le phénotype qu'elle provoque, à savoir une déformation de la région anale de son hôte, a longtemps été considéré comme étant une mutation apparue au laboratoire (Hodgkin et al., 2000). L'observation en microscopie optique à haute résolution a révélé la présence de bactéries adhérentes à la cuticule, positionnées au niveau du rectum de l'animal. La réponse de l'hôte induit un gonflement de la région en présence des bactéries (Gravato-Nobre et al., 2005). Bien que non létale, ces bactéries provoquent un ralentissement du développement ainsi qu'une constipation chez leur hôte (Hodgkin et al., 2000).

ii. *Leucobacter* Verde1 et Verde2

Les phénotypes associés à *Microbacterium nematophilum* ont été retrouvés par la suite dans un isolat naturel de *C. elegans* provenant du Japon et dans un isolat naturel de *Caenorhabditis tropicalis* provenant du Cap Vert (Hodgkin et al., 2013, 20013). *C. tropicalis* était en fait infecté par deux *Leucobacter*: Verde1 et Verde2. Ces deux bactéries peuvent être létales pour leur hôte et pour la souche de laboratoire N2 chacune à leur manière. Dans des conditions standard de culture, Verde1 seule provoque une diminution de la locomotion mais n'empêche pas la croissance ni la reproduction de l'hôte. Cependant, en conditions liquides, l'ajout de Verde1 induit la formation surprenante d'étoiles de vers: les individus sont collés les uns aux autres à l'extrémité de leur queue (Figure 26C). Ensuite, les bactéries pénètrent et prolifèrent dans l'animal causant ainsi sa mort. La souche Verde2 seule provoque un phénotype similaire à celui provoqué par *M. nematophilum* mais tue son hôte beaucoup plus rapidement que cette dernière (Figure 26C). De manière surprenante, la souche co-infectée de *C. tropicalis* et la souche de *C. elegans* co-infectée au laboratoire ne présentent pas un taux de mortalité très élevé, suggérant une meilleure tolérance de la co-infection que des infections simples (Hodgkin et al., 2013, 20013).

iii. *Nematocida parisii*

Nematocida parisii est le premier parasite intracellulaire connu naturellement infectant *C. elegans*. Il s'agit d'une microsporidie, un champignon parasite unicellulaire fréquemment retrouvé dans les isolats naturels de *C. elegans* (Troemel et al., 2008). Comme son nom l'indique, l'infection par ce parasite est létale pour le ver. L'infection de *C. elegans* débute par l'ingestion de spores de *N. parisii*. Une fois dans la lumière de l'intestin, les spores produisent des tubes polaires qui servent à injecter le noyau et le sporoplasme à

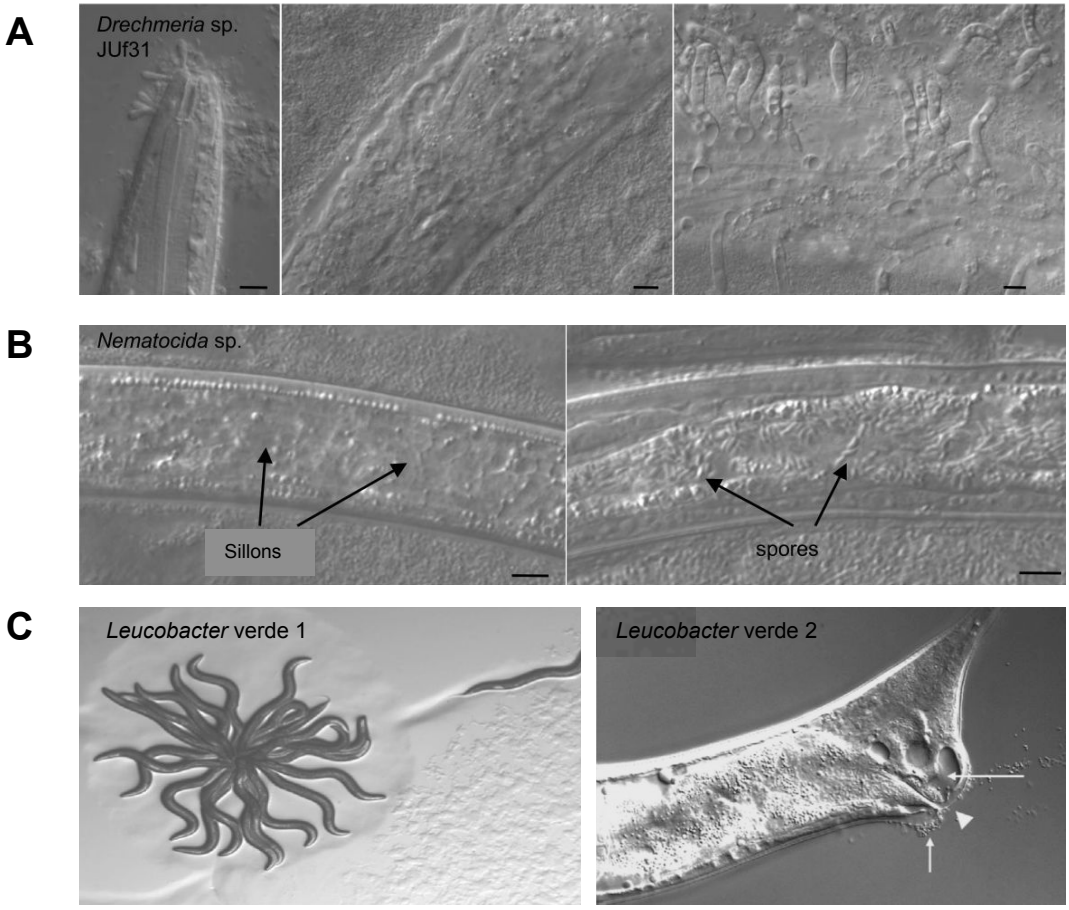


Figure 26. Parasites naturels de *C. elegans*. A: Infection par *Drechmeria coniospora.* Les spores s'attachent à la cuticule par l'intermédiaire des boutons adhésifs. Cette spore traverse la cuticule et pénètre dans l'organisme. Il forme ensuite une hyphe capable de produire des spores. Image issue de Félix et Duveau, 2012 **B:** Infection par une microsporidie. Une spore microsporidienne injecte son contenu dans la cellule intestinale. Une structure appelée méronte se forme et permet la production de spores qui sortent ensuite de la cellule de manière non lytique. Image issue de Troemel et al. (2008) **C:** Image de gauche: infection par la bactérie *Leucobacter* Verde1. Aucun phénotype n'est visible en milieu solide. En revanche, en milieu liquide, ces bactéries induisent la formation d'une étoile de vers. Cette formation est due à la forte capacité d'adhésion de la bactérie qui piège les vers. Image issue de Hodgkin et al. (2013). Image de droite: infection par la bactérie *Leucobacter* Verde2. Ces bactéries se retrouvent au niveau du rectum de l'animal et provoquent un gonflement de celui-ci. Le phénotype observé lors d'une infection par *Microbacterium nematophilum* est similaire. Image issue de Hodgkin et al. (2013).

l'intérieur des cellules intestinales. S'en suit la formation de la structure de multiplication appelée méronte. Ce méront peut se différencier en spores (Figure 26B). Une réorganisation du réseau d'actine de la cellule intestinale permet une sortie non lytique des spores dans le lumen et une propagation par défécation de l'hôte (Troemel et al., 2008).

iv. *Drechmeria coniospora*

Drechmeria coniospora est un champignon parasitant spécifiquement les nématodes (Jansson, 1994). Les conidies (spores des champignons) s'attachent à la cuticule des nématodes par l'intermédiaire de boutons adhésifs (Figure 26A) et un appressorium perce ensuite la cuticule. Ceci permet à une hyphe de pénétrer à l'intérieur de l'animal et de le coloniser. L'hyphe croît et transperce la cuticule de son hôte depuis l'intérieur. Les spores sont ensuite produites, permettant la transmission du parasite (Félix et Duveau, 2012).

v. Le rétrotransposon *Cer1*

Le rétrotransposon *Cer1* a été identifié initialement grâce à sa similitude de séquence avec les rétrotransposons gypsy présents chez *D. melanogaster*(Britten, 1995). Cependant, les preuves de son activité n'ont été révélées que récemment (Dennis et al., 2012). Lorsque *C. elegans* est cultivé à 15°C, température peu utilisée pour des expériences en laboratoire, des particules similaires à celles produites par des virus sont observées en microscopie électronique dans certaines cellules germinales. Ces particules, assimilables à des capsides virales, sont le produit de l'expression du rétrotransposon *Cer1*. Elles modifient le réseau de microtubules et l'utilisent pour migrer vers le noyau lors des phases pachytène et diplotène de la méïose des cellules germinales (Dennis et al., 2012).

Il est bien évidemment impossible que la liste précédente fasse état de manière exhaustive des différents parasites associés à ce nématode. La collecte de nombreux isolats naturels de *C. elegans* a déjà permis d'en isoler de nombreux et le laboratoire de Marie-Anne Félix possède une base de données les recensant. Cependant, ces parasites ne sont pas encore bien caractérisés et/ou identifiés même si certains ont été cités (Félix et Duveau, 2012).

IV.D. Systèmes de défense chez *C. elegans*

De nombreux parasites de l'homme, de plantes ou d'autres animaux provoquent

également des dégâts chez *C. elegans*. Remplacer la source de nourriture habituelle par le parasite d'intérêt suffit généralement à déclencher une infection chez *C. elegans*. De plus, la transparence de l'hôte, son anatomie relativement simple et la connaissance poussée de sa biologie rendent plus aisée l'observation du phénomène d'infection et de défense que pour les hôtes naturels. L'existence de nombreux mutants et la possibilité d'effectuer de la mutagénèse chez *C. elegans* permettent également l'identification des mécanismes impliqués dans la réponse immunitaire, parfois conservés chez l'homme. A ce jour, une quarantaine d'espèces de parasites a été testée sur des cultures de *C. elegans* (Sifri et al., 2005), en faisant ainsi un modèle pour l'étude de l'immunité innée (Tableau 2).
Deux principales voies d'entrée peuvent être empruntées par les parasites lors d'une infection: soit le parasite passe par des les ouvertures de la cuticule (en général par la bouche), soit le parasite parvient à se créer un passage par lui même à travers la cuticule. Par conséquent, en absence de cellules spécifiquement impliquées dans la réponse immunitaire, ce sont l'intestin et l'hypoderme qui assurent directement le rôle de défense (Irazoqui et Ausubel, 2010). Certains éléments conservés entre les insectes et les mammifères sont absents chez *C. elegans*: une seule protéine Toll, exprimée dans certains neurones, a été identifiée (Pujol et al., 2001), tandis que le facteur de transcription NF-κB est complètement absent, suggérant l'existence de voies alternatives. Nous passerons ici en revue les différents mécanismes impliqués dans la réponse immunitaire dans les différents tissus de l'animal.

IV.D.1. Défenses intestinales

La majorité des études visant à décrire les mécanismes immunitaires dans l'intestin de *C. elegans* utilisent des infections par des parasites non-naturels (*Pseudomonas aeruginosa*, *Enterococcus faecalis*, *Staphylococcus aureus*, etc…). L'infection de l'intestin se caractérise généralement par une accumulation de bactéries intactes dans le lumen (Sifri et al., 2005). De plus, la majorité des changements transcriptionnels induits par une infection semblent avoir lieu dans l'intestin (O'Rourke et al., 2006; Alper et al., 2007; Wong et al., 2007; Irazoqui et al., 2008; Engelmann et al., 2011), suggérant un rôle essentiel de cet organe dans la réponse immunitaire (Figure 27C et 27D).

IV.E.1.α. Voies de signalisation dans les cellules intestinales

Les protéines NSY-1, SEK-1 et PMK-1 sont impliquées dans une cascade de phosphorylation MAPK (Mitogen Activated Protein Kinase) similaire à celle des mammifères et semblent être un des composants essentiels de la réponse immunitaire intestinale (Huffman et al., 2004; Kim et al., 2002; Kim et al., 2004). Malgré l'absence de protéine Toll à la surface des cellules intestinales, une protéine TIR (Toll, interleukin

Espèce	Réservoir naturel	Hôte naturel	Mode d'infection *C. elegans*	Temps moyen de survie de *C. elegans*
Bactéries à Gram- des mammifères				
Acinetobacter baumannii	Sol et Eau	Opportuniste de l'Homme	I	5–7j
Aeromonas hydrophila	Eau	Poissons et opportuniste de l'Homme	n.d.	6–8j
Burkholderia cepacia	Sol et Eau	Opportuniste de l'Homme	T+IP	20h
B. mallei c	Tractus respiratoire des chevaux	Chevaux	?	20h
B. pseudomallei c	Sol et Eau	Opportuniste de l'Homme	T+I	20h
Escherichia coli	Tube digestif des animaux	Opportuniste des mammifères	I	4j
Pseudomonas aeruginosa	Sol et Eau	Opportuniste de l'Homme	T+I	4h-3j
Salmonella enterica serovar Dublin c	Tube digestif des animaux	Homme	?	4j
S. enterica serovar Enteritidis c	Tube digestif des animaux	Homme	?	5j
S. enterica serovar Typhimurium c	Tube digestif des animaux	Homme	IP	4–7j
Serratia marcescens	Sol et eau, tube digestif et tractus respiratoire des animaux	Opportuniste de l'Homme	IP	2–6j
Yersinia pestis d	Rongeurs	Mammifères	B	?
Y. pseudotuberculosis	Tube digestif des animaux	Mammifères et oiseaux	B	?
Bactéries à Gram- des végétaux				
Agrobacterium tumefaciens	Sol et Eau	Végétaux et opportuniste de l'Homme	?	7j
Erwinia chrysanthemi	Sol et Eau	Végétaux	?	3j
Pectobacterium carotovora subsp. *Carotovora*	Sol et Eau	Végétaux	?	5j
Autres bactéries à Gram-				
Burkholderia thailandensis	Sol et Eau	?	T+IP	10–42h
Photorhabdus luminescens	Nématodes	Insectes	?	2j
Shewanella oneidensis	Mer, poissons décomposés	Opportuniste rare de l'Homme	?	7j
Shewanella massilia	Mer, poissons décomposés	?	?	8j
Xenorhabdus nematophilus	Nématodes	Insectes	?	5j
Bactéries à Gram+, faible taux en G+C				
Bacillus megaterium	Sol	?	T	15m
Bacillus thuringiensis	Sol	Nématodes et insectes	T	?
Enterococcus faecalis	Tube digestif des animaux	Opportuniste de l'Homme	IP	3–7j
E. faecium	Tube digestif des animaux	Opportuniste de l'Homme	T	4–6h
Staphylococcus aureus	Tube digestif des animaux	Opportuniste de l'Homme	I	1,5–5j
Streptococcus dysgalactiae	Muqueuses animales	Mammifères	T	?
S. mitis	Cavité buccale de l'Homme	Opportuniste de l'Homme	T	?
S. oralis	Cavité buccale de l'Homme	Opportuniste de l'Homme	T	?
S. pneumoniae	Tractus respiratoire des animaux	Homme	T	6j
S. pyogenes	Muqueuses humaines	Homme	T	6j
Streptococcus, Group G	Muqueuses animales	Mammifères	T	?
Bactéries à Gram+, fort taux en G+C				
Leucobacter Verde1	Sol	*Caenorhabditis*	V+B	1j si piégé
Leucobacter Verde2	Sol	*Caenorhabditis*	V+B	1j
Microbacterium nematophilum	?	*C. elegans*	B	?
Streptomyces albireticuli	Sol	Nématodes	V	3h
S. avermitilis	Sol	Nématodes	T	?
Champignons				
Cryptococcus neoformans	Sol	Opportuniste de l'Homme	T+I	4–8j
Drechmeria coniospora	Sol	Nématodes	V	4–5j
Paecilomyces lilacinus	Sol	Nématodes et opportuniste de l'Homme	T	6–24h
Nematocida parisii	Sol	Nématodes	IC	5j

Tableau 2. Parasites capables d'infecter *C. elegans.* Les parasites naturels sont représentés en vert. Les différentes valeurs pour le mode d'infection sont: B pour production de biofilm; I pour colonisation de l'intestin; IC pour infection intracellulaire; IP pour colonisation durable de l'intestin; T pour production de toxines et V pour invasif (destruction de la cuticule). Les unités pour le temps moyen de survie sont: m pour minute(s); h pour heure(s); j pour jour(s). Tableau adapté de Sifri et al. (2005).

receptor), interagissant avec Toll chez les insectes et les mammifères, agit en amont de la voie MAPK (Liberati et al., 2004). TIR-1 n'est pas activée directement par un récepteur membranaire mais par une autre cascade de phosphorylation impliquant TPA-1 et DKF-2 (Ren et al., 2009). Le récepteur activant cette voie reste cependant inconnu. Il semblerait néanmoins que la voie déclenchée par le récepteur FSHR-1 (protéine G) et la voie MAPK précédemment décrite régulent indépendamment l'expression de gènes de résistance de manière commune (Powell et al., 2009). Le ligand reconnu par FSHR-1 est pour le moment inconnu du fait de l'absence de ligand similaire à la FSH des mammifères chez *C. elegans*. La voie PMK-1 est également impliquée dans la réponse aux protéines dépliées (UPR) effectuée par le réticulum endoplasmique suite à l'infection par des bactéries sécrétant des toxines qui perméabilisent la membrane des cellules intestinales (Barrows et al., 2007; Haskins et al., 2008).
L'activation de la voie de l'insuline par DAF-2/InsR permet la répression de gènes impliqués dans la réponse immunitaire par l'activation de DAF-16/FOXO (Murphy et al., 2003). Néanmoins, cette voie n'est pas spécifique d'une infection par un parasite puisqu'elle est également activée en cas de stress cellulaire.

IV.D.1.β. Effecteurs de l'immunité produits dans l'intestin

Ces voies de signalisation induisent la production de protéines antimicrobiennnes.

- Les lectines de type C

Les protéines possédant des domaines lectines de type C se fixent à des carbohydrates spécifiques des surfaces bactériennes (Schulenburg et al., 2008). Chez *C. elegans*, il existe près de 265 lectines de type C, suggérant une diversification lié au parasitisme. Cependant la manière dont elles interagissent avec le parasite reste incomprise.

- Les caenopores

Les caenopores (ou SPP pour saposin-like protein) sont des protéines formant des pores dans les membranes des bactéries (Roeder et al., 2010). Elles partagent des caractéristiques structurelles et fonctionnelles avec les amoebapores et certaines protéines sécrétées par les cellules cytotoxiques des vertébrés.

- Les lysozymes

Les lysozymes sont des protéines sécrétées dans la lumière de l'intestin où elles agissent directement sur les bactéries (O'Rourke et al., 2006).

- Les Facteurs antimicrobiens

Les facteurs antimicrobiens (ABF) détruisent la membrane des bactéries à Gram- chez

les nématodes (Zhang et al., 2000). Six ABF ont été découverts chez *C. elegans* dont ABF-1 et ABF-3 qui sont produites constitutivement par l'intestin, constituant une des premières défenses intestinales du nématode (Alper et al., 2007). La surexpression de ces ABF semble dépendre spécifiquement du parasite rencontré. Ainsi ABF-1 est surexprimée en présence de *Cryptococcus neoformans* (Means et al., 2009) tandis qu'ABF-3 est surexprimée lors d'une infection par *Staphylococcus aureus* (Alper et al., 2007), suggérant une spécificité de la réponse immunitaire intestinale.

IV.D.2. Défenses hypodermiques

D'autres parasites sont capables d'infecter *C. elegans* par perforation de la cuticule comme par exemple *Drechmeria coniospora* (Jansson, 1994; Pujol et al., 2008a; Félix et Duveau, 2012). En réponse à cette perforation ou à une blessure quelconque de l'hypoderme, ce dernier monte une défense immunitaire (Figure 27A et 27B).

IV.D.2.α. Voies de signalisation de l'hypoderme

La cascade de phosphorylation MAPK est encore ici impliquée dans le signal. Cependant, les composants en amont de la voie sont plus nombreux. Ainsi, on trouve toujours TPA-1 activée par DAG mais aussi PKC-3, EGL-8 et PLC-3 (phospholipases C), GPA-12 et RACK-1 (sous-unités de protéines G) (Ziegler et al., 2009). Le signal induisant cette voie a été identifié mais n'a pas encore fait l'objet d'une publication.
Une autre voie de signalisation est activée par le ligand DBL-1 sécrété par les neurones. DBL-1 se lie à DAF-4, un récepteur TGF-β, qui active le facteur de transcription SMA-3 (Zugasti et Ewbank, 2009). Plus récemment, (Dierking et al., 2011) ont montré qu'une protéine similaire au STAT des autres animaux, STA-2, est impliquée dans la réponse contre *D. coniospora* en étant activée par SNF-12 qui fait partie d'une grande famille de protéines incluant des transporteurs de neurotransmetteur. Cependant aucun récepteur Jak, acteur fondamentale de la réponse antivirale chez les insectes et les vertébrés, n'a été retrouvé chez *Caenorhabditis elegans*. En revanche, un récepteur couplé à une protéine G, DCAR-1, vient d'être identifiée comme étant en amont de la voie PMK-1 au niveau de l'hypoderme (Zugasti et al., 2014). Ce récepteur, un des premiers identifiés comme étant impliqué dans la réponse immuntaire, reconnait le ligand HPLA présent chez *C. elegans* lors d'une infection par D. coniospora et lorsque des dégats sont causés à la cuticule.

IV.D.2.β. Effecteurs de l'immunité dans l'hypoderme

Comme précédemment, ces voies de signalisation induisent la production de peptides antimicrobiens qui accroissent la survie de l'hôte lors d'une infection.

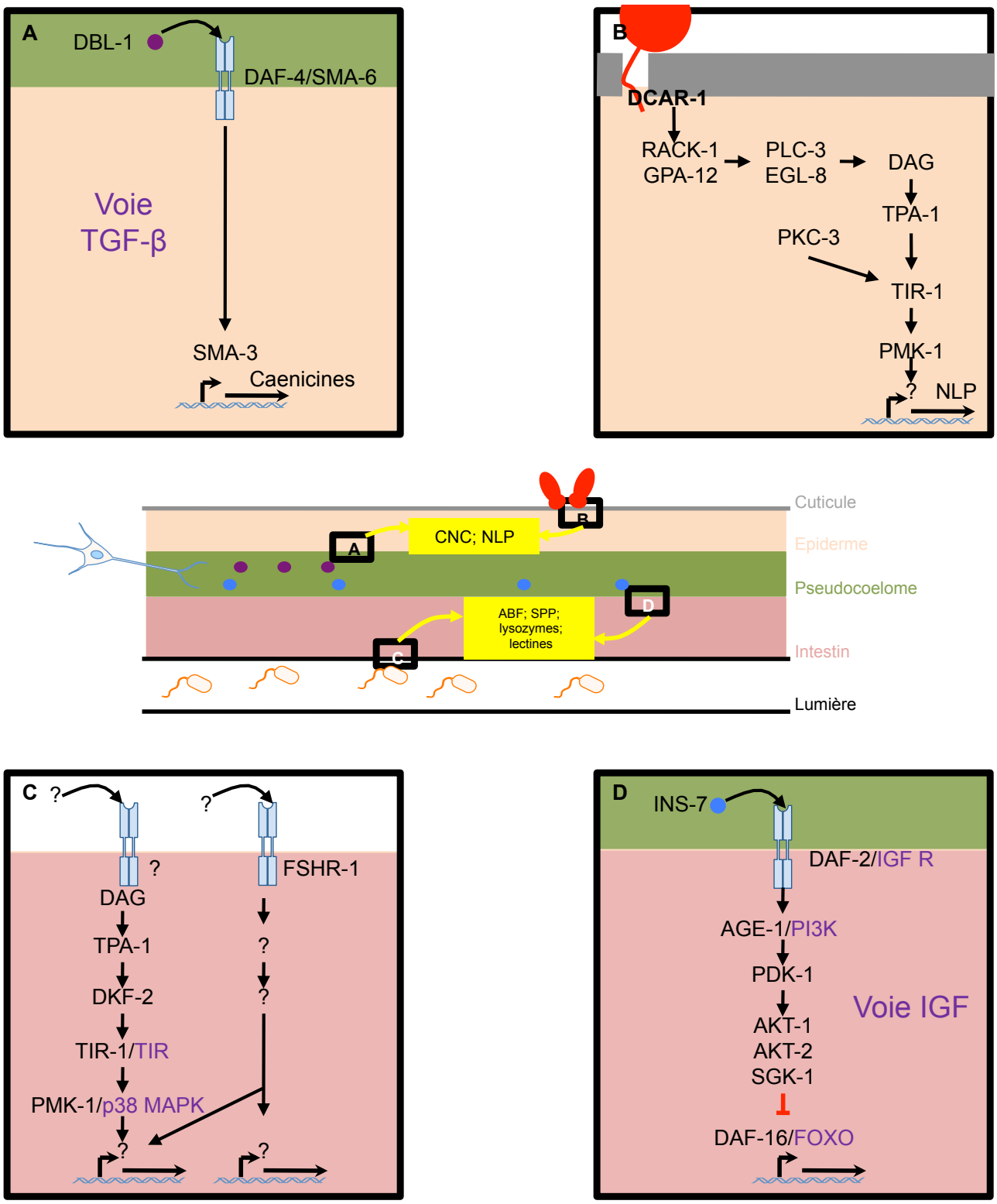


Figure 27. Défense immunitaire de différents tissus chez *C. elegans*. A: Voie de signalisation de l'hypoderme déclenchée par un signal nerveux circulant dans le pseudocoelome. **B:** Voie de signalisation de l'hypoderme durant une infection par *Drechmeria coniospora*. **C:** Voies de signalisation de l'intestin déclenchées par un ligand provenant du lumen. **D:** Voie de signalisation intestinale déclenchée par un signal nerveux circulant dans le pseudocoelome. Toutes ces voies produisent des protéines antimicrobiennes (ABF: facteur antibactérien; SPP: protéine dérivée de la saposine; lectine; lysozyme; CNC: caenecine et NLP: dérivé de neuropeptide). En mauve les orthologues retrouvés chez l'Homme. D'après Irazoqui et Ausubel (2010) et Ewbank et Zugasti (2011).

La voie MAPK induit ici la production de peptides proches des neuropeptides (NLP) appartenant au cluster de gènes *nlp-29* (Pujol et al., 2008) alors que la voie DAF-4 induit la production de caenecines (Zugasti et Ewbank, 2009). Les modes d'action de ces AMP ne sont pour le moment pas identifiés mais semblent être très divers (Pujol et al., 2012). Chez *C. elegans*, le cluster de gènes *nlp-29* regroupe 6 gènes codant pour des AMP (Pujol et al., 2008) dont 5 d'entre eux serait issus d'évènements de duplication d'un *nlp* ancestral postérieurs à la spéciation entre *C. elegans* et *C. briggsae* (Pujol et al., 2008), suggérant une forte pression de sélection exercé par les parasites susceptibles de percer la cuticule afin de coloniser le nématode.

IV.D.3. Défenses du rectum

Certains parasites sont ingérés mais n'infectent pas les cellules intestinales (Parsons et Cipollo, 2014). *Microbacterium nematophilum* est retrouvé au niveau du rectum de son hôte et induit donc une réponse localisée légèrement différente des réponses intestinales et épidermiques. C'est la MAPK MPK-1/ERK qui est alors activée par l'intermédiaire de LIN-45 et MEK-2. La réponse transcriptionnelle est sensiblement la même que celle produite par PMK-1 (Nicholas et Hodgkin, 2004).

IV.D.4. Rôle du système nerveux

IV.D.4.α. Régulation de la réponse immunitaire

L'un des avantages de *C. elegans* est la parfaite connaissance que nous avons de son anatomie liée à sa simplicité. Ainsi, il est possible ici d'envisager ici le lien entre la réponse immunitaire et le système nerveux.
Nous avons vu précédemment que la production de caenecines par l'hypoderme dépend du ligand DBL-1 produit par les neurones (Zugasti et Ewbank, 2009). De manière similaire, la réponse amorcée par DAF-2 dans les cellules intestinales dépend de la fixation d'INS-7 (insuline), également sécrétée par les neurones. Une absence de sécrétion d'INS-7 par les neurones provoque une dé-répression de DAF-16 et donc l'expression d'AMP (Kawli et Tan, 2008). De la même manière, les animaux mutants pour le gène *npr-1*, codant pour un récepteur couplé aux protéines G dans les neurones, montrent une dérégulation des gènes impliqués dans l'immunité (production d'INS-7, de DBL-1...) (Mallo et al., 2002; Kawli et Tan, 2008).

IV.D.4.β. Evitement des bactéries pathogènes

Il semblerait que le seul récepteur Toll de *C. elegans* soit impliqué dans le comportement

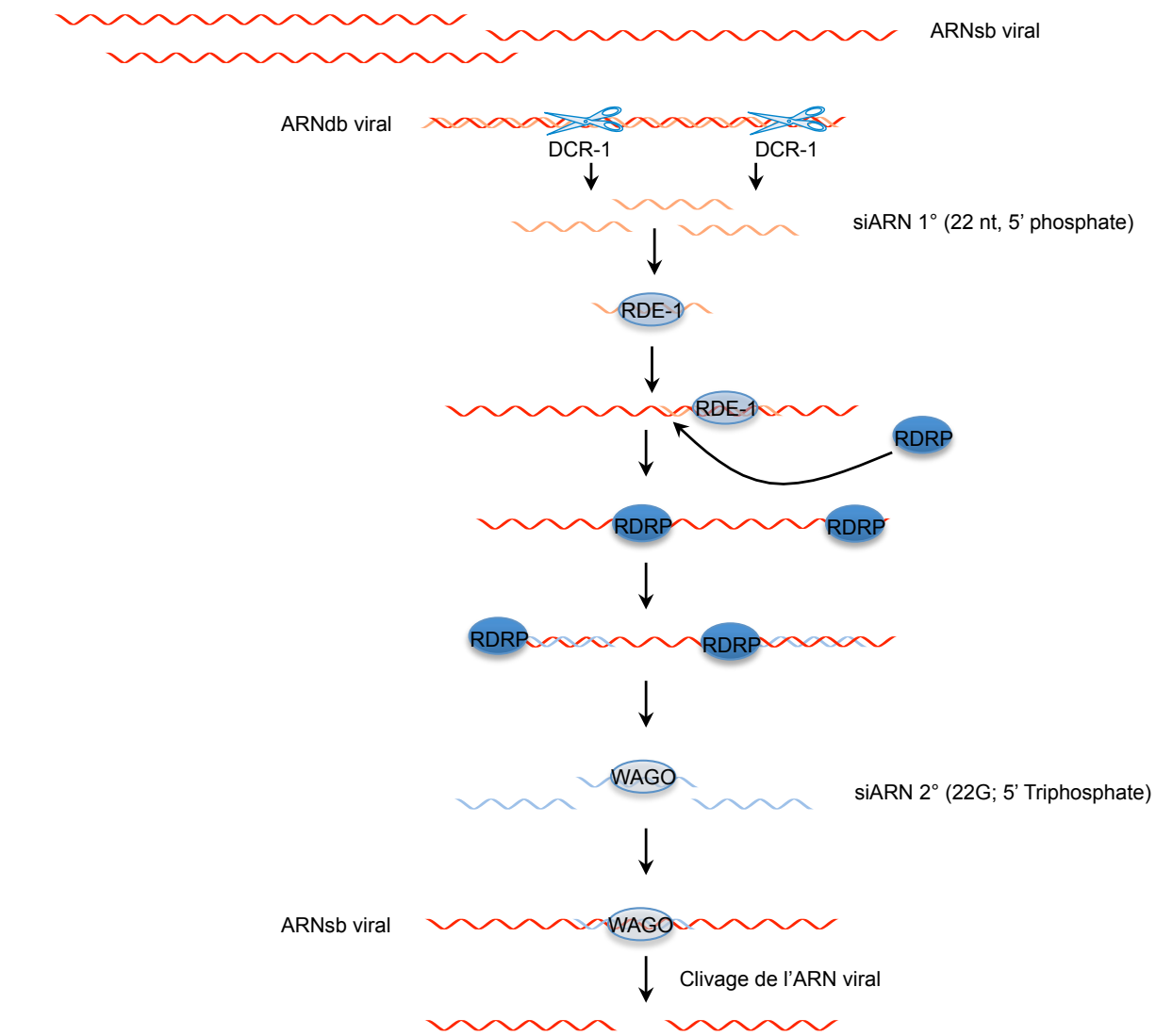


Figure 28. Réponse antivirale par les petits ARN chez *C. elegans*. La présence d'un ARN double brin (dbARN) est détectée par un complexe protéique incluant une dicer (DCR-1, RDE-1 et DRH-1). Cet ARNdb est clivé par ce même complexe en petits ARN simple brin (siARN 1°) d'une taille de 23 nucléotides et d'une extrémité 5' monophosphate. Ces si ARN 1° sont pris en charge par la protéine argonaute RDE-1. Le complexe ainsi formé se fixe à un ARN viral complémentaire du siARN 1° et permet le recrutement d'un complexe composé d'une RDRP (RRF-1 ou EGO-1).Cet RDRP permet la production d'ARN secondaires d'une longueur de 22 nucléotides, commençant par une guanine et dont l'extrémité 5' est triphosphate (siARN 2°). Ce phénomène permet une amplification de la réponse des petits ARN. Les siARN 2° sont ensuite pris en charge par la protéine Argonaute WAGO qui permet un clivage spécifique de l'ARN viral simple brin.

d'évitement de ces bactéries (Pujol et al., 2001; Pradel et al., 2007).

NPR-1 aurait également un rôle majeur dans l'évitement des bactéries pathogènes. Les animaux mutants pour *npr-1* tendent à rester sur la plage bactérienne, tandis que les animaux témoins se placent autour de la plage (Reddy et al., 2009; Styer et al., 2008).

IV.D.5. Les voies des petits ARN chez C. elegans

Aucun virus naturel n'avait été identifié chez *C. elegans* avant la découverte du virus d'Orsay par notre laboratoire. Cependant des études sur la réponse antivirale ont été effectuées en utilisant soit un Flock House Virus (FHV) artificiel, exprimé après un choc thermique (Lu et al., 2005; Lu et al., 2009), soit une infection de cellules primaires par le Vesicular Stomatitis Virus (VSV), un virus de mammifère (Schott et al., 2005; Wilkins et al., 2005). En absence des principales protéines impliquées dans l'ARN interférence, notamment des protéines argonautes et dicer, les différents virus sont capables de se répliquer (Lu et al., 2005; Schott et al., 2005; Wilkins et al., 2005), En revanche, une augmentation de l'expression d'autres protéines de cette voie diminue fortement la réplication des virus (Wilkins et al., 2005). L'interférence par les petits ARN semble donc jouer un rôle majeur dans la défense antivirale chez *C. elegans,* comme chez les plantes et les insectes. Quelques différences sont néanmoins notables.

IV.D.5.α. La voie des petits ARN exogènes

La voie des petits ARN exogènes est très utilisée pour éteindre un gène d'intérêt chez *C. elegans* (Fire et al., 1998a). Lorsqu'un ARN double brin se retrouve dans une cellule de *C. elegans,* il est reconnu par des protéines RDE-4 (Parker et al., 2006) qui recrutent un complexe composé de la protéine dicer DCR-1, de DRH-1 et de la protéine argonaute RDE-1 (Tabara et al., 2002). Malgré le fait que ce complexe soit essentiel à la mise en place de l'ARN interférence, un animal mutant pour le gène *drh-1* est toujours capable de produire des petits ARN à partir d'ARN double brin ciblant un gène de l'animal. Ce complexe clive l'ARN double brin en petits ARN double brin reconnus par RDE-1 (Yigit et al., 2006). Ces petits ARN primaires mesurent 23 nucléotides et ne possèdent qu'un phosphate à leur extrémité 5'. Le duplexe constitué par les deux brins d'ARN est ensuite séparé en deux simple brin dont un sert de guide à la fixation du complexe ARN-RDE-1 à l'ARN messager correspondant. Ce complexe recrute alors les polymérases à ARN dépendantes de l'ARN RRF-1 et EGO-1 (Sijen et al., 2001; Smardon et al., 2000) ainsi que leurs cofacteurs DRH-3 et EKL-1. RRF-1 produit alors des ARN simple brins, nommés siARN secondaire, d'une longueur de 22 nucléotides, débutant par une guanine et dont l'extrémité 5' est triphosphatée à partir de l'ARN messager (Pak et Fire, 2007). Ces petits ARN secondaires sont recrutés par la protéine argonaute WAGO (Worm

ArGOnaute) qui permet le clivage des ARN messager correspondant (Yigit et al., 2006). Lors d'une infection virale ce ne sont pas les ARN messager du nématode qui sont la cible du clivage mais les ARN messager viraux. Les ARN secondaires peut également être pris en charges par l'argonaute NRDE-3 (Nuclear RDE-3) et être menés dans le noyau (Guang et al., 2008). La reconnaissance spécifique d'un ARNm complémentaire en cours de transcription provoque le recrutement des protéines NRDE-1/2/4 (Burton et al., 2011). Ce complexe favorise lamise en place d'une marque H3K9me (méthylation de la lysine 9 de l'histone 3), inhibitrice de la transcription par une histone méthyltransférase (Burkhart et al., 2011).

IV.D.5.β. L'implication de drh-1 *dans la réponse antivirale*

Un crible par ARN interférence, reposant sur la capacité de réplication du FHV dans les cellules intestinales de *C. elegans,* a permis d'identifier les gènes *drh-1* et *drh-2* (dicer related helicase) comme spécifiquement impliqués dans la réponse antivirale (Lu et al., 2009). La protéine DRH-1 est un orthologue des protéines RLR (RIG-I-Like Receptor) présents chez les vertébrés. Ces protéines sont impliquées dans la reconnaissance des ARN viraux et amorcent la réponse par les interférons. Cependant, dans cette étude, DRH-1 semble agir en aval de la détection de l'ARN viral et *C. elegans* ne produit pas d'interférons, suggérant un rôle différent de DRH-1 dans la réponse antivirale. Lu et al. suggèrent également que *drh-2*, probablement un paralogue de *drh-1*, inhibe la réponse antivirale. L'annotation disponible classe *drh-2* en tant que pseudogène (www.wormbase.org release WS243) tandis que d'autres études ont pu identifier un peptide correspondant à sa séquence (Duchaine et al., 2006). La fonction de *drh-2* dans la réponse antivirale chez *C. elegans* n'est donc pas certaine.

IV.E.5.γ. La voie des microARN

Avant la découverte des petits ARN exogènes, d'autres ARN non-codants avaient été identifiés à cause de leur rôle important dans la régulation de l'expression des gènes chez *C. elegans* : les miARN (ou microARN) (Lee et al., 1993; Reinhart et al., 2000). La production d'un miARN débute par la transcription d'un pri-miARN (miARN primaire) qui est par la suite coiffé et polyadénylé. Il est ensuite pris en charge par laRNase III DRSH-1 et sa partenaire PASH-1 pour former le pré-miARN (Denli et al., 2004). A l'issue de cette étape, ce dernier forme une structure en épingle à cheveux d'une longueur d'environ 70 nt (Lee et al., 2002) et est exporté vers le cytoplasme. Il est alors reconnu et clivé par DCR-1 en un duplex d'ARN une longueur d'environ 22 pb (Ketting et al., 2001). L'un des deux brins d'ARN est reconnu par une protéine Argonaute (généralement ALG-1, pour Argonaute-Like-Gene et ALG-2) (Grishok et al., 2001). Le complexe miARN-Argonaute ainsi formé est recruté par le complexe miRISC

(microRNA induced RISC) (Zhang et al., 2007). Le miARN reconnaît de manière imparfaite la partie 3'UTR de l'ARNm cible et recrute des protéines réprimant la traduction ou induisant la déadénylation de ce dernier. Les miARN sont impliqués dans la régulation de l'expression de gènes lors du développement embryonnaire et post embryonnaire (comme par exemple les premiers miARN découverts lin-4 et let-7 impliqués dans la régulation des transitions développementales (Chalfie et al., 1981; Reinhart et al., 2000)) mais aussi dans des processus physiologiques tels que la communication entre muscles et neurones (Simon et al., 2008) ou encore dans le processus de vieillissement impliquant la voie DAF-16/FOXO (Shen et al., 2012). Enfin, de récentes études ont mis en évidence un rôle des miARN dans la réponse immunitaire de *C. elegans*. Il semblerait en effet que DCR-1 joue un rôle important dans la réponse contre l'infection par la souche DB27 de Bacillus thuringiensis (Iatsenko et al., 2013). Il apparait également que de nombreux miARN ciblant à des gènes impliqués dans la réponse immunitaire contre Pseudomonas aeruginosa sont associés aux complexes RISC présents cellules intestinales (Kudlow et al., 2012), suggérant un rôle répresseur des miARN sur la réponse basale de ces gènes.

La plupart des protéines impliquées dans les deux voies (siARN exogènes et miARN) sont différentes, exception faite de DCR-1. Sawh et Duchaine, en 2013, ont identifié un produit de clivage stable de la protéine DCR-1 qui à la fois augmente la production de petits ARN exogènes mais qui réduit également la production des miARN en séquestrant les Argonautes ALG-1 et ALG-2, révélant ainsi une forte compétition entre les deux voies.

IV.E.5.δ. La voie endogène des petits ARN

La voie endogène des petits ARN est produite à partir d'un ARNm endogène à l'animal. Le mécanisme responsable de la reconnaissance spécifique de l'ARNm engageant la voie est encore inconnue. Un complexe contenant la RDRP RRF-3 (Lee et al., 2006), DRH-3 et ERI-5 produit le brin complémentaire à partir de l'ARNm. Ce duplex est ensuite reconnu et clivé par la protéine DCR-1. Il se compose d'un brin guide de 26 nt, débutant par un G (Ambros et al., 2003; Vasale et al., 2010), et d'un brin passager de 19 nt, et est pris en charge par des protéines Argonautes : ALG-3/4 dans les spermatozoïdes et ERGO-1 (Yigit et al., 2006) dans les ovocytes et dans les embryons. Les complexes argonautes-ARN-26G ainsi formés reconnaissent l'ARNm complémentaire et recrutent une autre RDRP (RRF-1) , associée à DRH-3 et EKL-1, qui produit des ARN secondaires de 22 nt débutant par un G (Duchaine et al., 2006; Gu et al., 2009). Ces derniers sont recrutés par une protéine WAGO (Vasale et al., 2010) et le complexe ainsi formé reconnaît un ARNm complémentaire et inhibe soit sa traduction, soit sa transcription.

Les animaux mutants pour un des éléments de la voie des petits ARN endogènes montrent une mobilisation plus importante de transposons, une augmentation d'ARNm aberrants et une augmentation de l'expression de gènes dupliqués (Gu et al., 2009) dans les gamètes et dans l'embryon, suggérant un rôle fondamental dans le système de surveillance du génome.

IV.E.5.ε. La voie piARN

Les piARN, pour petits ARN associés aux protéines PIWI, se caractérisent par une longueur de 21 nt et un biais de U en 5' (d'où leur autre nom ARN 21U) et sont exprimés exclusivement dans les cellules de la lignée germinale chez *C. elegans*. Les précurseurs de ces petits ARN sont transcrits par la polymérase à ARN II à partir de séquences précédées par un motif spécifique, organisées en deux clusters sur le Chromosome IV (Ruby et al., 2006). Ils présentent une coiffe 7-methylguanylate en 5' et leur taille est généralement de 28 ou 29 nt (Goh et al., 2014; Weick et al., 2014). Les enzymes impliquées dans la maturation du précurseur vers un ARN 21U mature ne sont pas clairement identifiées mais les protéines TOFU-1 et TOFU-2 (Twenty One u-rna biogenesis Fouled Up) pourraient y participer (Goh et al., 2014). Une fois mature, le piARN est pris en charge par l'unique protéine PIWI découverte jusqu'à ce jour chez *C. elegans* : PRG-1 (Batista et al., 2008; Wang et Reinke, 2008)(Batista et al., 2008; Wang et Reinke, 2008). Ce complexe piARN-PRG-1 se fixe à un ARNm correspondant et permet une amplification des piARN par l'intermédiaire d'une RDRP (Das et al., 2008; Lee et al., 2012). Cette amplification produit des piARN secondaires d'une longueur de 22 nt, possédant une extrémité 5' triphosphate et débutant par un G. Ces ARN-22G sont recrutés par l'argonaute secondaire HRDE-1 (pour "Heritable RNAi DEficient-1") et le complexe ainsi formé est transporté dans le noyau (Ashe et al., 2012). Les protéines NRDE-1/2/4 sont alors recrutées lors de la fixation du complexe sur un ARNm correspondant en cours de production par la polymérase à ARN II de manière similaire à ce qui a été observé pour la voie exogène des petits ARN (Burton et al., 2011a). Ce recrutement permet le dépôt de marques H3K9me3, inhibitrices de la transcription, par des méthyltransférases (voir Bélicard et Félix, 2012 en annexe).

Ces piARN semblent jouer un rôle important dans le maintien de l'intégrité de la lignée germinale. Les animaux mutant pour le gène prg-1 montrent des problèmes de fertilité liés à la diminution du nombre de noyaux mitotiques et méiotiques dans les gonades (Cox et al., 1998). Les mutants présentent apparemment également une diminution de l'expression de gènes impliqués dans la spermatogénèse (Wang et Reinke, 2008)(Wang et al., 2008). Le gène prg-1 est également impliqué dans la répression de la mobilisation des transposons Tc3 dans la lignée germinale (Batista et al., 2008; Das et al., 2008)(Batista et al., 2008; Das et al., 2008), jouant ainsi un rôle clé, au côté de la voie

endogène des petits ARN, dans le maintien de l'intégrité du génome lors de la gamétogénèse.

Toutes ces voies ne sont pas totalement indépendantes les unes des autres, et possèdent des acteurs communs. Par exemple, la protéine DCR-1 intervient dans la voie exogène, la voie endogène et la voie des miARN (Grishok et al., 2001; Hutvágner et al., 2001; Knight et Bass, 2001), suggérant une potentielle compétition entre ces voies (Duchaine et al., 2006). De même, DRH-3 intervient dans la production de petits ARN secondaires dans les voies endo et exogènes et dans la production des ARN primaires dans celle des piARN (Aoki et al., 2007; Gu et al., 2009). Mais les interactions ne s'arrêtent pas à l'implication de des mêmes protéines dans ces différentes voies. Par exemple, les animaux mutants pour protéines RRF-3 et ERI-1 (Enhanced RNA Interference), qui interviennent spécifiquement dans la voie endogène des petits ARN, montrent une augmentation de la production des petits ARN provenant de la voie exogène, tout en étant dispensables à cette voie (Lee et al., 2006). Ceci suggère que non seulement la quantité de protéines commune à chacune des voies est limitantes mais également que les voies interagissent entre elles directement ou indirectement.

De nombreux effecteurs ont été identifiés chez *C. elegans*, révélant parfois une spécificité d'action. Cependant, la plupart des récepteurs à l'origine de la production de ces effecteurs sont inconnus et la manière dont sont reconnues les infections reste incomprise.

IV. Conclusion

C. elegans est un très bon modèle pour l'étude de la biologie cellulaire et moléculaire. Néanmoins, la compréhension du fonctionnement de cet organisme dans son environnement naturel reste insuffisante. Les récentes campagnes d'échantillonnage, que ce soit à l'échelle mondiale ou locale ont permit d'avoir vision plus précise de son écologie. *C. elegans* est un animal vivant dans des substrats riches en bactéries, comme les fruits et les plantes en décomposition.

Une meilleure compréhension de l'habitat de *C. elegans* nous apporte de nombreux éléments essentiels à la compréhension de son évolution dans la nature. L'aptitude qu'a *C. elegans* à survivre à des conditions environnementales très variées par l'intermédiaire des L1 et des dauers est depuis longtemps admise. Cette grande résistance, associée au comportement migratoire des dauers permettrait d'expliquer l'absence de structure génétique des populations tant à l'échelle globale.

C. elegans est également depuis une vingtaine d'année devenu un modèle pour l'étude du système immunitaire innée car il montre une sensibilité à certains parasites pathogènes d'intérêt médical ou agronomique. Ces études ont mis en évidence l'absence d'éléments pourtant très conservés chez les métazoaires tels que les récepteurs Toll, LRR ou encore le facteur de transcription NF-κB. Même si la description des voies de signalisation devient de plus en plus précise, la manière dont elles sont activées reste dans la majorité des cas inconnue. La découverte de ces récepteurs pourrait nous en apprendre plus sur la spécificité du système immunitaire inné chez *C. elegans* et par conséquent nous en apprendre plus sur les interactions qu'il entretient avec ses parasites.

L'association de données concernant sa génétique des populations, son écologie et son immunité nous permet aujourd'hui d'envisager *C. elegans* comme un modèle d'intérêt en génétique écologique et plus particulièrement dans l'étude des interactions hôte-parasite.

Lors de mon arrivée au laboratoire, l'objectif était de récolter des informations sur les différents parasites naturels de *C. elegans*. Ma grande chance fut la découverte des premiers virus infectant naturellement des *Caenorhabditis*.

Cette découverte a permis d'orienter ma thèse vers la relation *Caenorhabditis*-virus.

Dans un premier temps j'ai participé à la **description des infections naturelles et expérimentales de *Caenorhabditis* par des virus proches des Nodavirus**. J'ai plus particulièrement participé à la description des symptômes causés par l'infection mais aussi à l'étude de la spécificité d'infection (infections croisées entre espèces de *Caenorhabditis* et infections intra-spéciques). Ces travaux sont le fruit d'une collaboration avec les laboratoires de David Wang et d'Eric Miska.

Mon second objectif a été d'**évaluer la variabilité de sensibilité au virus d'Orsay chez *C. elegans*** et de **déterminer les bases génétiques de cette variabilité.** Ces travaux ont été réalisés en étroite collaboration avec le laboratoire d'Eric Miska.

RESULTATS

I. Infection naturelle et expérimentale de nématodes *Caenorhabditis* par de nouveaux virus proches des Nodavirus

Notre étude rapporte le premier cas de *Caenorhabditis* naturellement infectés par des virus. La souche de *C. elegans* JU1580, en provenance d'Orsay, et la souche de *Caenorhabditis briggsae* JU1264, en provenance de Santeuil, sont infectées par le virus d'Orsay et le virus de Santeuil, respectivement. Ces virus, transmis horizontalement, provoquent chez leurs hôtes une désorganisation des cellules intestinales mais n'affectent pas drastiquement leur survie ou leur fécondité.

Ces virus possèdent une organisation proche des Nodavirus : deux ARN simple brin positifs, le premier (ARN1) codant pour Polymérase à ARN dépendante de l'ARN et le second (ARN2) codant pour une protéine de capside et une protéine sans homologue connu. Ils sont retrouvés principalement dans les cellules intestinales et la gonade somatique de leurs hôtes.

Les infections de *C. elegans* par le virus de Santeuil et celle de *C. briggsae* par le virus d'Orsay montrent une spécificité liée d'espèce. Les tentatives infructueuses d'infection des souches standards de *C. elegans* et *C. briggsae* de laboratoire soulignent également une variation intraspécifique de sensibilité. Cette variation se retrouve lorsque l'on teste la sensibilité de différentes souches naturelles de *C. elegans* au virus d'Orsay.

La variation de sensibilité au virus d'Orsay chez *C. elegans* pourrait être associée à la capacité de l'animal à produire une réponse par les petits ARN. L'isolat JU1580 ainsi que plusieurs autres souches déficientes pour l'ARN interférence (souches naturelles ou mutantes pour des éléments de l'ARN interférence) montrent une sensibilité accrue au virus d'Orsay. Ces résultats confirment le rôle majeur de cette voie dans la réponse antivirale observé par les études précédemment effectuées sur les virus artificiels ou non-naturels.

En bleu, ma participation à l'étude.
Figures réalisées : Figure 1E-1H' ; Figure 1K ; Figure 4A-B ; Figure 5A-B ; Figure 7B ; Figure 8E ; Figure S3.

OPEN ACCESS Freely available online

Natural and Experimental Infection of *Caenorhabditis* Nematodes by Novel Viruses Related to Nodaviruses

Marie-Anne Félix[1*☯], Alyson Ashe[2☯], Joséphine Piffaretti[1☯], Guang Wu[3], Isabelle Nuez[1], Tony Bélicard[1], Yanfang Jiang[3], Guoyan Zhao[3], Carl J. Franz[3], Leonard D. Goldstein[2], Mabel Sanroman[1], Eric A. Miska[2*], David Wang[3*]

1 Institut Jacques Monod, CNRS-University of Paris-Diderot, Paris, France, **2** Gurdon Institute, University of Cambridge, Cambridge, United Kingdom, **3** Departments of Molecular Microbiology and Pathology & Immunology, Washington University in St. Louis School of Medicine, St. Louis, Missouri, United States of America

Abstract

An ideal model system to study antiviral immunity and host-pathogen co-evolution would combine a genetically tractable small animal with a virus capable of naturally infecting the host organism. The use of *C. elegans* as a model to define host-viral interactions has been limited by the lack of viruses known to infect nematodes. From wild isolates of *C. elegans* and *C. briggsae* with unusual morphological phenotypes in intestinal cells, we identified two novel RNA viruses distantly related to known nodaviruses, one infecting specifically *C. elegans* (Orsay virus), the other *C. briggsae* (Santeuil virus). Bleaching of embryos cured infected cultures demonstrating that the viruses are neither stably integrated in the host genome nor transmitted vertically. 0.2 μm filtrates of the infected cultures could infect cured animals. Infected animals continuously maintained viral infection for 6 mo (~50 generations), demonstrating that natural cycles of horizontal virus transmission were faithfully recapitulated in laboratory culture. In addition to infecting the natural *C. elegans* isolate, Orsay virus readily infected laboratory *C. elegans* mutants defective in RNAi and yielded higher levels of viral RNA and infection symptoms as compared to infection of the corresponding wild-type N2 strain. These results demonstrated a clear role for RNAi in the defense against this virus. Furthermore, different wild *C. elegans* isolates displayed differential susceptibility to infection by Orsay virus, thereby affording genetic approaches to defining antiviral loci. This discovery establishes a bona fide viral infection system to explore the natural ecology of nematodes, host-pathogen co-evolution, the evolution of small RNA responses, and innate antiviral mechanisms.

Citation: Félix M-A, Ashe A, Piffaretti J, Wu G, Nuez I, et al. (2011) Natural and Experimental Infection of *Caenorhabditis* Nematodes by Novel Viruses Related to Nodaviruses. PLoS Biol 9(1): e1000586. doi:10.1371/journal.pbio.1000586

Academic Editor: Jonathan Hodgkin, Oxford University, United Kingdom

Received August 5, 2010; **Accepted** December 10, 2010; **Published** January 25, 2011

Copyright: © 2011 Félix et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported in part by the CNRS (Félix lab), National Institutes of Health grant U54 AI057160 to the Midwest Regional Center of Excellence for Biodefense and Emerging Infectious Disease Research (Wang lab), and a Cancer Research UK Programme Grant (Miska lab). DW holds an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund. AA was supported by a fellowship from the Herchel-Smith Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Eric Miska's spouse is a member of the PloS Biology editorial staff; in accordance with the PLoS policy on competing interests she has been excluded from all stages of the review process for this article.

* E-mail: felix@ijm.univ-paris-diderot (MAF); eam29@cam.ac.uk (EAM); davewang@wustl.edu (DW)

☯ These authors contributed equally to this work.

Introduction

Model organisms such as *D. melanogaster* [1,2] and *C. elegans* [3,4] have been increasingly used in recent years to examine features of the host immune system and host-pathogen co-evolution mechanisms, due to the genetic tractability and ease of manipulation of these organisms. A prerequisite to fully exploit such models is the identification of an appropriate microbe capable of naturally infecting the host organism. Analysis in *C. elegans* of bacterial pathogens such as *Pseudomonas*, *Salmonella*, or *Serratia* has been highly fruitful, in some instances revealing the existence of innate immune pathways in *C. elegans* that are also conserved in vertebrates [3]. The recent report of natural infections of *C. elegans* intestinal cells by microsporidia makes it a promising model for microsporidia biology [5]. Efforts to use *C. elegans* to understand anti-viral innate immunity, however, have been hampered by the lack of a natural virus competent to infect and replicate in *C. elegans*.

In the absence of a natural virus infection system, some efforts to define virus-host responses in *C. elegans* have been pursued using artificial methods of introducing viruses or partial virus genomes into animals [6,7]. For example, the use of a transgenic Flock House virus RNA1 genome segment has clearly established a role for RNAi in counteracting replication of Flock House virus RNA [7] and has defined genes essential for the RNAi response [8]. However, this experimental system can only examine replication of the viral RNA and is fundamentally unable to address the host response to other critical aspects of the virus life cycle such as virus entry, virion assembly, or egress. The ability of a host to target steps other than genome replication to control viral infections is highlighted by recent discoveries such as the identification of tetherin, which plays a critical role at the stage of viral egress by blocking the release of fully assembled HIV virions from infected human cells [9]. Furthermore, the artificial systems used to date for analysis of virus-nematode interactions cannot be used to examine transmission dynamics of virus infection. These limitations underscore the need to establish an authentic viral infection and replication system in nematodes.

Natural populations of *C. elegans* have proven hard to find until recent years. The identification of *C. elegans* habitats and the

Author Summary

The nematode *C. elegans* is a robust model organism that is broadly used in biology. It also has great potential for the study of host-microbe interactions, as it is possible to systematically knockout almost every gene in high-throughput fashion to examine the potential role of each gene in infection. While *C. elegans* has been successfully applied to the study of bacterial infections, only limited studies of antiviral responses have been possible since no virus capable of infecting any *Caenorhabditis* nematode in laboratory culture has previously been described. Here we report the discovery of natural viruses infecting wild isolates of *C. elegans* and its relative *C. briggsae*. These novel viruses are most closely related to the ssRNA nodaviruses, but have larger genomes than other described nodaviruses and clearly represent a new taxon of virus. We were able to use these viruses to infect a variety of laboratory nematode strains. We show that mutant worms defective in the RNA interference pathway, an antiviral system known to operate in a number of organisms, accumulate more viral RNA than wild type strains. The discovery of these viruses will enable further studies of host-virus interactions in *C. elegans* and the identification of other host mechanisms that counter viral infection.

development of simple isolation methods (MAF, unpublished) [10] has now enabled extensive collection of natural isolates of *C. elegans*. Here we report the discovery of natural populations of *C. elegans* and of its close relative *C. briggsae* that display abnormal morphologies of intestinal cells. These abnormal phenotypes can be maintained in permanent culture for several months, without detectable microsporidial or bacterial infection. We show that these populations are infected by two distinct viruses, one specific for *C. elegans* (Orsay virus), one for *C. briggsae* (Santeuil virus). These viruses resemble viruses in the *Nodaviridae* family, with a small, bipartite, RNA(+sense) genome. Infection by each virus is transmitted horizontally. In both nematode species, we find intraspecific variation in sensitivity to the species-specific virus. We further show that infected worms mount a small RNA response and that RNAi mechanisms act as antiviral immunity in nematodes. Finally, we demonstrate that the *C. elegans* isolate from which Orsay virus was isolated is incapable of mounting an effective RNAi response in somatic cells. We thus find natural variation in host antiviral defenses. Critically, these results establish the first experimental viral infection system in *C. elegans* suitable for probing all facets of the host antiviral response.

Results

Natural Viral Infections of *C. briggsae* and *C. elegans*

From surveys of wild nematodes from rotting fruit in different regions of France, multiple *Caenorhabditis* strains were isolated that displayed a similar unusual morphology of the intestinal cells and no visible pathogen by optical microscopy. Intestinal cell structures such as storage granules disappeared (Figures 1A–J, 2A–C) and the cytoplasm lost viscosity and became fluid (Figure 1B,I), moving extensively during movement of the animal. The intestinal apical border showed extensive convolutions and intermediate filament disorganization (Figures 1A, 2H, as described in some intermediate filament mutants, [11]). Multi-membrane structures were sometimes apparent in the cytoplasm (Figure 1C). Elongation of nuclei and nucleoli, and nuclear degeneration, were observed using Nomarski optics, live Hoechst 33342 staining, and electron microscopy (Figures 1E–H, 2D–F). Finally, some intestinal cells fused together (Figure 1I). This suite of symptoms was first noticed during sampling of *C. briggsae*. Indeed, more individuals appeared affected in *C. briggsae* than in *C. elegans* cultures, and to a greater extent (Figure 1K).

One representative, stably infected, strain of each nematode species, *C. elegans* JU1580 (isolated from a rotting apple in Orsay, France) and *C. briggsae* JU1264 (isolated from a snail on a rotting grape in Santeuil, France), were selected for detailed analysis. Bleaching of adult animals resulted in phenotype-free progeny from both strains, demonstrating that the phenotype was not vertically transmitted (embryos are resistant to the bleaching treatment) (Figure 1K). Addition of dead infected animals, or homogenates from infected animals after filtration through 0.2 μm filters, to plates containing previously bleached animals recapitulated the morphological phenotype, raising the possibility that a virus might play a role in inducing the morphological phenotype (Figure 1K). We found that the infectious agent could be passed on horizontally through live animals by incubating GFP-labeled animals (strain JU1894, Table 1) with 10 non-GFP-infected worms (JU1580), checking that the latter did not die before removing them 24 h later. The GFP-labeled culture displayed the intestinal symptoms after a week. One possibility is that the intestinal infectious agent is shed from the intestine through the rectum and may enter the next animal during feeding.

In support of the hypothesis that these wild *Caenorhabditis* were infected by a virus, small virus-like particles of approximately 20 nm diameter were visible by electron microscopy of the intestinal cells (Figures 2H, S1). Such particles were not observed in bleached animals, nor in *C. elegans* animals infected by bacteria, which showed a strong reduction of intestinal cell volume (strain JU1409, unpublished data).

While a clear morphological phenotype was visible by microscopy, infection did not cause a dramatic decrease in adult longevity (unpublished data), nor a change in brood size (Figure S2A,B). However, progeny production was significantly slowed down during adulthood, most clearly in the infected *C. briggsae* JU1264 isolate compared to the uninfected control (Figure S2D).

Molecular Identification of Two Divergent Viruses

An unbiased high-throughput pyrosequencing approach was used to determine whether any known or novel viruses were present in the animals. From JU1264, 28 unique sequence reads were identified initially that shared 30%–48% amino acid sequence identity to known viruses in the family *Nodaviridae*. Nodaviruses are bipartite positive strand RNA viruses. The RNA1 segment of all previously described nodaviruses is ~3.1 kb and encodes ORF A, the viral RNA-dependent RNA polymerase. Some nodaviruses also encode ORFs B1/B2 at the 3′ end of the RNA1 segment. The B1 protein is of unknown function while the B2 protein is able to inhibit RNAi [12]. The RNA2 segment of all previously described nodaviruses is ~1.4 kb and possesses a single ORF encoding the viral capsid protein. Assembly of the initial JU1264 pyrosequencing reads followed by additional pyrosequencing, RT-PCR, 5′ RACE, and 3′ RACE yielded two final contigs, which were confirmed by sequencing of overlapping RT-PCR amplicons. The two contigs corresponded to the RNA1 and RNA2 segments of a novel virus. The first contig (3,628 nt) encoded a predicted open reading frame of 982 amino acids that shared 26%–27% amino acid identity to the RNA-dependent RNA polymerase of multiple known nodaviruses by BLAST alignment. All known nodavirus B2 proteins overlap with the C-terminus of the RNA-dependent RNA polymerase and are

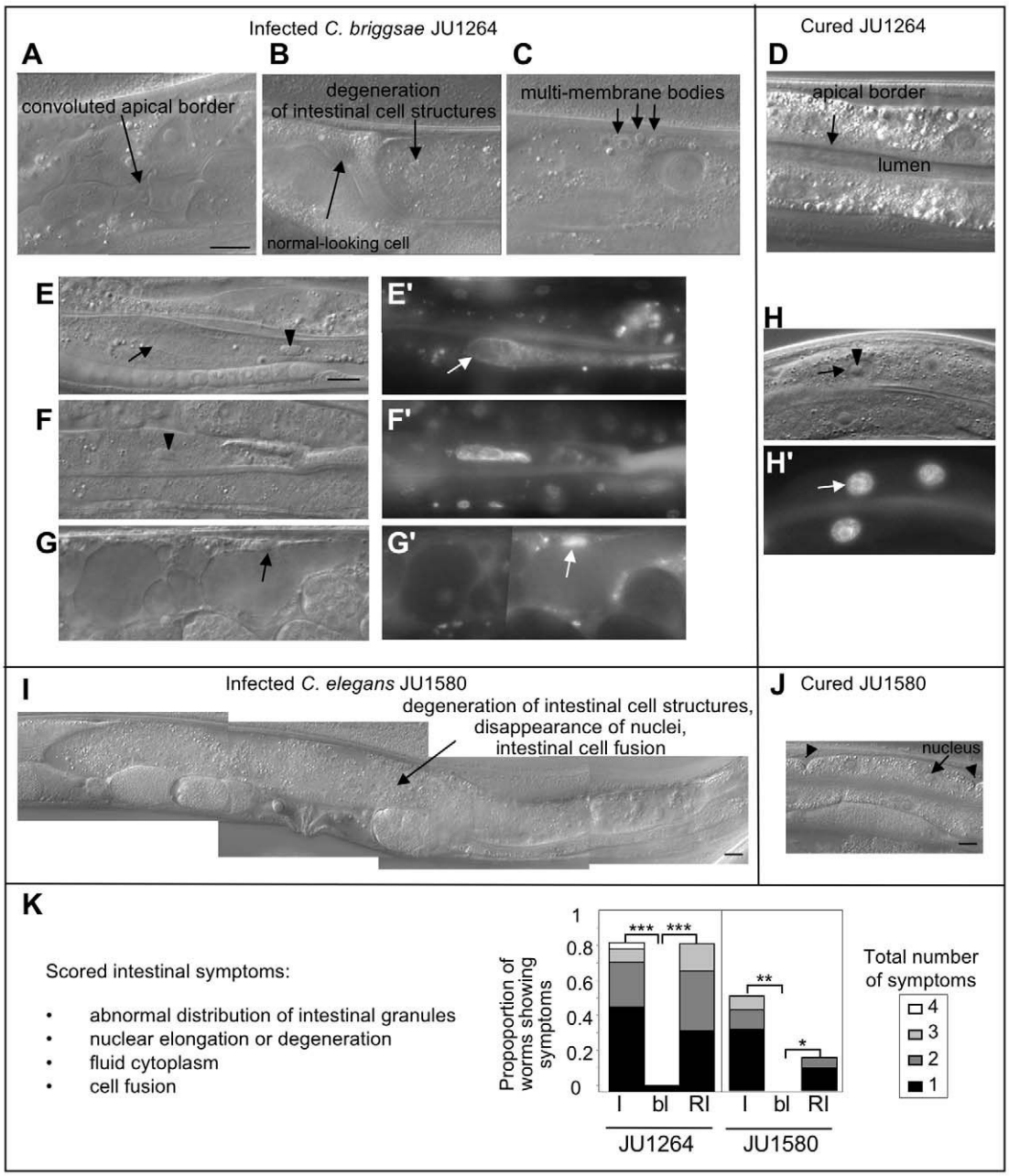


Figure 1. Intestinal cell infection phenotypes in wild *Caenorhabditis* isolates. (A–H) *C. briggsae* JU1264 and (I,J) *C. elegans* JU1580 observed by Nomarski microscopy. (A–C, E–G, I) Infected adult hermaphrodites from the original cultures, with the diverse infection symptoms: convoluted apical intestinal border (A), degeneration of intestinal cell structures and liquefaction of the cytoplasm (B, G, I), presence of multi-membrane bodies (C). The animals in (E–H) were also observed in the fluorescence microscope after live Hoechst 33342 staining of the nuclei, showing the elongation and degeneration of nuclei (E′–H′). In (E), the nucleus and nucleolus are abnormally elongated. In (F), the nuclear membrane is no longer visible by Nomarski optics. In (G), the cell cytoplasmic structures are highly abnormal (apparent vacuolisation) and the nucleus is very reduced in size. In (E–H′), arrows denote nuclei and arrowheads nucleoli. The infected animal in (I) displays an abnormally large intestinal cell that is probably the result of cell fusions, with degeneration of cellular structures including nuclei. (D, H, J) Uninfected (bleached) adults. Arrowheads in (J) indicate antero-posterior boundaries between intestinal cells, each of which generally contains two nuclei. Bars: 10 μm. (K) Proportion of worms showing the indicated cumulative number of morphological infection symptoms in at least one intestinal cell, in the original wild isolate (I), after bleaching (bl) and after re-infection by a 0.2 μM filtrate (RI). Note that not all symptoms shown in (A–I) were scored, because some are difficult to score or may also occur in healthy animals. The animals were scored 4 d after re-infection for *C. briggsae* JU1264, and 7 d after re-infection for *C. elegans* JU1580, at 23°C. The symptoms are similar in both species, and generally more frequent in JU1264. *** p value on number of worms showing infection symptoms $<7.10^{-11}$, Fisher's exact test; ** p value$<3.10^{-6}$; * p value$<3.10^{-2}$.
doi:10.1371/journal.pbio.1000586.g001

encoded in the +1 frame relative to the polymerase. No open reading frame with these properties was predicted in the 3′ end of the RNA1 segment. The second contig of 2,653 nt, which was presumed to be the near-complete RNA2 segment, encoded at its 5′ end a predicted protein with ~30% identity to known nodavirus capsid proteins (Figure 3A). This contig was ~1 kb

Table 1. Strain list.

Strain	Genotype
JU1264	*C. briggsae* wild isolate, Santeuil, France
JU1580	*C. elegans* wild isolate, Orsay, France
AF16	*C. briggsae* wild reference isolate, India
N2	*C. elegans* wild reference isolate, England
CB4856	*C. elegans* wild isolate, Hawaii
AB1	*C. elegans* wild isolate, Australia
PB303	*C. elegans* wild isolate, USA
PB306	*C. elegans* wild isolate, USA
JU258	*C. elegans* wild isolate, Madeira
PS2025	*C. elegans* wild isolate, California, USA
JU1894	*mfEx50*[*let858*::GFP, *myo-2*::DsRed] in JU1580 background
JU1895	*mfEx51*[*let858*::GFP, *myo-2*::DsRed] in N2 background
WM27	*rde-1(ne219)* V
WM29	*rde-2(ne221)* I
WM49	*rde-4(ne301)* III
NL936	*unc-32(e189) mut-7(pk204)* III

doi:10.1371/journal.pbio.1000586.t001

larger than the RNA2 segment of all previously described nodaviruses and appeared to encode a second ORF of 332 amino acids at the 3′ end. This second predicted ORF, named ORF δ, had no significant BLAST similarity to any sequence in Genbank.

Pyrosequencing of JU1580 demonstrated the presence of a second distinct virus that shared the same general genomic organization as the virus detected in JU1264. Partial genome sequences of 2,680 nucleotides of the RNA1 segment and 2,362 nucleotides of the RNA2 segment were obtained and confirmed by RT-PCR. The putative RNA-dependent RNA polymerases of the two viruses shared 44% amino acid identity by BLAST analysis. Like the virus in JU1264, the virus in JU1580 was predicted to encode a capsid protein at the 5′ end of the RNA2 segment as well as a second ORF in the 3′ half of the RNA2 segment. The ORF δ encoded proteins from the two viruses shared 37% amino acid identity when compared using BLAST. Thus, the genomic organization of these two viruses, while sharing substantial commonality with known nodaviruses, also displayed novel genomic features. Phylogenetic analysis of the predicted RNA polymerase and capsid proteins demonstrated that the virus sequences in JU1580 and JU1264 were highly divergent from all previously described nodaviruses and most closely related to each other (Figure 3B,C). We propose that these sequences represent two novel virus species and have tentatively named them Santeuil virus (from JU1264) and Orsay virus (from JU1580).

Viral Detection and Confirmation of Viral Infection

RT-PCR assays were used to analyze RNA extracted from JU1580, JU1264, their corresponding bleached control strains JU1580bl and JU1264bl, and the same strains following reinfection with viral filtrates. Orsay virus RNA could be detected by RT-PCR in the original JU1580 culture, disappeared in the bleached strains, and stably reappeared following re-infection with the corresponding viral filtrate (Figure 4A). The same pattern applied for the Santeuil virus and JU1264 animals (Figure 4B). JU1580 and JU1264 cultures continuously propagated for 6 mo by transferring a piece of agar (approx. 0.1 cm^3) to the next plate twice a week continued to yield positive RT-PCR results (unpublished data).

Northern blotting confirmed the presence of Orsay and Santeuil virus RNA sequences in the infected animals. Hybridization with a DNA probe targeting the RNA1 segment of Santeuil virus yielded multiple bands in JU1264 animals but not in the corresponding bleached control strain. The strongest band detected migrated between 3.5 and 4 kb consistent with the 3,628 nt sequence we generated for the putative complete RNA1 segment (Figure 4C). Multiple higher molecular weight bands were also detected that may represent multimeric forms of the viral genomic RNAs, which have previously been described for some nodaviruses [13,14]. Northern blotting with a probe targeting the RNA2 segment (Figure 4C) yielded a major band that migrated at ~2.5 kb as well as fainter, higher molecular weight bands. Similar patterns were seen for both segments of Orsay virus (unpublished data).

To demonstrate virus replication in the infected animals, we performed Northern blotting using strand-specific riboprobes. For positive sense RNA viruses like nodaviruses, the negative sense RNA is only synthesized during active viral replication. It is not packaged in virions and typically exists in much lower quantities than the positive strand. Robust levels of the positive strand of the Santeuil virus RNA1 segment were detected (Figure 4D). Northern blotting with a riboprobe designed to hybridize to the negative sense strand detected a band of ~3.5 kb as well as higher molecular weight bands and a lower band of ~1.5 kb (~30-fold longer exposure than the positive sense blot; Figure 4D). While the precise nature of the high and low molecular weight species remains to be defined, the presence of multiple RNA species of negative sense polarity in JU1264 animals demonstrates bona fide replication of Santeuil virus in JU1264.

In order to determine the localization of Orsay viral RNA in infected animals, we performed RNA fluorescent in situ hybridization (FISH) using a probe complementary to the positive sense RNA1 segment of Orsay virus. Viral RNA was robustly detected in intestinal cells of JU1580bl animals infected 4 d previously with Orsay viral filtrate (Figure 4E, top panels). Interestingly, some animals also showed localization of viral RNA in the somatic gonad (Figure 4E, middle panels). JU1580bl animals not treated with the viral filtrate displayed no fluorescent signal (Figure 4E, bottom panels).

High Specificity of Infection by Orsay and Santeuil Nodaviruses

We tested whether the Orsay and Santeuil viruses could cross-infect the cured wild isolate of the other *Caenorhabditis* species, as well as the reference laboratory strains of *C. elegans* and *C. briggsae*. The Orsay and Santeuil viruses could only infect strains of *C. elegans* and *C. briggsae*, respectively (Figure 5A,B and Figure S3). Furthermore, each virus showed intraspecific specificity of infection. Indeed, we could not detect any replication of the Santeuil virus in *C. briggsae* AF16. The N2 laboratory *C. elegans* strain, while infectable by Orsay virus, appeared to be more resistant to viral infection than JU1580bl. Quantitative RT-PCR demonstrated that viral RNA accumulated in the N2 strain at levels above background but 50–100-fold lower than in JU1580bl (Figure 5C).

Small RNA Response upon Infection

One key defense mechanism of plants and animals against RNA viruses is the small RNA response [15]. We therefore determined by deep sequencing of small RNAs whether the infected animals produced small RNAs in response to viral infection. We generated small RNA libraries from mixed-stage JU1580 animals infected

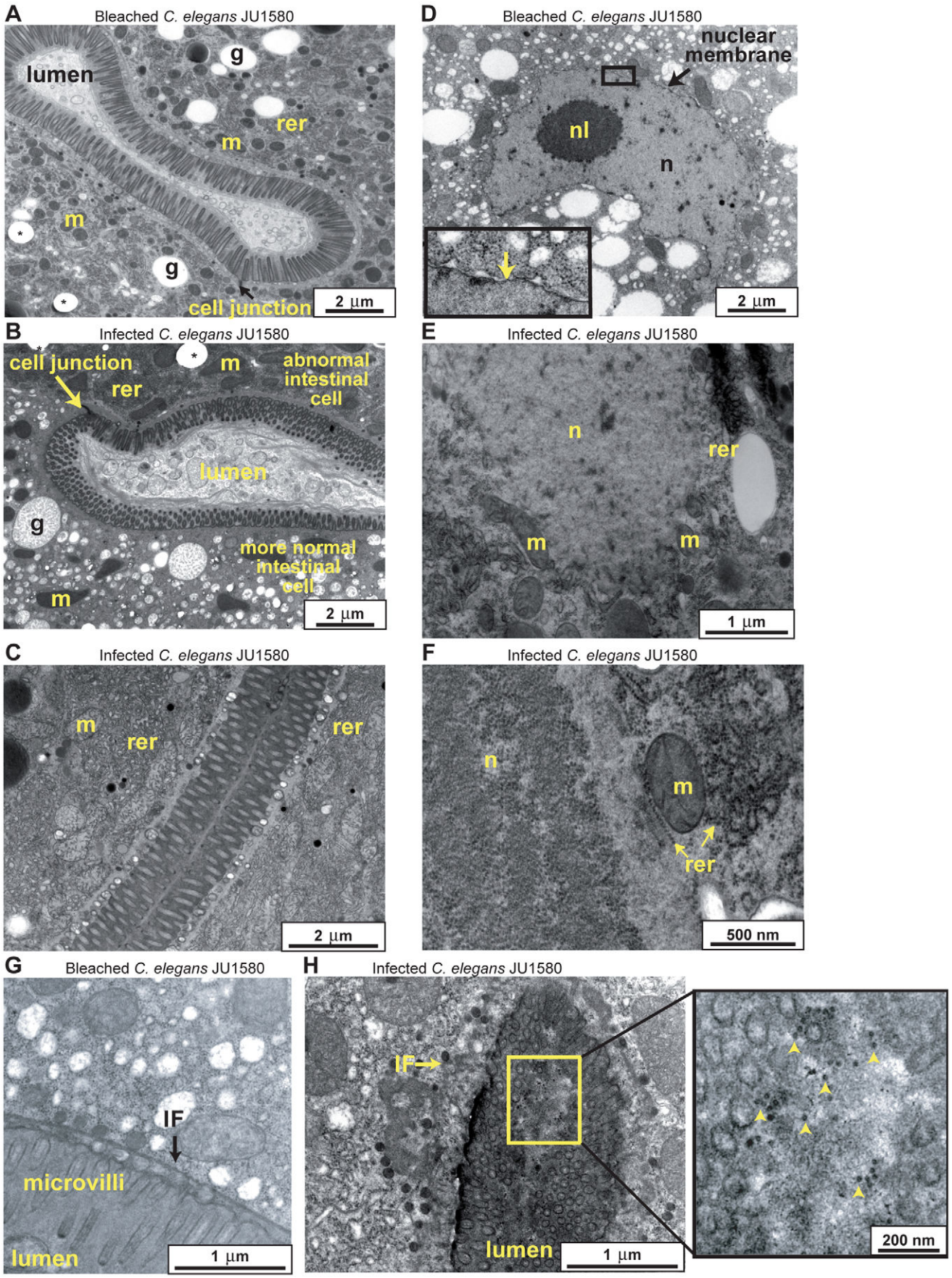

A
Bleached C. elegans JU1580
lumen
g
rer
m
m
g
cell junction
2 μm
D
Bleached C. elegans JU1580
nuclear membrane
nl
n
2 μm
B
Infected C. elegans JU1580
cell junction
m
rer
abnormal intestinal cell
lumen
g
more normal intestinal cell
m
2 μm
E
Infected C. elegans JU1580
n
rer
m
m
1 μm
C
Infected C. elegans JU1580
m
rer
rer
2 μm
F
Infected C. elegans JU1580
n
m
rer
500 nm
G
Bleached C. elegans JU1580
IF
microvilli
lumen
1 μm
H
Infected C. elegans JU1580
IF
lumen
1 μm
200 nm

Figure 2. Transmission electron micrographs of intestinal cells of *C. elegans* JU1580 adult hermaphrodites. (A,D,G) Bleached animals. (B–C, E–F, H) Naturally infected animals. (A–C) The infection provokes a reorganization of cytoplasmic structures, most visibly the loss of intestinal lipid storage granules (g). The cytoplasm of infected intestinal cells mostly contains rough endoplasmic reticulum (rer) and mitochondria (m). * hole in the resin used for inclusion in electron microscopy. (D–F) A nucleus in a non-infected animal is surrounded by a nuclear membrane (see inset in D), whereas the nuclear membrane disappears upon infection (E–F). Absence or incomplete nuclear membrane was observed repeatedly in infected animals, while the nuclear membrane could be observed on bleached animals (using both fixation methods). The nuclear material (n) in (F) may represent nucleolar material and at lower magnification (not shown) matches the shape of elongated nucleoli as observed by Nomarski optics (Figure 1E–G). The rough endoplasmic reticulum (rer) on the left of the mitochondrion (m) in (F) may be a remnant of the nuclear envelope. (G–H) The infection may result in disorganization of the intermediate filament (IF) network normally located below the apical plasma membrane. On the right of (H) is shown a higher magnification of the intestinal lumen, showing putative viral particles (arrowheads). The animals were fixed using high-pressure freezing (A–C, E–F) or conventional fixation (D, G, H).
doi:10.1371/journal.pbio.1000586.g002

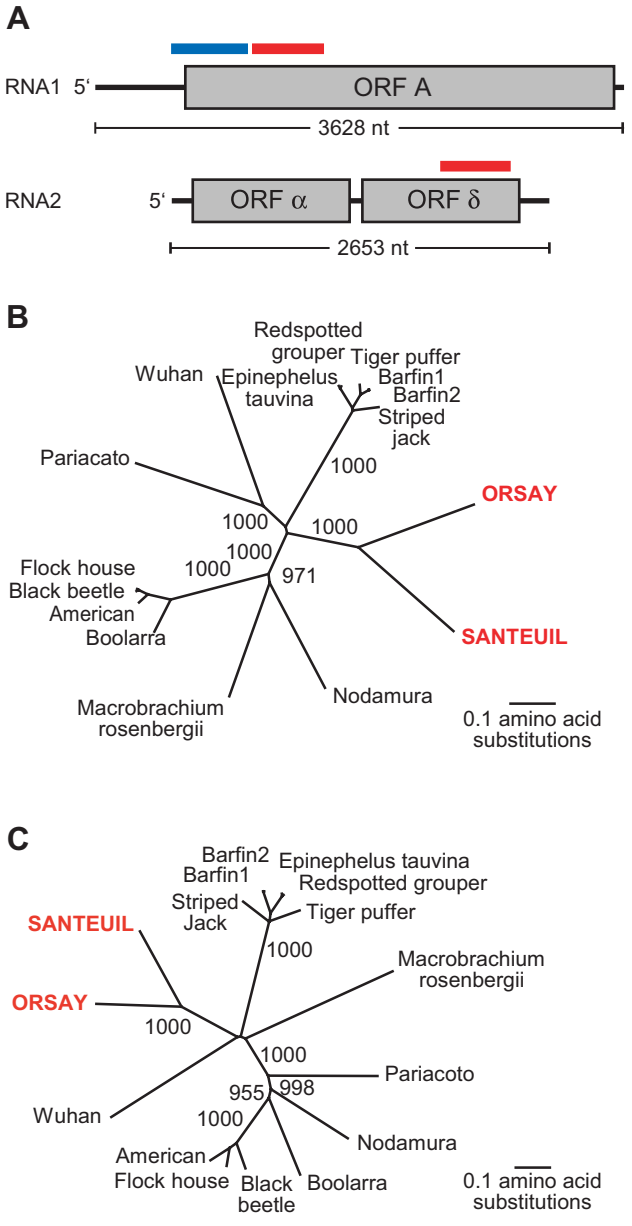


Figure 3. Genomic organization and phylogenetic analysis of novel viruses. (A) Schematic of genomic organization of Santeuil virus. Predicted open reading frames are displayed in gray boxes. Red bar indicates sequence used to generate double-stranded DNA probes for Northern blotting. Blue bar indicates sequence used to generate single-stranded riboprobes. (B) Neighbor-joining phylogenetic analysis of the predicted RNA-dependent RNA polymerases encoded by the RNA1 segments. (C) Neighbor-joining phylogenetic analysis of the predicted capsid proteins encoded by the RNA2 segments.
doi:10.1371/journal.pbio.1000586.g003

with the Orsay virus and from the bleached control strain and analyzed them using Illumina/Solexa high-throughput sequencing. These libraries represent small RNAs of 18–30 nucleotides in length independent of their 5′ termini. Small RNAs from infected JU1580 animals that mapped to viral RNA1 or RNA2 and had no match to the *C. elegans* genome are shown in Figure 6A and 6B, respectively. Of a total of 1,149,633 unambiguously mapped unique sequences, almost 2% (21,392) mapped to the two RNA segments of Orsay virus. Such RNAs were virtually absent from a library generated from bleached JU1580 animals (<0.001%) (unpublished data). Small RNAs that corresponded to the sense strand of the viral RNAs had a broad length distribution and no 5′ nucleotide preference. These sense small RNAs might represent Dicer cleavage products or other viral RNA degradation intermediates. In contrast, most antisense small RNAs were 22 nt long and showed a bias for guanidine as the first base (Figure 6A,B). This signature is reminiscent of a class of secondary RNAs named 22G RNAs that are thought to be downstream effectors of exogenous and endogenous small RNA pathways [16–18]. Such RNAs are not associated with transgenes expressed in the soma of *C. elegans* from extrachromosomal arrays [17] nor generally a feature associated with active transcription of endogenous genes [16–18]. These data suggest that JU1580 animals raise a small RNA response to viral infection. We also detected small RNAs of both sense and antisense polarity that mapped to the Santeuil virus genome in the JU1264 wild *C. briggsae* isolate but not in bleached animals (unpublished data).

RNAi Competency of the Host Is an Antiviral Defense

As viral infection appears to invoke a small RNA response in JU1580 animals, we next tested if mutations in small RNA pathways could affect replication of the Orsay virus. Orsay virus infection of the N2 reference strain was reduced compared to JU1580, as assayed by viral RNA qRT-PCR (Figure 5C) and infection symptoms (Figures S3A and 7B). Mutation of the *rde-1* gene—which encodes an Argonaute protein required for the initiation of exogenous RNAi [19]—in the N2 background increased viral RNA abundance and morphological symptoms to levels comparable to JU1580 using both assays (Figure 7A,B). The infected *rde-1* strain produced infectious viral particles, as reinfection of the cured JU1580 strain by filtrates of infected *rde-1* animals yielded positive RT-PCR results (unpublished data). In addition, mutation of other exogenous RNAi pathway genes including *rde-2*, *rde-4*, and *mut-7* (Table 1) also led to increased viral RNA accumulation as determined by quantitative RT-PCR (Figure 7A). We thus conclude that RNAi mechanisms provide antiviral immunity to *C. elegans* and that Orsay virus infection of mutant animals can be used to define genes important for antiviral defense.

Natural Variation in Somatic RNAi Efficiency in *C. elegans*

Since a functional RNAi pathway limits the accumulation of viral RNA in the N2 reference strain, we assessed the exogenous

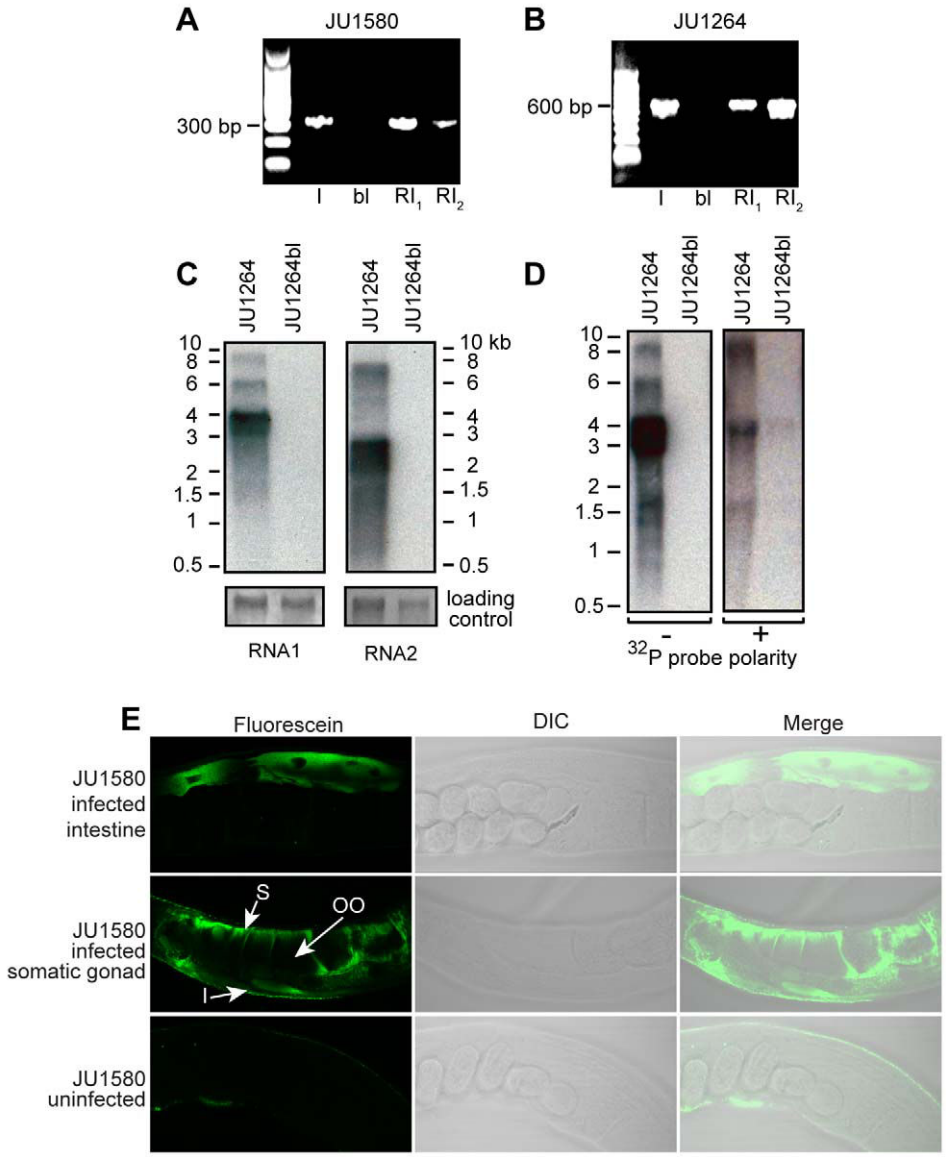


Figure 4. Molecular evidence of viral infection. (A) RT-PCR detection of the Orsay virus in the original JU1580 wild isolate (I), after bleaching (bl) and after re-infection by a 0.2 µM filtrate after 7 d (RI_1) and 3 wk (RI_2) of culture at 23°C. (B) RT-PCR detection of the Santeuil virus in the original wild isolate (I), after bleaching (bl) and after re-infection by a 0.2 µM filtrate after 4 d (RI_1) and 4 wk (RI_2) at 23°C. (C) Northern blots of Santeuil virus RNA1 and RNA2 segments hybridized using a double-stranded DNA probe. (D) Northern blots of Santeuil virus RNA1 segment using + and − sense riboprobes. (E) RNA FISH with a probe targeting Orsay virus RNA1 segment. Representative JU1580bl animals following infection by Orsay virus (top and middle rows) or uninfected (bottom row). S corresponds to ovary sheath cells, OO is an oocyte, and I is an intestinal cell.
doi:10.1371/journal.pbio.1000586.g004

RNAi competency of the bleached culture of JU1580 (JU1580bl) relative to the reference N2 strain. Using external application of dsRNAs by feeding, JU1580bl was found to be highly resistant to RNAi of a somatically expressed gene (*unc-22*) but competent for RNA inactivation of a germline-expressed gene (*pos-1*) (Figure 8A,B). *C. elegans* wild isolates, such as CB4856, were previously known to be variably sensitive to germline RNAi [20]. Here we thus observed for the first time a large variation in sensitivity to somatic RNAi, which does not correlate with germline RNAi sensitivity and thus cannot be due to inability to intake dsRNA from the intestinal lumen. We confirmed insensitivity to somatic RNAi of the JU1580bl isolate using a ubiquitously expressed GFP transgene (*let-858::GFP*), which was inactivated by GFP RNAi in the *C. elegans* N2 reference background, but only modestly repressed in the JU1580bl isolate (Figure 8C, Figure S4). We confirmed that the insensitivity to somatic RNAi also applied when *unc-22* dsRNAs were directly injected into the syncytial germline (Figure 8D). Therefore, the robust accumulation of Orsay virus RNA observed in infected JU1580 may be rendered possible in part by the partial defect in the somatic RNAi pathway of this wild isolate. The accumulation of small RNAs in response to the virus in infected JU1580 indicates, however, that its RNAi response is at least partially active in some tissues, perhaps including the germline.

The germline RNAi competence of JU1580 together with the presence of Orsay virus RNA1 in the somatic gonad raises the possibility that vertical transmission of viral infection could occur in a strain defective for germline RNAi. To examine this

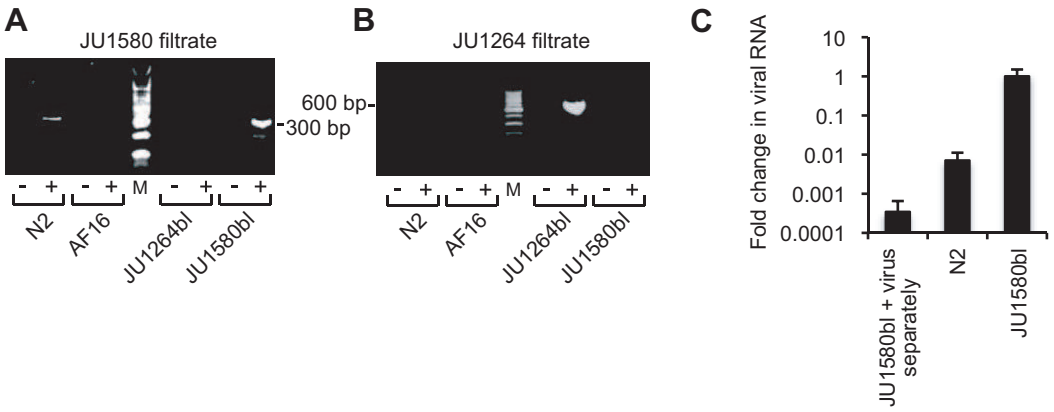


Figure 5. Specificity of infection by the Orsay and Santeuil viruses. (A) Specificity of infection by the Orsay virus. Each *Caenorhabditis* strain (name indicated below the gel) was mock-infected (−) or infected with a virus filtrate (+). RT-PCR on cultures after 7 d at 23°C. See Figure S3 for corresponding morphological symptom scoring. (B) Specificity of infection by the Santeuil virus. RT-PCR results after 4 d at 23°C. (C) Quantitative variation in viral replication N2 versus JU1580. N2 and JU1580 were tested by qRT-PCR for infection with Orsay virus extract ($n = 10$ independent replicates for each strain). By conventional RT-PCR assay, Orsay virus infection of N2 yielded positive bands in 3 out of 10 replicate infections whereas 7 out of 10 replicate infections of JU1580bl were positive in these conditions. Control RNA ($n = 6$) was extracted from JU1580bl animals grown in parallel without virus filtrate, and to which filtrate was added at the time of sample collection. RNA levels were normalized to *ama-1* and shown as average fold-change relative to JU1580bl. Error bars represent SEM.
doi:10.1371/journal.pbio.1000586.g005

possibility, JU1580bl, N2, and *rde-1* were exposed to Orsay virus filtrate. A subset of adult animals from each plate was bleached and their adult offspring collected 4 d later. No evidence for vertical transmission was observed by qRT-PCR for Orsay virus RNA in any strain (Figure S5).

We further tested the efficiency of the RNAi response in six other wild C. *elegans* isolates representative of its worldwide diversity (Figure 8A,B,D). Our results suggest that the somatic RNAi response varies quantitatively in *C. elegans* and is not correlated with germline RNAi sensitivity. Under experimental conditions that yield efficient infection of JU1580bl by Orsay virus, none of the other strains yielded significant levels of morphological symptoms (Figure 8E). Only JU1580bl and JU258 were positive by RT-PCR (unpublished data). Thus, factors other than RNAi competency also contribute to the sensitivity of *C. elegans* to the Orsay virus.

Discussion

The First Viruses Infecting *Caenorhabditis*

Here we report the first molecular description, to our knowledge, of viruses that naturally infect nematodes in the wild. The two novel viruses we identified, while clearly related to known nodaviruses, possess unique genomic features absent from all other previously described nodaviruses. These viruses may thus define a novel genus within the family *Nodaviridae* or may even represent prototype species of a new virus family (pending formal classification by the International Committee for the Taxonomy

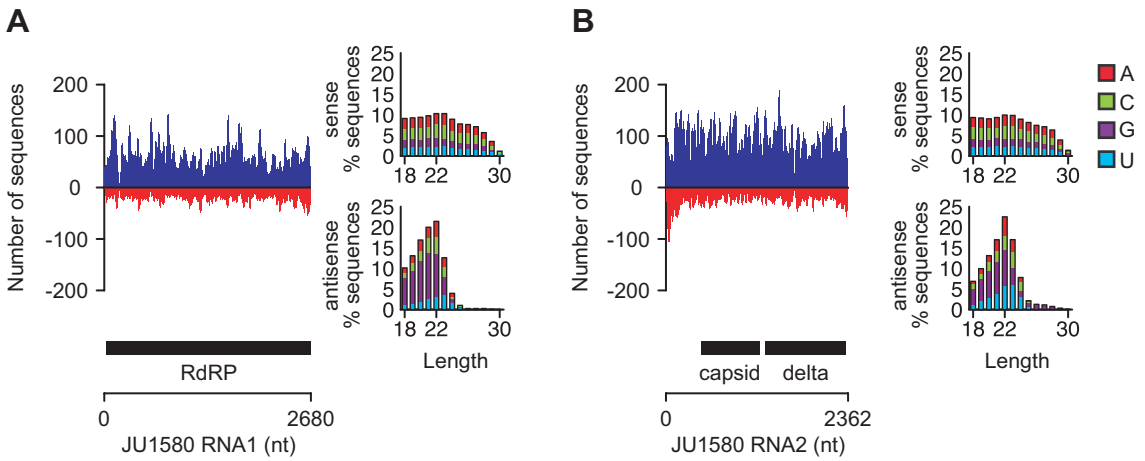


Figure 6. Small RNAs produced upon viral infection. Number of unique sequences obtained by Illumina/Solexa high-throughput sequencing of a 5′-independent small RNA library from JU1580 matching a given position in the Orsay virus segment RNA1 (A) or RNA2 (B). The number of sequences in sense and antisense orientation are shown on the positive (blue) and negative (red) *y*-axis, respectively. Only sequences with a perfect and unambiguous match to the virus genome were considered. The location of virus protein-coding genes is indicated below each graph as black bars and the RNA genome as a line. Features of sense and antisense sequences (length and identity of first nucleotide) are shown to the right of each graph.
doi:10.1371/journal.pbio.1000586.g006

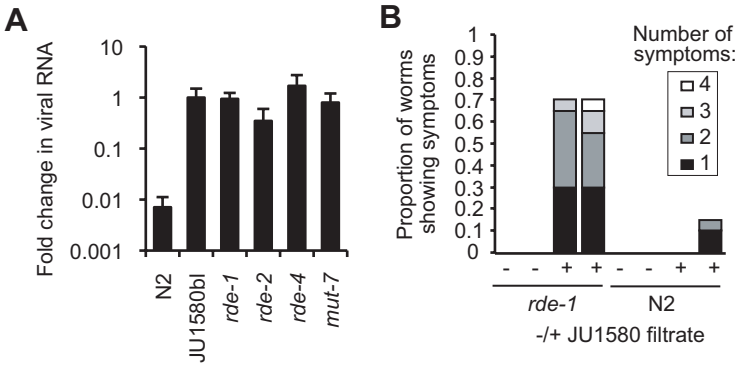


Figure 7. RNAi-deficient mutants of *C. elegans* can be infected by the Orsay virus. (A) JU1580bl, N2, *rde-1(ne219)* ($n = 10$ independent replicates each), *rde-2(ne221)*, *rde-4(ne301)*, and *mut-7(pk204)* ($n = 5$ independent replicates each) were tested by qRT-PCR for infection with Orsay virus extract. RNA levels were normalized to *ama-1* and shown as average fold-change relative to JU1580bl. Error bars represent SEM. Same results as displayed in Figure 5C for N2 and JU1580. (B) Scoring of symptoms in two independent replicates of infection of *rde-1* mutant and wild-type N2 animals by the Orsay virus filtrate, after 4 d.
doi:10.1371/journal.pbio.1000586.g007

of Viruses). The same range of intestinal symptoms was observed in animals that were infected by the Orsay and Santeuil viruses, further suggesting that these viral infections were causing the cellular symptoms. We observed putative viral particles of the size expected for nodaviruses, and a strong RNA FISH signal in intestinal cells and the somatic gonad of infected animals demonstrating that the virus is present intracellularly. It is likely that further sampling of natural populations of *Caenorhabditis* will yield other viruses of this and other groups. In fact, these symptoms were seen repeatedly in *C. briggsae* animals sampled from different locations in France, and in one instance, a Santeuil virus variant has been identified (unpublished data).

A characteristic feature of these two viruses is the presence of the novel ORF δ. Conservation of sequence length and identity of the ORF δ in these two viruses, and the absence of this ORF in all other described nodaviruses, suggests that this predicted protein is likely to be important for the ability of the virus to infect or replicate in nematodes. Its function is currently unknown, but it is tempting to speculate that this protein may play a role in antagonizing an innate antiviral pathway.

A Laboratory Viral Infection of a Small Model Animal

The infection of *C. elegans* by the Orsay nodavirus provides an exciting prospect for studies in virology, host cell biology, and antiviral innate immunity. Genetic screens to identify anti-viral factors in model organisms have been limited in large part by the lack of natural infection systems. Although *Drosophila* has been used with great success to examine host-virus interactions for various insect viruses [21] and influenza [1], none of these studies has examined viral infection of the host organism by natural transmission routes. Here we present a novel association between *C. elegans* and a virus that persists in culture through horizontal transmission, causing high damage in intestinal cells yet remarkably little effect on the animal, which continues moving, eating, and producing progeny, although at a lower rate.

De novo infection of naïve animals can be affected by the simple addition of either dead infected animals or homogenized lysates made from infected animals to culture dishes. This is sufficient to seed sustained complete cycles of viral replication, shedding, and infection. With this system, it is now possible to embark on whole genome genetic screens to identify host factors that block any facet of the viral life cycle. Using the current experimental conditions, infection of JU1580bl and *rde-1* mutants in N2 background was highly reproducible. The fact that the reference wild type N2 strain may only sustain a very low yet detectable viral titer makes it a particularly favorable genetic background in which to screen for genes involved in interaction with the virus.

The intestine is a tissue that is particularly exposed to microbes through ingestion, and is a main entry point for pathogens in *C. elegans* as in other animals. In *C. elegans*, the intestinal cells are large and easily amenable to observations by optical microscopy. The viral parasites affect the organization of the polarized epithelial intestinal cells and will likely provide interesting mechanisms and tools to study their cell biology. Clear reorganization occurs in the intermediate filaments that line the apical brush border, as well as in the lipid storage granules, the nuclear membrane, and other intracellular compartments.

The abnormal state of the intestinal cells may slow down progeny production by decreasing the food intake. Alternatively, the presence of viral RNA in the somatic gonad may explain the delay in progeny production, although no gonadal cellular phenotypes have been observed. The presence of viral RNA in the somatic gonad is particularly interesting given the lack of vertical transmission.

Targeted Mutant Screens with Orsay Virus Confirm a Role for RNAi in Antiviral Defense

Although prior studies have clearly demonstrated a role for *C. elegans* RNAi in counteracting viral infection, these studies utilized either a transgenic system of viral RNA expression [7] or primary culture cells [22,23]. The observed susceptibility of Orsay virus RNA to RNAi processing in JU1580 animals provides the first evidence in a completely natural setting, without any artificial manipulations, that RNAi serves an antiviral role in nematodes. Coupled to the increase in accumulation of Orsay virus RNA in RNAi pathway mutant strains as compared to wild type N2, these studies demonstrate that the RNAi pathway is an important antiviral defense against Orsay virus. Moreover, these results demonstrate the feasibility of identifying antiviral genes or pathways in this experimental infection system. The mechanism by which the animals prevent transmission to their offspring is unclear, but our initial results with *rde-1* mutants suggest that perturbing germline RNAi is not sufficient to enable vertical transmission.

Evolution of Viral Sensitivity and Specificity in Natural Populations

The quantitative difference in Orsay nodavirus sensitivity between the N2 and JU1580 wild *C. elegans* genetic backgrounds

will allow the identification of a set of host genes that modulate viral sensitivity during evolution of natural host populations. Based on the defect in exogenous RNAi of the JU1580 strain, we speculate that this set will include, but is unlikely to be limited to, genes involved in exogenous RNAi pathways. Support for the role of other genes outside the RNAi pathway comes from our data on natural isolates. Despite the fact that the magnitude of the somatic RNAi defect of the natural isolate PS2025 was comparable to that of JU1580, no evidence of viral RNA accumulation or morphological symptoms was observed following addition of

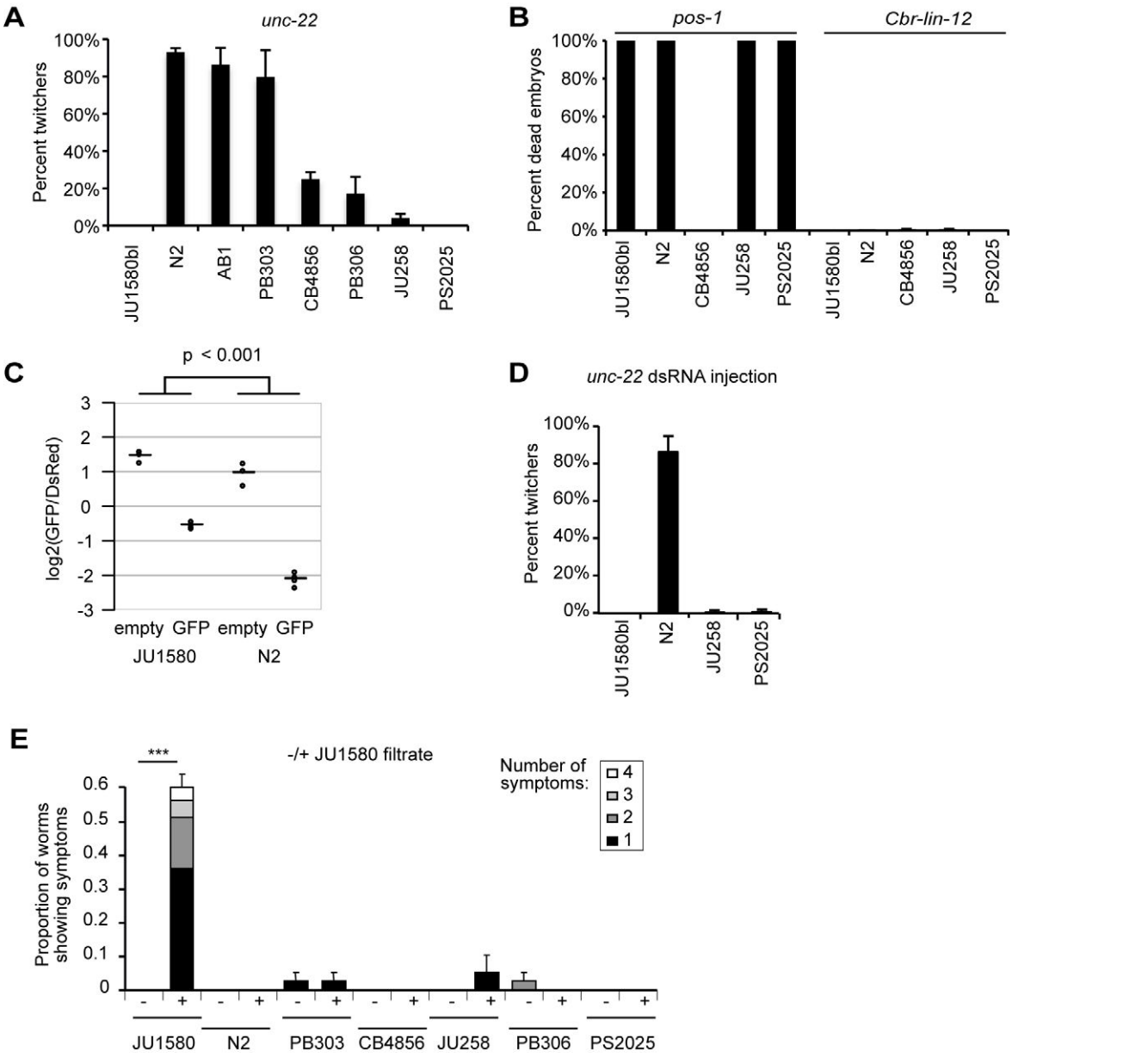


Figure 8. Natural variation in somatic RNAi efficacy in *C. elegans*. (A) Somatic RNAi was tested using bacteria expressing dsRNA specific for the *unc-22* gene (acting in muscle; [37]). The percentage of animals with the corresponding twitcher phenotype is shown for different *C. elegans* wild isolates (representative of the species' diversity; [38]). Bar: standard error over four replicate plates. (B) Germline RNAi was tested by feeding the animals with bacteria expressing dsRNA specific for the *pos-1* gene. The percentage of animals with the corresponding embryonic-lethal phenotype is shown for five wild genetic backgrounds of *C. elegans*. *Cbr-lin-12* RNAi is a negative control. Bar: standard error over six replicate plates (too small to be seen). $n>450$ observed individuals for each treatment. (C) Somatic RNAi was tested using bacteria expressing dsRNA specific for GFP. Each point corresponds to the median $\log_2$(GFP/DsRed) intensity ratio from one flow cytometry run of strains carrying the *let-858::GFP* transgene in the JU1580 and N2 backgrounds, after treatment with GFP RNAi or empty vector. Horizontal bars indicate group means. The difference in $\log_2$ intensity ratios between GFP RNAi and empty vector is reduced in JU1580 compared to N2 ($p<0.001$, see Methods). (D) *unc-22* dsRNA was administered by injection into the syncytial germline of the mother. 10–14 animals of each genotype were injected and 30 progeny were scored for the twitcher phenotype on each plate. (E) Orsay virus sensitivity of seven wild *C. elegans* isolates representative of the species' diversity. Morphological symptoms were scored 5 d after infection of clean cultures by the Orsay virus filtrate at 23°C. The JU1580 control was performed in duplicate. Bar: standard error on total proportion. *** $p<0.001$.

doi:10.1371/journal.pbio.1000586.g008

Orsay virus filtrate. Whether PS2025 lacks one or more crucial receptors for viral infection or has alternative antiviral pathways that suppress viral replication is currently unknown.

In addition, the Orsay and Santeuil viruses appear to specifically infect *C. elegans* and *C. briggsae*, respectively. Moreover, the *C. elegans rde-1* mutation in the N2 background confers susceptibility to the Orsay virus, but not to the Santeuil virus (Figure S3C). The two viruses thus provide a system to study host-parasite specificity and its evolution. With the isolation of additional variants of each virus (our unpublished data), viral evolution studies can also be undertaken. Host-parasite evolutionary and ecological interactions can thus be explored at two evolutionary scales, within and between species of both host and parasite. The rapid life cycle of *C. elegans* also allows experimental evolution in the laboratory [24,25]. This model system, which can include both natural and engineered variants of both virus and host, is thus favorable for combining studies of host-pathogen co-evolution in the laboratory and in natural populations.

Materials and Methods

Nematode Field Isolation

Caenorhabditis nematodes were isolated on *C. elegans* culture plates seeded with *E. coli* strain OP50 using the procedures described in [10]. JU1264 was isolated from a snail collected on rotting grapes in Santeuil (Val d'Oise, France) on 14 Oct 2007. JU1580 was isolated from a rotting apple sampled in Orsay (Essonne, France) on 6 Oct 2008. When required, cultures were cleared of natural bacterial contamination by frequent passaging of the animals and/or antibiotic treatment (LB plates with 50 µg/ml tetracycline, ampicilline, or kanamycine for 1 h). Infected cultures were kept frozen at −80°C and in liquid N2 as described in [26]. Bleaching was performed as in [26].

Light Microscopy

When observed with a transillumination dissecting microscope, infected animals displayed a paler intestine than healthy worms. This lack of intestinal coloration occurred all along the entire intestinal tract in *C. briggsae* JU1264 and preferentially in the anterior intestinal tract in *C. elegans* JU1580. Intestinal cells were observed with Nomarski optics with a 63× or 100× objective. The four symptoms used for scoring were 1, the disappearance of gut granules in at least part of a cell; 2, degeneration of the nucleus including a very elongated nuclear or nucleolus (when the rest of the nucleus has degenerated) or the apparent disappearance of the nucleus; 3, the loss of cytoplasmic viscosity visible as a very fluid flow of cytosol within the cell; and 4, the fusion of intestinal cells. Some of these traits may sometimes appear in uninfected animals. We systematically tested for a significant increase after infection of the proportion of animals with symptoms (Fisher's exact test). Note that some of these symptoms can also be caused by microsporidial and bacterial infections. Thus, the diagnostic of a viral infection based on the cellular symptoms requires an otherwise clean culture.

Live Hoechst 33342 Staining of Nuclei

Animals were washed off a culture plate in 10 ml of ddH_20, pelleted and incubated in 10 ml of 10 µg/ml Hoechst 33342 in ddH_20 for 45 min with soft agitation, protecting the tube from light with an aluminum foil. The animals were then pelleted and transferred to a new culture plate seeded with *E. coli* OP50. After 2 h, they were mounted and observed with a fluorescence microscope.

Electron Microscopy

A few adults were washed in 0.2 ml of M9 solution, suspended in 2% paraformaldehyde +0.1% glutaraldehyde, and cut in two on ice under a dissecting microscope for better reagent penetration [27]. Worm pieces were then resuspended overnight in 2% OsO_4 at 4°C, washed, embedded in 2% low melting point agar, dehydrated in solutions of increasing ethanol concentrations, and embedded in resin (Epon-Araldite). High-pressure freezing was performed using a Leica PACT2 high-pressure freezer [28].

Progeny Counts

The time course was started by isolating single L4 larvae for *C. elegans* JU1580 and single L3 larvae for *C. briggsae* JU1264. The parent animal then transferred every day to a new plate until the end of progeny production. The plates were incubated at 20°C for 2 d and kept at 4°C until scoring. The few cases where the parent died before the end of its laying period were not included. Some progeny died as embryos in both infected and non-infected cultures (non-significant effect of treatment; unpublished data). The timing of progeny production was analyzed in R using a Generalized Linear Model using infection status, day, individual (nested in infection status), and Infection Status×Day as explanatory variables, assuming a Poisson response variable and a log link function. Individual, day and Infection Status×Day were the significant explanatory variables for both JU1264 and JU1580 ($p<0.001$).

Infectious Filtrate Preparation and Animal Infections

Nematodes were grown on 10 plates (90 mm diameter) until just starved, resuspended in 15 ml of 20 mM Tris-Cl pH 7.8, and pelleted by low-speed centrifugation (5,000 g). The supernatant was centrifuged twice at 21,000 g for 5 min (4°C) and pellets discarded. The supernatant was passed on a 0.2 µm filter. 55 mm culture plates were prepared with 2–5 young adults of N2, *rde-1(ne219)*, or JU1580bl. At the same time (Figures 1K, 4A,B, 5, 7B, and S3), or the following day (Figures 5C, 7A), 30 µl of infectious filtrate was pipetted onto the bacterial lawn. The cultures were incubated at 20°C except otherwise indicated. When both *C. elegans* and *C. briggsae* were grown in parallel, an incubation temperature of 23°C (indicated in the figure legends) was used so that both species developed at similar speeds. Maintenance over more than 4 d after re-infection was performed by transferring a piece of agar (approx. 0.1 cm^3) every 2–3 d to a new plate with food.

High-Throughput Sequencing

Phenol-chloroform purified DNA and RNA from infected JU1580 and JU1264 animals were subject to random PCR amplification as described [29]. The amplicons were then pyrosequenced following standard library construction on a Roche Titanium Genome Sequencer. Raw sequence reads were filtered for quality and repetitive sequences. BLASTn and BLASTx were used to identify sequences with limited similarity to known viruses in Genbank. Contigs were assembled using the Newbler assembler. To confirm the assembly, primers for RT-PCR were designed to amplify overlapping fragments of ~1.5 kb. Amplicons were cloned and sequenced.

5′ and 3′ RACE

5′ RACE was performed according to standard protocols (Invitrogen 5′ RACE kit). 3′ RACE was performed by first adding a polyA tail using PolyA polymerase (Ambion) and then using Qiagen 1-step RT-PCR kit with gene specific primers and an oligo-dT-adapter primer. Products were cloned into pCR4 and sequenced using standard Sanger chemistry.

Small RNA Sequencing

4–6 90 mm plates with 15–20 adults (JU1580 or bleached JU1580) were grown for 4 d at 20°C. Mixed stage animals from all plates were collected, pooled, and frozen at −80°C. Total RNA was extracted using the mirVana miRNA isolation kit (Ambion). Small RNAs were size selected to 18–30 bases by denaturing polyacrylamide gel fractionation. A cDNA library that did not depend on 5′-monophosphates was constructed by tobacco acid pyrophosphatase treatment using adapters recommended for Solexa sequencing as described previously [30]. Each sample was labeled with a unique four base pair barcode. cDNA was purified using the NucleoSpin Extract II kit (Macherey & Nagel). Small RNA libraries were sequenced using the Illumina/Solexa GA2 platform (Illumina, Inc., San Diego, CA). Fastq data files were processed using custom Perl scripts. Reads with missing bases or whose first four bases did not match any of the expected barcodes were excluded. Reads were trimmed by removing the first four nucleotides and any 3′ As. The obtained inserts were collapsed to unique sequences, retaining the number of reads for each sequence. Sequences in the expected size range (18–30 nucleotides) were aligned to the *C. elegans* genome (WS190) downloaded from the UCSC Genome Browser website (http://genome.ucsc.edu/) [31] and the JU1580 partial virus genome using the ELAND module within the Illumina Genome Analyzer Pipeline Software, v0.3.0. Figure 6 is based on unique sequences (multiple reads of the same sequence were collapsed) with perfect and unambiguous alignment to the Orsay virus genome. Small RNA sequence data were submitted to the Gene Expression Omnibus under accession number GSE21736.

Neighbor-Joining Phylogenetic Analysis

The predicted amino acid sequences from Orsay and Santeuil nodaviruses were aligned using ClustalW to the protein sequences of the following nodaviruses. Capsid Protein: Barfin1 flounder nervous necrosis virus NC_013459, Barfin2 flounder virus BF93Hok RNA2 NC_011064, Black beetle virus NC_002037, Boolarra virus NC_004145, *Epinephelus tauvina* nervous necrosis virus NC_004136, Flock house virus NC_004144, *Macrobrachium rosenbergii* nodavirus RNA-2 NC_005095, Nodamura virus RNA2 NC_002691, Pariacoto virus RNA2 NC_003692, Redspotted grouper nervous necrosis virus NC_008041, Striped Jack nervous necrosis virus RNA2 NC_003449, Tiger puffer nervous necrosis virus NC_013461, Wuhan nodavirus ABB71128.1, and American nodavirus ACU32796.1. 1,000 bootstrap replicates were performed.

RNA Polymerase: Barfin flounder nervous necrosis virus YP_003288756.1, Barfin flounder virus BF93Hok YP_002019-751.1, Black beetle virus YP_053043, Boolarra virus NP_689439, *Epinephelus tauvina* nervous necrosis virus NP_689433.1, Flock house virus NP_689444.1, Nodamura virus NP_077730, Pariacoto virus NP_620109.1, Redspotted grouper nervous necrosis virus YP_611155.1, Striped Jack nervous necrosis virus NP_599-247.1, Tiger puffer nervous necrosis virus YP_003288759.1, *Macrobrachium rosenbergii* nodavirus NP_919036.1, Wuhan_Nodavirus AAY27743, and American nodavirus SW-2009a ACU32-794.1. 1,000 bootstrap replicates were performed.

RT-PCR

Nematodes from two culture plates were resuspended in M9 and then washed three times in 10 ml M9. RNA was extracted using Trizol (Invitrogen) (5–10 vol:vol of pelleted worms) and resuspended in 20 µl in RNAse-free ddH_2O. 5 µg of RNA were reverse transcribed using SuperscriptIII (Invitrogen) in a 20 µl volume. 5 µl were used for PCR in a 20 µl volume (annealing temperature 60°C, 35 cycles). For the Orsay nodavirus, the reverse transcription used the GW195 primer (5′ GACGCTTCCAA-GATTGGTATTGGT) and the PCR oTB3 (5′ CGGATTCTC-GACATAGTCG) and oTB4 (5′GTAGGCGAGGAAGGAG-ATG). For the Santeuil nodavirus, reverse transcription used oTB6RT (5′ GGTTCTGGTGGTGATGGTG) and PCR oTB5 (5′ GCGGATGTTCTTCACGGAC) and oTB6 (5′ GTCAG-TAGCGGACCAGATG).

One-Step RT-PCR

Animals from one 55 mm culture plate plus viral filtrate (see infection procedure) were washed twice in M9. RNA was extracted using 1 ml Trizol (Invitrogen) and resuspended in 10 µl DEPC-treated H_2O. 0.1 µl was used for RT-PCR using the OneStep RT-PCR Kit (Qiagen). Primers annealed to viral RNA1 (GW194 and GW195).

qRT-PCR

cDNA was generated from 1 µg total RNA with random primers using Superscript III (Invitrogen). cDNA was diluted to 1:100 for qRT-PCR analysis. qRT-PCR was performed using either QuantiTect SYBR Green PCR (Qiagen) or ABsolute Blue SYBR Green ROX (Thermo Scientific). The amplification was performed on a 7300 Real Time PCR System (Applied Biosystems). Each sample was normalized to *ama-1*, and then viral RNA1 (primers GW194: 5′ ACC TCA CAA CTG CCA TCT ACA and GW195: 5′ GAC GCT TCC AAG ATT GGT ATT GGT) levels were compared to those present in re-infected bleached JU1580 animals.

Northern Blotting

For Northern blots, 0.5 µg of total RNA extracted from JU1264 and JU1264bl animals were electrophoresed through 1.0% denaturing formaldehyde-MOPS agarose gels. RNA was transferred to Hybond nylon membranes and then subject to UV cross-linking followed by baking at 75°C for 20 min. Double stranded DNA probes targeting the RNA1 segment of Santeuil nodavirus (nt 1141–1634) and the RNA2 segment of Santeuil nodavirus (nt 1833–2308) were generated by random priming in the presence of α-^{32}P dATP using the Decaprime kit (Ambion). Blots were hybridized for 4 h at 65°C in Rapid hyb buffer (GE Healthcare) and washed in 2XSSC/0.1%SDS 5 min×2 at 25°C, 1XSSC/0.1%SDS 10 min×2 at 25°C, 0.1XSSC/0.1%SDS 5 min×4 at 25°C, and 0.1XSSC/0.1%SDS 15 min×2 at 42°C and 0.1XSSC/0.1%SDS 15 min×1 at 68°C. For strand specific riboprobes, ^{32}P labeled RNA was generated by in vitro transcription with either T7 or T3 RNA polymerase (Ambion) in the presence of α-^{32}P UTP. The target plasmid contained a cloned region of the Santeuil nodavirus RNA1 segment (nt 523–1022) and was linearized with either PmeI or NotI, respectively. For the riboprobes, blots were hybridized at 70°C and then sequentially washed as follows: 2XSSC/0.1%SDS 5 min×2 at 68°C, 1XSSC/0.1%SDS 10 min×2 at 68°C, 0.1XSSC/0.1%SDS 10 min×2 at 68°C, and 0.1XSSC/0.1%SDS 20 min×1 at 73°C. The Santeuil RNA1 segment migrates at approximately the same position as the 28S ribosomal RNA. Under the extended exposure time (72 h) needed to visualize the negative sense genome, low levels of non-specific binding to the 28S RNA become apparent (Figure 4D).

RNA Interference

For *pos-1* and *unc-22* RNAi using bacteria as the dsRNA source, bacterial clones from the Ahringer library expressing dsRNAs [32]

(available through MRC Geneservice) were used to feed *C. elegans* on agar plates. For the *pos-1* experiment, bacteria were concentrated 10-fold by centrifugation prior to seeding the plates. A *C. briggsae Cbr-lin-12* fragment [33] was used as a negative control as it does not match any sequence in *C. elegans*. Three or four L4s were deposited on an RNAi plate, singly transferred the next day to a second RNAi plate, and their progeny scored after 2 d (*pos-1*) or 3 d (*unc-22*) at 23°C.

For *unc-22* dsRNA synthesis and injection, the *unc-22* fragment in the Ahringer library clone was amplified by PCR using the T7 primer and in vitro transcribed with the T7 polymerase using the Ambion MEGAscript kit, according to the manufacturer's protocol [34]. *Cel-unc-22* dsRNAs were injected at 50 ng/µl into both gonadal arms of young hermaphrodite adults of the relevant strain. The animals were incubated at 20°C. The adults were transferred to a new plate individually on the next day, and the proportion of twitching progeny scored 3 d later, touching each animal with a platinum-wire pick to induce movement.

For GFP RNAi, transgenic N2 and JU1580 strains were generated expressing the ubiquitously expressed *let-858::GFP* and the pharyngeal marker *myo-2::DsRed* as an extrachromosomal array. Bacteria expressing dsRNA against GFP cDNA were used to feed animals on agar plates. An empty vector was used as a negative control. Two or three L4s were deposited on a 55 mm RNAi plate, grown at 20°C for 3 d, and the GFP/DsRed expression levels in their offspring measured using flow cytometry (Union Biometrica) as described previously [35]. Offspring from two RNAi plates were combined for sorting. Each combination of RNAi vector and strain was repeated in at least triplicate. GFP and DsRed intensities were obtained from 14 wormsorter runs including 3–4 replicate runs for N2 and JU1580 after treatment with GFP RNAi or empty vector. A larger proportion of N2 animals showed reporter expression compared to JU1580 animals (Figure S4, top). To control for this difference between strains, animals with no reporter expression were excluded by requiring DsRed intensities to exceed a cutoff set to the median 99th percentile from three control runs of animals with no array present (Figure S4). A linear regression model was fitted to the median log2(GFP/DsRed) intensity ratios including strain, treatment, and an interaction term as explanatory variables. The interaction term was significantly different from zero at $p<0.001$.

RNA Fluorescent In Situ Hybridization (FISH)

A segment of Orsay virus RNA1 was generated with primers GW194 and GW195 and cloned into pGEM-T Easy (Promega). Fluorescein labeled probe was generated from linearized plasmid using the Fluorescein RNA Labeling Mix (Roche) and MEGAscript SP6 transcription (Ambion). JU1580bl animals were infected with Orsay virus filtrate and grown for 4 d at 20°C on 90 mm plates. Control animals were grown under the same conditions in the absence of virus. In situ hybridization was performed essentially as previously described [36]. The fluorescent RNA probe was visualized directly on an Olympus FV1000 Upright microscope.

Genbank Sequences: Accession numbers for Orsay and Santeuil virus contigs: HM030970–HM030973. Small RNA sequencing data at GEO: GSE21736.

Supporting Information

Figure S1 Putative viral particles in transmission electron microscopy of intestinal cells of infected *C. elegans* JU1580 adult hermaphrodites. On the right are shown higher magnifications of the parts delimited by a black rectangle in the left micrograph, showing putative viral particles (arrows). (A) Putative viral particles are found in an intracellular multi-membrane compartment. The particles in the upper left of the inset of (A) are ribosomes; the putative viral particles in the lower two-thirds are characterized by a slightly larger and more regular ring appearance (ca. 20 nm diameter). (B) Putative viral particles similar to those in Figure 2H are visible in the intestinal lumen, close to microvilli. These particles are clearly larger and distinct from the ribosomes (r) seen on the lower left of the inset. The animals were fixed using conventional fixation in (A) and high-pressure freezing in (B).
Found at: doi:10.1371/journal.pbio.1000586.s001 (3.41 MB PPT)

Figure S2 Infection does not alter brood size, but results in delayed progeny production in *C. briggsae* JU1264. (A,B) Boxplots of the distribution of brood size in naturally infected and bleached ("bl") cultures of *C. elegans* JU1580 (A) and *C. briggsae* JU1264 (B) at 20°C. The line indicates the median, the box the lower and upper quartiles, and the whiskers the 10th and 90th percentiles. Brood size is not significantly different in infected versus non-infected cultures (Wilcoxon test $p = 1.0$ for JU1264, $p = 0.81$ for JU1580). (C,D) Progeny number over time in JU1580 (C) and JU1264 (D). Infection results in a significant change in the timing of progeny production (Generalized linear model, Treatment×Day: $p<0.001$ in both cases). Note that time 0 corresponds to the L4 stage in the experiment in (A,C) and the L3 stage in (B,D).
Found at: doi:10.1371/journal.pbio.1000586.s002 (0.33 MB PDF)

Figure S3 Scoring of morphological symptoms after exposure of various wild isolates and *rde-1* mutants. (A) Specificity of infection by the Orsay nodavirus. Each *Caenorhabditis* strain was mock-infected (−) or infected with a virus filtrate (+). The proportion of worms with morphological infection symptoms after 7 d at 23°C is shown for the same experiment as in Figure 5A. (B) Specificity of infection by the Santeuil nodavirus. The proportion of worms with morphological infection symptoms after 4 d at 23°C is shown for the same experiment as in Figure 5B. (C) Santeuil virus sensitivity of *rde-1* mutants in the *C. elegans* N2 background. Morphological symptoms were scored 5 d after infection at 23°C by the Santeuil virus filtrate. * $p<0.05$; *** $p<0.001$.
Found at: doi:10.1371/journal.pbio.1000586.s003 (0.29 MB PDF)

Figure S4 Quantification of GFP transgene expression and silencing. (Top) Reporter genes are expressed in a larger proportion of animals from N2 compared to JU1580. Shown are the number of animals according to binned $\log_2$ DsRed intensities. Each line corresponds to one flow cytometry run with colors indicating strain and treatment as explained in the color legend. The extrachromosomal array was inherited more efficiently in N2 than in JU1580, making it necessary to analyze only those animals carrying the array. (Bottom) Each data point corresponds to the difference between treatment with GFP RNAi and empty vector in N2 (blue) and JU1580 (red) observed for animals with $\log_2$ DsRed intensity in a given bin. Dots indicate the difference between the means of median $\log_2$(GFP/DsRed) ratios for treatment with GFP RNAi and empty vector. Vertical bars indicate standard errors. The cutoff for DsRed intensities is indicated in both panels by a vertical dotted line.
Found at: doi:10.1371/journal.pbio.1000586.s004 (0.29 MB AI)

Figure S5 Quantitative analysis of vertical transmission of Orsay virus. N2, JU1580bl, and *rde-1* (5 replicates each) were infected with Orsay virus. After 4 d 25 adults from each plate were bleached onto a new, uninfected plate. The remaining adults were collected for RNA extraction. The offspring from the bleached

adults were collected for RNA extraction after 4 d. As a control, virus was added to plates and incubated in the absence of animals for 4 d. Viral RNA levels were determined by qRT-PCR and normalized to *gapdh*. Viral RNA level is shown on a log scale, using a reference value of 1 for the infection of JU1580.
Found at: doi:10.1371/journal.pbio.1000586.s005 (0.27 MB AI)

Acknowledgments

M.A.F. thanks the caretakers of the Orsay orchard for access to it and J.-L. Bessereau, F. Duveau, R. Legouis, N. Naffakh, and B. Samuel for helpful discussions.

Author Contributions

The author(s) have made the following declarations about their contributions: Conceived and designed the experiments: MAF EAM DW. Performed the experiments: MAF AA JP GW IN TB YJ CJF MS. Analyzed the data: MAF AA JP GW IN TB GZ CJF LDG MS EAM DW. Wrote the paper: MAF AA EAM DW.

References

1. Hao L, Sakurai A, Watanabe T, Sorensen E, Nidom CA, et al. (2008) Drosophila RNAi screen identifies host genes important for influenza virus replication. Nature 454: 890–893.
2. Sabin LR, Zhou R, Gruber JJ, Lukinova N, Bambina S, et al. (2009) Ars2 regulates both miRNA- and siRNA-dependent silencing and suppresses RNA virus infection in *Drosophila*. Cell 138: 340–351.
3. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, et al. (2002) A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. Science 297: 623–626.
4. Powell JR, Kim DH, Ausubel FM (2009) The G protein-coupled receptor FSHR-1 is required for the *Caenorhabditis elegans* innate immune response. Proc Natl Acad Sci U S A 106: 2782–2787.
5. Troemel ER, Felix MA, Whiteman NK, Barriere A, Ausubel FM (2008) Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. PLoS Biol 6: 2736–2752. doi:10.1371/journal.pbio.0060309.
6. Liu WH, Lin YL, Wang JP, Liou W, Hou RF, et al. (2006) Restriction of vaccinia virus replication by a *ced-3* and *ced-4*-dependent pathway in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A 103: 4174–4179.
7. Lu R, Maduro M, Li F, Li HW, Broitman-Maduro G, et al. (2005) Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. Nature 436: 1040–1043.
8. Lu R, Yigit E, Li WX, Ding SW (2009) An RIG-I-Like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in *Caenorhabditis elegans*. PLoS Pathog 5: e1000286. doi:10.1371/journal.ppat.1000286.
9. Neil SJ, Zang T, Bieniasz PD (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 451: 425–430.
10. Barrière A, Félix M-A (2006) Isolation of *C. elegans* and related nematodes. In: Community The *C. elegans* Research Community, editor. Wormbook.
11. Hüsken K, Wiesenfahrt T, Abraham C, Windoffer R, Bossinger O, Leube RE (2008) Maintenance of the intestinal tube in *Caenorhabditis elegans*: the role of the intermediate filament protein IFC-2. Differentiation 76: 881–896.
12. Li H, Li WX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. Science 296: 1319–1321.
13. Ball LA (1994) Replication of the genomic RNA of a positive-strand RNA animal virus from negative-sense transcripts. Proc Natl Acad Sci U S A 91: 12443–12447.
14. Johnson KN, Zeddam JL, Ball LA (2000) Characterization and construction of functional cDNA clones of Pariacoto virus, the first Alphanodavirus isolated outside Australasia. J Virol 74: 5123–5132.
15. Aliyari R, Ding SW (2009) RNA-based viral immunity initiated by the Dicer family of host immune receptors. Immunol Rev 227: 176–188.
16. Gu W, Shirayama M, Conte D, Jr., Vasale J, Batista PJ, et al. (2009) Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans germline*. Mol Cell 36: 231–244.
17. Pak J, Fire A (2007) Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. Science 315: 241–244.
18. Sijen T, Steiner FA, Thijssen KL, Plasterk RH (2007) Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science 315: 244–247.
19. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, et al. (1999) The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99: 123–132.
20. Tijsterman M, Okihara KL, Thijssen K, Plasterl RHA (2002) PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of *C. elegans*. Curr Biol 12: 1535–1540.
21. Huszar T, Imler JL (2008) Drosophila viruses and the study of antiviral host-defense. Adv Virus Res 72: 227–265.
22. Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A 102: 18420–18424.
23. Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, et al. (2005) RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. Nature 436: 1044–1047.
24. Azevedo RB, Keightley PD, Lauren-Maatta C, Vassilieva LL, Lynch M, et al. (2002) Spontaneous mutational variation for body size in *Caenorhabditis elegans*. Genetics 162: 755–765.
25. Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H (2010) Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. Proc Natl Acad Sci U S A 107: 7359–7364.
26. Wood WB (1988) The nematode *Caenorhabditis elegans*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory. 667 p.
27. Hall DH (1995) Electron microscopy and three-dimensional image reconstruction. In: Epstein HF, Shakes DC, eds. Methods in cell biol, *Caenorhabditis elegans* modern biological analysis of an organism. San Diego: Academic Press. pp 395–436.
28. Weimer RM (2006) Preservation of *C. elegans* tissue via high-pressure freezing and freeze-substitution for ultrastructural analysis and immunocytochemistry. Methods Mol Biol 351: 203–221.
29. Wang D, Urisman A, Liu YT, Springer M, Ksiazek TG, et al. (2003) Viral discovery and sequence recovery using DNA microarrays. PLoS Biol 1: E2. doi:10.1371/journal.pbio.0000002.
30. Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, et al. (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. Mol Cell 31: 79–90.
31. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, et al. (2002) The human genome browser at UCSC. Genome Res 12: 996–1006.
32. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, et al. (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421: 231–237.
33. Félix M-A (2007) Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. Curr Biol 17: 103–114.
34. Ahringer J, ed (2006) Reverse genetics. Wormbook, ed The *C elegans* research community. Available http://www.wormbook.org. Accessed 29 December 2010.
35. Lehrbach NJ, Armisen J, Lightfoot HL, Murfitt KJ, Bugaut A, et al. (2009) LIN-28 and the poly(U) polymerase PUP-2 regulate *let-7* microRNA processing in *Caenorhabditis elegans*. Nat Struct Mol Biol 16: 1016–1020.
36. Motohashi T, Tabara H, Kohara Y (2006) Protocols for large scale in situ hybridization on *C. elegans* larvae. WormBook. pp 1–8. Available: http://wormbook.org. Accessed 29 December 2010.
37. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, et al. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391: 806–811.
38. Milloz J, Duveau F, Nuez I, Félix M-A (2008) Intraspecific evolution of the intercellular signaling network underlying a robust developmental system. Genes Dev 22: 3064–3075.

Supporting Information

Figure S1

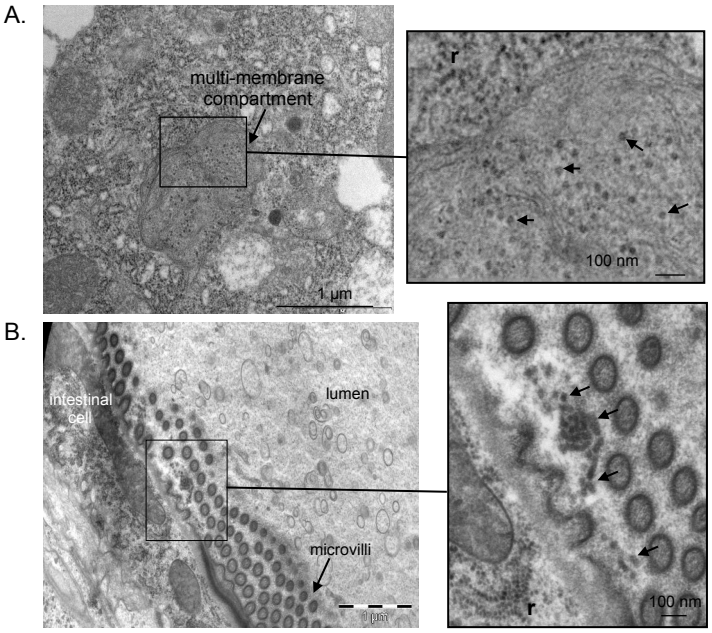

A.
multi-membrane
compartment
1 µm
r
100 nm
B.
Intestinal
cell
lumen
microvilli
1 µm
r
100 nm

Figure S2

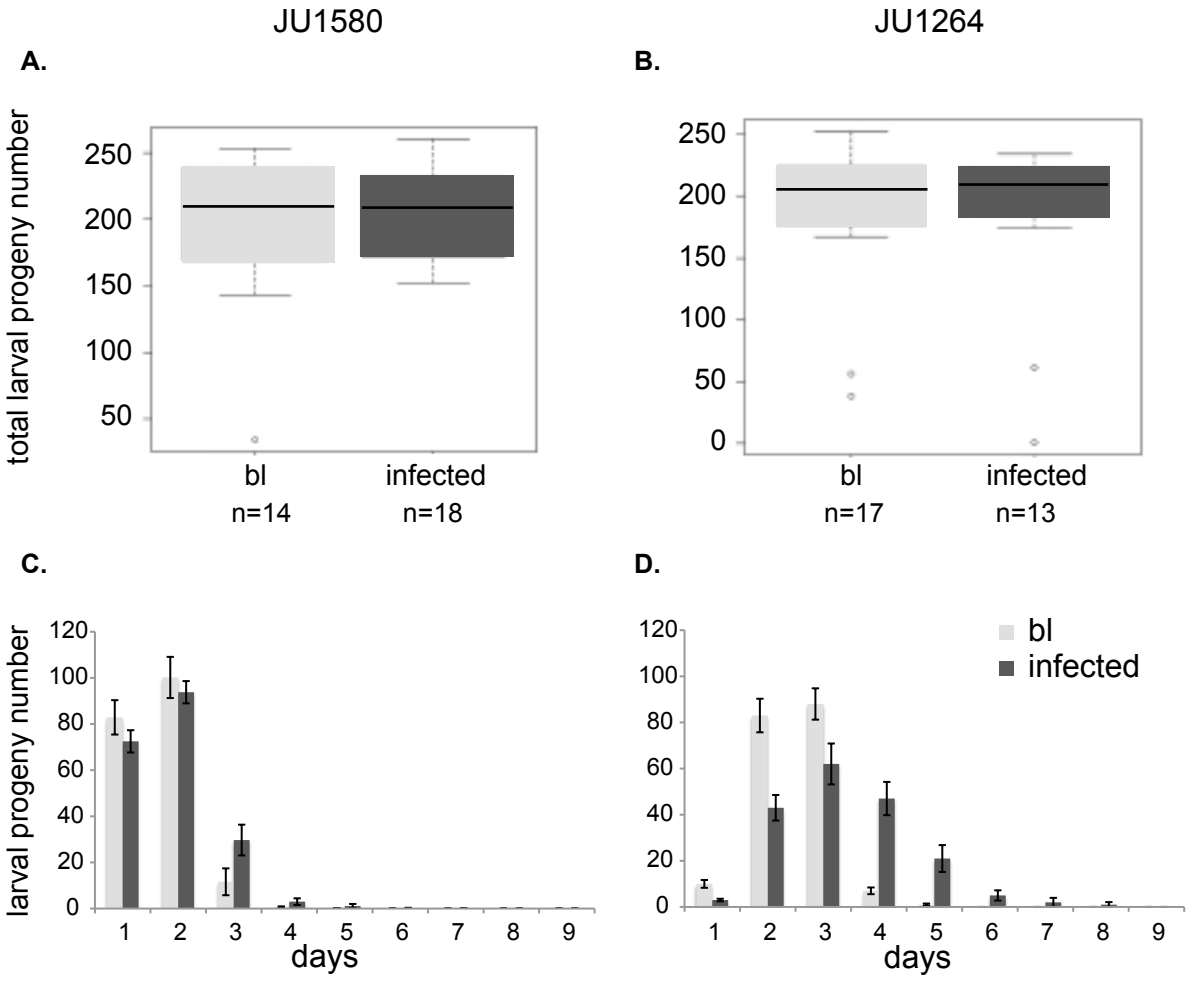

JU1580
JU1264
A.
B.
total larval progeny number
250
200
150
100
50
0
bl
infected
n=14
n=18
n=17
n=13
C.
D.
larval progeny number
120
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
days
bl
infected

Figure S3

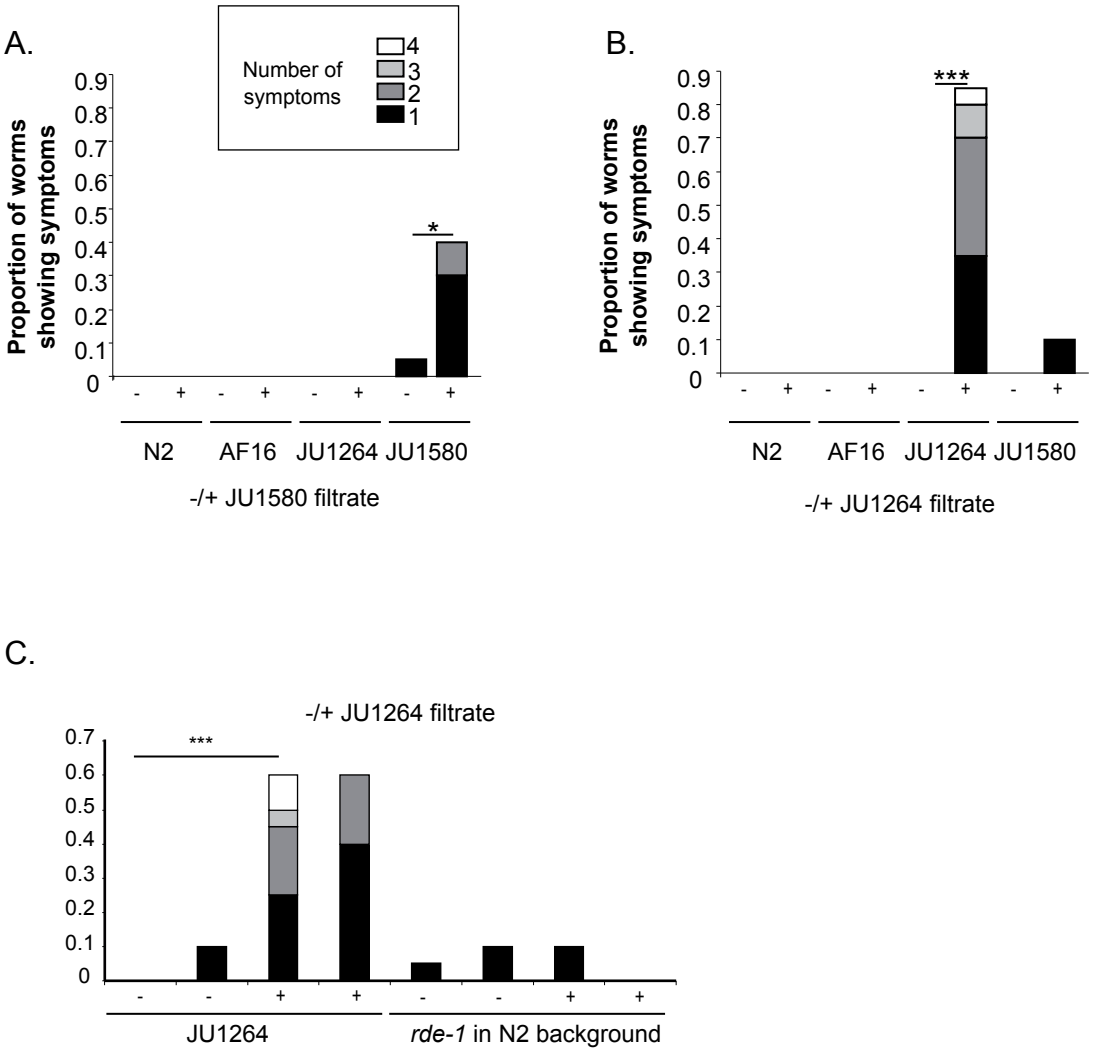
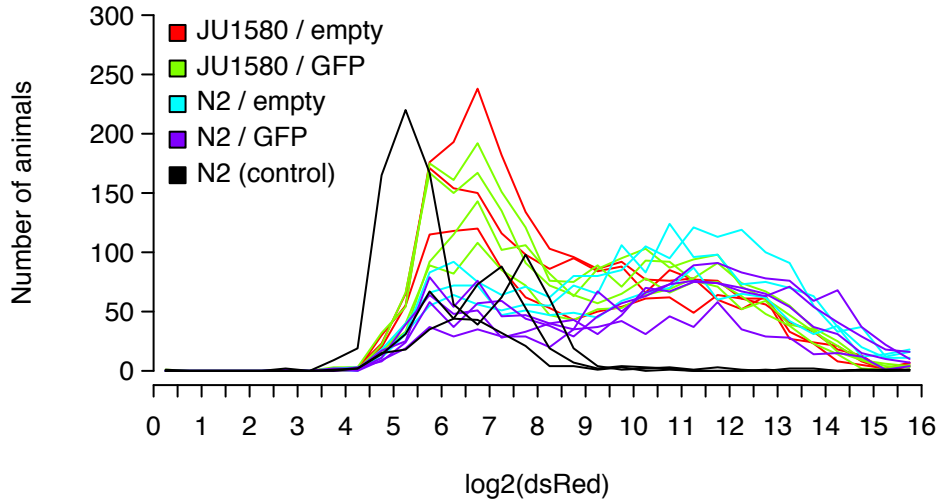

A.
Number of symptoms
4
3
2
1
Proportion of worms showing symptoms
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
*
- + - + - + - +
N2 AF16 JU1264 JU1580
-/+ JU1580 filtrate
B.
***
Proportion of worms showing symptoms
- + - + - + - +
N2 AF16 JU1264 JU1580
-/+ JU1264 filtrate
C.
-/+ JU1264 filtrate
***
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
- - + + - - + +
JU1264
rde-1 in N2 background

Figure S4


JU1580 / empty
JU1580 / GFP
N2 / empty
N2 / GFP
N2 (control)
Number of animals
log2(dsRed)

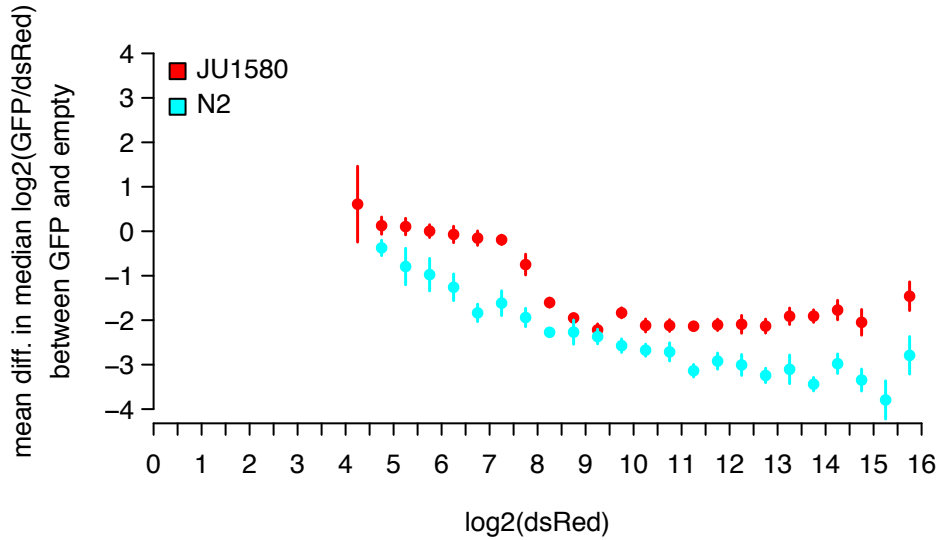
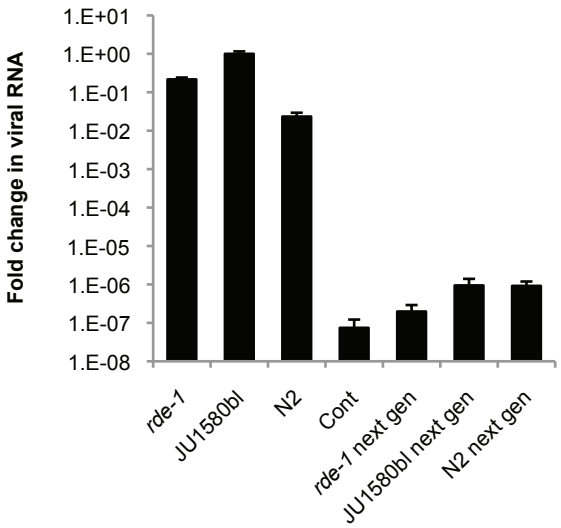

JU1580
N2
mean diff. in median log2(GFP/dsRed) between GFP and empty
log2(dsRed)

Figure S5


Fold change in viral RNA
1.E+01
1.E+00
1.E-01
1.E-02
1.E-03
1.E-04
1.E-05
1.E-06
1.E-07
1.E-08
rde-1
JU1580bl
N2
Cont
rde-1 next gen
JU1580bl next gen
N2 next gen

II. Une délétion naturelle de l'homologue de RIG-I abolit le clivage de l'ARN viral et l'immunité antivirale chez *Caenorhabditis elegans*

L'étude précédente suggère une variation de sensibilité au virus d'Orsay entre différentes souches naturelles de *C. elegans*. Le but de cette seconde étude est de déterminer les bases génétiques de cette variabilité. Nous avons pour cela testé la capacité du virus d'Orsay à se répliquer dans 97 isolats naturels de *C. elegans* représentant la diversité mondiale, partiellement séquencés par séquençage de l'ADN associé au site de restriction EcoRI. La variation observée entre les différentes souches est importante. Une analyse d'association pangénomique révèle un fort signal sur une région de 6 Mb au centre du chromosome IV. L'infection de F2 issus du croisement initial entre la souche standard de laboratoire N2 et JU1580 confirme cette localisation. Une cartographie réalisée à l'aide de recombinants entre N2 et JU1580 dans les 6 Mb d'intérêt a permis d'identifier une région plus restreinte de 155 kb contenant une délétion du gène *drh-1* chez JU1580.

Cette délétion de 159 pb abolit le fonctionnement de la protéine DRH-1 alors qu'une complémentation par un fosmide dans JU1580 restaure la résistance au virus. Cette protéine est un homologue des protéines RIG-I-like des mammifères, impliquées dans la reconnaissance des ARN viraux et dans le déclenchement d'une réponse interféron (absente chez *C. elegans*). De manière surprenante, cette délétion *a priori* délétère se retrouve à une fréquence intermédiaire dans la population (22/97 isolats). Une très faible diversité de SNP ainsi qu'un très faible taux de recombinaison autour du locus *drh-1* suggèrent un balayage sélectif récent dans la région. Cependant, il ne semble pas que cette délétion soit avantageuse pour son porteur en conditions de laboratoire. La délétion naturelle dans le gène *drh-1* pourrait donc avoir été emportée à cause d'un fort déséquilibre de liaison avec un gène positivement sélectionné dans cette région.

Cette délétion dans *drh-1* abolit presque totalement la production de petits ARN primaires à partir d'ARN viral, suggérant un rôle majeur de *drh-1* dans la production de petits ARN impliqué dans la réponse antivirale. Le profil de production de petits ARN spécifiques du virus d'Orsay dans des doubles mutants entre *drh-1* et de multiples acteurs de la voie des petits ARN place DRH-1 tout en amont de la voie. Ainsi, DRH-1 jouerait le même rôle de reconnaissance d'ARN viraux que chez les mammifères mais

déclencherait un mécanisme de production de petits ARN viraux similaire à celui des plantes et des arthropodes.

En bleu, ma participation à l'étude.
Figures réalisées : Figure 1A-B ; Figure 1-figure supplement 1A ; Figure 1-figure supplement 2A ; Figure 2 ; Figure 2-figure supplement 1 ; Figure 2-figure supplement 2.

eLIFE
elife.elifesciences.org

RESEARCH ARTICLE

A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity

Alyson Ashe[1,2†], Tony Bélicard[3,4†], Jérémie Le Pen[1,2†], Peter Sarkies[1,2†], Lise Frézal[3,4], Nicolas J Lehrbach[1,2‡], Marie-Anne Félix[3,4*], Eric A Miska[1,2*]

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, United Kingdom; [2]Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom; [3]Institute of Biology of Ecole Normale Supérieure (IBENS), Centre National de la Recherche Scientifique, UMR 8197, Paris, France; [4]Institut National de la Santé et de la Recherche Médicale U 1024, Paris, France

Abstract RNA interference defends against viral infection in plant and animal cells. The nematode *Caenorhabditis elegans* and its natural pathogen, the positive-strand RNA virus Orsay, have recently emerged as a new animal model of host-virus interaction. Using a genome-wide association study in *C. elegans* wild populations and quantitative trait locus mapping, we identify a 159 base-pair deletion in the conserved *drh-1* gene (encoding a RIG-I-like helicase) as a major determinant of viral sensitivity. We show that DRH-1 is required for the initiation of an antiviral RNAi pathway and the generation of virus-derived siRNAs (viRNAs). In mammals, RIG-I-domain containing proteins trigger an interferon-based innate immunity pathway in response to RNA virus infection. Our work in *C. elegans* demonstrates that the RIG-I domain has an ancient role in viral recognition. We propose that RIG-I acts as modular viral recognition factor that couples viral recognition to different effector pathways including RNAi and interferon responses.
DOI: 10.7554/eLife.00994.001

***For correspondence:** felix@biologie.ens.fr (M-AF); eric.miska@gurdon.cam.ac.uk (EAM)

†These authors contributed equally to this work

‡Present address: Department of Molecular Biology, Harvard Medical School, Boston, United States

Competing interests: The authors declare that no competing interests exist.

Funding: See page 19

Received: 24 May 2013
Accepted: 20 August 2013
Published: 08 October 2013

Reviewing editor: Detlef Weigel, Max Planck Institute for Developmental Biology, Germany

Copyright Ashe et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

The arms races between pathogens and their hosts have led to the evolution of sophisticated mechanisms to provide immunity against infection. Whilst adaptive immunity is specific to vertebrates, innate mechanisms are present in all multicellular organisms, allowing cells to recognize specific pathogens and instigate appropriate responses. RNA viruses are important pathogens of many multicellular organisms, which replicate without a DNA intermediate using RNA dependent RNA polymerase. Successful neutralization of invading RNA viruses by cells thus requires the viral genome to be recognized within the sea of endogenous RNA.

The primary innate immune sensors for RNA viruses in mammals are RIG-I and its homolog MDA-5 (***Schlee, 2013***). Viral recognition by RIG-I and MDA-5 triggers activation of downstream signaling, mediated by the proteins' N-terminal CARD domains, and results in the activation of the interferon pathway (***Yoneyama et al., 2004***). Initial recognition of viral RNA is likely to be mediated by the DExD/H-box helicase domain and the C-terminal RIG-I domain. Though the precise ligands that activate RIG-I family proteins are not fully defined, MDA-5 appears to bind long dsRNA (***Gitlin et al., 2006***), whilst RIG-I seems to recognize the 5′ end of double stranded RNA, but only if it has a 5′ triphosphate (***Pichlmair et al., 2006***). As all known RNA polymerases leave a triphosphate at the 5′ end of newly synthesized RNA, the presence of a 5′ triphosphate is likely to be a signature of RNA virus replication

eLife digest Most organisms—from bacteria to mammals—have at least a rudimentary immune system that can detect and defend against pathogens, particularly viruses. This defense mechanism, which is known as the innate immune system, uses sensor proteins to recognize viral RNA, and then mobilizes other immune components to attack the invaders.

The specific mechanisms used to destroy viruses differ between species. In mammals, a protein called RIG-1 binds to viral RNA and activates a signaling pathway that leads to the production of interferons: immune proteins named after their ability to 'interfere' with viral replication. Plants and insects do not use interferons, but instead use a mechanism called RNA interference, in which long double-stranded RNAs are cleaved into shorter fragments.

The nematode worm *C. elegans* also deploys RNA interference against viruses but, in contrast to insects and plants, worms do not possess a specific set of RNA interference enzymes that participate solely in the antiviral response. They do, however, express a protein called DRH-1 that is related to the RIG-I protein found in mammals.

To investigate whether DRH-1 contributes to innate immunity in *C. elegans*, Ashe et al. infected 97 strains of *C. elegans* from around the world with a virus, and showed that some strains were more sensitive to the virus than others, with certain strains showing complete resistance. By comparing a sensitive strain with a resistant one, Ashe et al. revealed that viral sensitivity was caused by a mutation in the gene encoding DRH-1.

Further experiments showed that DRH-1 is required for the first step in RNA interference. Ashe et al. have thus identified a conserved role for RIG-1 in initiating antiviral responses, and propose that the protein couples virus recognition to distinct defense mechanisms in different evolutionary groups.

DOI: 10.7554/eLife.00994.002

and thus allow viral replication intermediates to be distinguished from endogenous mRNA, which will predominantly display a 5′ cap (***Rehwinkel and Reis e Sousa, 2010***).

In plants and insects, interferon signaling is not involved in antiviral defense. Instead, the RNA interference (RNAi) pathway provides robust defense against RNA viral infection (***Ding and Voinnet, 2007***; ***Ding, 2010***). The initial step in protection against positive strand RNA virus infection in plants and insects is detection and subsequent cleavage of the double-stranded replication intermediate by members of the Dicer family of endonucleases. Insects and plants possess dedicated Dicer enzymes responsible specifically for this antiviral response, Dicer-like 4 (and to a lesser extent Dicer-like 1) in plants, and Dicer2 in insects (***Bouché et al., 2006***; ***Deleris et al., 2006***; ***Fusaro et al., 2006***; ***Galiana-Arnoux et al., 2006***; ***van Rij et al., 2006***; ***Diaz-Pendon et al., 2007***). The small RNAs thus generated feed into the canonical RNAi machinery and can be used to silence the viral genome.

In the nematode *C. elegans*, the RNAi pathway is best characterized as a response to artificial introduction of dsRNA by feeding or injection. However, exposing *C. elegans* to virus-derived dsRNA using transgenes or infection with mammalian viruses also triggers an RNAi response (***Lu et al., 2005***; ***Schott et al., 2005***; ***Wilkins et al., 2005***; ***Yigit et al., 2006***). Additionally, we have shown previously that this response is also initiated upon infection with a positive strand RNA virus, named the Orsay virus, which infects *C. elegans* in the wild through horizontal transmission. Disruption of this pathway through mutation of the core components of the RNAi machinery results in greatly increased viral infection levels. Interestingly, the Orsay virus infects a wild isolate named JU1580 (from which the Orsay virus was isolated) to much higher levels than the N2 strain (***Félix et al., 2011***).

Despite these advances, our understanding of how RNAi-mediated antiviral defense in *C. elegans* is orchestrated is still limited. One major unsolved problem is how viral dsRNA within the cell is recognized in order to initiate the antiviral response. In contrast to the situation in both plants and insects, *C. elegans* only has one Dicer enzyme (***Knight and Bass, 2001***; ***Yigit et al., 2006***); it is expected therefore that the activity of Dicer in different small RNA pathways will be controlled by different partner proteins within distinct complexes (***Yigit et al., 2006***; ***Thivierge et al., 2012***).

In this regard it is intriguing that *C. elegans* encodes three homologues of the mammalian viral recognition protein RIG-I: DRH-1, DRH-2 and DRH-3. The three RIG-I family members do not contain the CARD domains required for interferon induction; consistently they have yet to be implicated in

signaling pathways. However, they do possess the RIG-I C-terminal domain and the helicase domains, potentially enabling them to recognize viral RNA. So far study of these genes has connected them to the RNA interference response. DRH-1 was initially characterized as a protein interacting with DCR-1, RDE-1 and RDE-4 (**Tabara et al., 2002**; **Duchaine et al., 2006**; **Sijen et al., 2007**). RNAi of *drh-1* was not found to be required for RNAi per se but result in a defect in RNAi of a second gene (**Tabara et al., 2002**), however it remains unclear whether this effect was solely due to knockdown of *drh-1*. More recent studies did not observe defects in endogenous or exogenous small RNA pathways in *drh-1* mutants (**Gu et al., 2009**; **Lu et al., 2009**). However, *drh-1* mutants were found to be defective in the silencing of a flockhouse virus-derived replicon (**Lu et al., 2009**). *drh-1* and *drh-2* are the result of a recent gene duplication event that occurred after the last common ancestor of *C. elegans* and its closest known sister clade including *C. briggsae* (www.wormbase.org) (**Stein et al., 2003**), and *drh-2* is the upstream gene in an operon containing *drh-1*. *drh-2* has lost some of its functional domains due to frame-shift mutations and its function remains unclear, although it has been suggested to act as a negative regulator of RNAi (**Lu et al., 2009**). DRH-3 is required for endogenous small RNA pathways in the germline and efficient exogenous RNAi (**Gu et al., 2009**), and is found in at least two distinct protein complexes including the ERI Complex (ERIC) (**Gu et al., 2009**; **Thivierge et al., 2012**). Thus, an intriguing possibility is that RIG-I family genes in *C. elegans* initiate RNAi rather than the interferon response.

Here we discover a naturally occurring deletion in the gene *drh-1* that is widespread in *C. elegans* populations despite predisposing individuals to viral sensitivity. We show that the increased viral sensitivity caused by this deletion results from a failure of the RNAi pathway. Mutations in *drh-1* almost completely abolish the production of primary siRNAs, allowing us to place DRH-1 at the top of a hierarchical RNAi response. Our data supports a model whereby the conserved RNA virus recognition capability of *drh-1* allows it to recruit the RNAi machinery to defend *C. elegans* from viral infection.

Results

A natural *drh-1* deletion is a major determinant of viral sensitivity in *C. elegans*

To explore intraspecific variation in viral resistance in *C. elegans*, we assayed a worldwide set of 97 wild *C. elegans* isolates that had previously been genotyped (**Andersen et al., 2012**). To assess viral sensitivity, we infected each isolate in triplicate and quantified the viral load after 7 days by qRT-PCR. Viral loads of the 97 isolates varied widely over five orders of magnitude (**Figure 1—figure supplement 1A**). Genome-wide association for viral load revealed a single peak covering a 6 Mb region in the middle of chromosome IV (**Figure 1A**). However, further mapping resolution is limited by the low natural recombination frequency in the species (**Cutter et al., 2009**; **Andersen et al., 2012**).

We therefore focused on genetic variation in viral sensitivity between the N2 and JU1580 isolates. We first assayed viral sensitivity in 110 F2 recombinant families (**Figure 1—figure supplement 1B**). Whole-genome sequencing of a pool of the 20 most sensitive families revealed linkage to chromosome IV (**Figure 1—figure supplement 1C**). To narrow down the candidate region, we introgressed the center of chromosome IV (IV:3,329,219 to IV:11,083,410) of JU1580 into N2 animals (yielding strain JU2170). As expected, JU2170 showed similar viral load to JU1580 (**Figure 1—figure supplement 2**). We then screened for recombinants in this region after crossing JU2170 with N2. This allowed us to restrict the candidate region to a 155 kb interval carried by a virus-sensitive recombinant strain named JU2196 (**Figure 1B**, **Figure 1—figure supplement 2**).

Genome sequencing of the JU1580 isolate and alignment to the N2 reference strain revealed 20 single nucleotide polymorphisms (SNPs), including one non-synonymous SNP, and one 159 base indel in the 155 kb region (**Figure 1—figure supplement 2**). The 159 base deletion in the JU1580 genome, named *niDf250* (IV:6,607,635-6,607,793), lies within the *drh-1* gene locus (**Figure 1C**). *drh-1* is a homolog of the mammalian *RIG-I* family genes, whose products bind virus-derived RNAs, acting as pattern recognition receptors (**Parameswaran et al., 2010**; **Yoneyama and Fujita, 2010**), and trigger antiviral innate immune responses in mammals (**Yoneyama et al., 2004**). *C. elegans* expresses three RIG-I-like proteins: DRH-1, DRH-2 and DRH-3. The *niDf250* deletion covers most of *drh-1* exon 19 and part of exon 20 (**Figure 1C**) and was confirmed using genomic PCR (**Figure 1D**). *drh-1* mRNA levels are unchanged in JU1580 (**Figure 1E**) and the resulting transcript is predicted to encode a truncated protein of 987 amino acids, identical to the N2 form of DRH-1 up to amino acid 973 but with

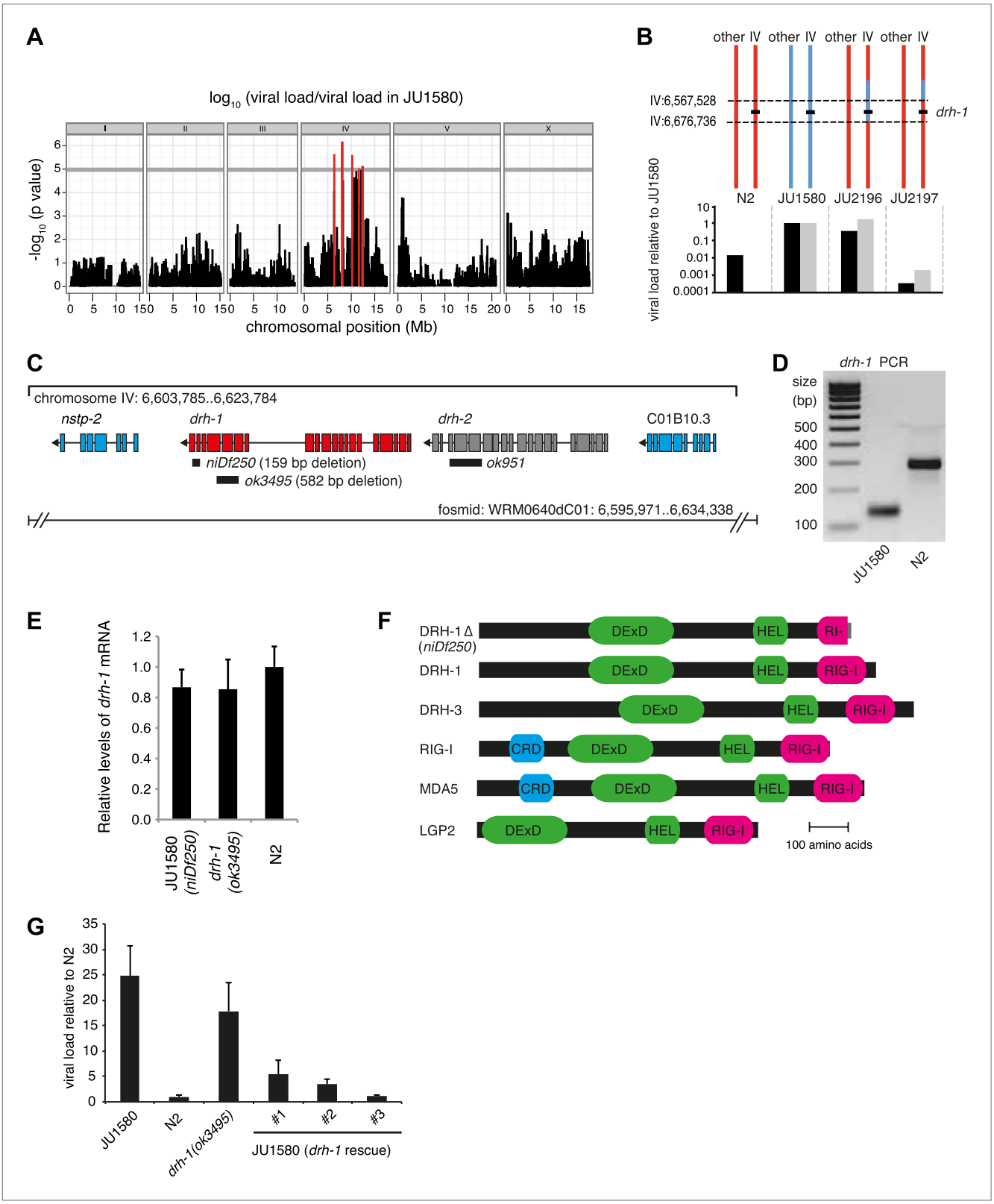


Figure 1. A deletion polymorphism in *drh-1* is a major determinant of Orsay virus sensitivity in wild isolates of *C. elegans*. (**A**) Genome-wide association analysis of Orsay virus sensitivity in 97 wild isolates of *C. elegans*. The mapped trait is the viral load of animals, measured by qRT-PCR on the Orsay virus RNA2 genome after 7 days of infection, using three independent infection experiments. The horizontal grey line is a Bonferroni-corrected threshold of

Figure 1. Continued on next page

Figure 1. Continued

significance at p=0.05. Peaks reaching above this threshold are colored in red. (**B**) Fine mapping of the candidate region causative for virus hypersensitivity observed in JU1580 animals. The genotypes of chromosome IV and other chromosomes are represented for parental (N2 and JU1580) and informative recombinant (JU2196 and JU2197) strains. Regions in red or blue are identical to N2 or JU1580, respectively. The inferred candidate region is delimited by dotted lines. Below each genotype are viral load measured by qRT-PCR of Orsay virus RNA2, in two independent infections (black and grey bars) and normalized to JU1580. (**C**) Diagram of the *drh-1* locus. Positions of deletion alleles and a rescuing fosmid are indicated. (**D**) PCR analysis of *niDf250* deletion in N2 and JU1580 strains. (**E**) *drh-1* mRNA level in different strains (as indicated), measured by RT-qPCR. (**F**) Diagram of *C. elegans* and human RIG-I like genes. DeXD = Pfam:DEAD, Hel = Pfam:Helicase_C, RIG-I = Pfam:RIG-I_C-RD, CRD = Pfam:CARD. (**G**) Viral load in different strains (as indicated), measured by RT-qPCR of the Orsay virus RNA1 genome after 4 days of infection. JU1580 (*drh-1* rescue) refers to JU1580 strains carrying three independent transgenic lines (SX2375, SX2376, SX2377). Transgenes include the fosmid WRM0640dC01 and a co-injection marker, they were integrated into the genome using X-rays. Error bars represent the standard error of the mean (SEM) of five biological replicates.

DOI: 10.7554/eLife.00994.003

The following figure supplements are available for figure 1:

Figure supplement 1. Variation in the ability of the Orsay virus to replicate in *C. elegans*.

DOI: 10.7554/eLife.00994.004

Figure supplement 2. Genotype and sensitivity to the Orsay virus of recombinants in the chromosome IV region.

DOI: 10.7554/eLife.00994.005

a novel C-terminus (***Figure 1F***). Importantly, this truncates the RIG-I C-terminal domain (amino acids 885–1014), thought to be required for RNA recognition specificity (***Kowalinski et al., 2011***; ***Figure 1F***).

To establish a potential role for DRH-1 in Orsay antiviral resistance we tested whether a *drh-1* deletion in the N2 background was sufficient to impart viral sensitivity, as assayed by viral load. Indeed, the *drh-1(ok3495)* mutant displayed an increased viral load compared to N2 animals, similar to that of JU1580 (***Figure 1G***). Conversely, transgenic JU1580 animals carrying a fosmid containing the N2 allele of *drh-1* were resistant to Orsay infection (***Figure 1G***). Therefore, variation at the *drh-1* locus explains the difference in viral load between N2 and JU1580.

An inactivating mutation in a pathogen-resistance gene could be expected to be a rare deleterious variant in natural populations, yet the wild *niDf250* allele is found at an intermediate frequency at the global level, in 22/97 (23%) of the tested wild isolates. The deletion is found in about one third of isolates from Europe and Africa (21/64, 33%) with a high incidence in France (14/30, 47%), but is rarer (1/30) in those from the Americas and the Pacific regions (***Figure 2A***). As expected, the presence of the deletion correlated strongly with viral load in the infection experiment (***Figure 1—figure supplement 1A***; Wilcoxon test on viral load of isolates carrying each *drh-1* allele, $p=1.3\times10^{-9}$). Thus, surprisingly, the derived *drh-1* allele has spread to intermediate frequency in natural populations, despite rendering the animals susceptible to viral infection.

A possible interpretation for the spread of the sensitive *drh-1* allele might be that high Orsay viral load has no deleterious effect on fitness. However, in laboratory conditions, we found that viral infection leads to delayed and decreased total progeny of JU1580 and *drh-1(ok3495)* mutant animals relative to uninfected animals, whilst infection of N2 had no significant effect (***Figure 2—figure supplement 1***). Furthermore, we performed a competition experiment between N2 and JU2196, which contains the *drh-1* region from JU1580 introgressed into the N2 background (***Figure 1B***). N2 rapidly outcompetes JU2196 in the presence of viral infection, but not its absence, confirming that the increased viral infection resulting from the *drh-1* deletion is indeed detrimental for fitness (***Figure 2B***). In the absence of viral infection, we could not detect in standard laboratory conditions over 10 generations of competition any positive or negative effect of the *drh-1* deletion and the introgressed surrounding region. Thus the natural *drh-1* deletion impairs fitness only in the presence of viral infection.

To characterize the evolutionary history of the *drh-1* region, we first focused on the 6 Mb region detected by genome-wide association. This region presents three main haplotypes among the 97 wild isolates (***Andersen et al., 2012***): the N2 haplotype, the JU1580 haplotype and a distant haplotypic group (including RW7000 and QX1211), as well as a few recombinants (***Figure 2C***, ***Figure 2—figure supplement 2***). The *drh-1(niDf250)* allele is exclusively found in isolates carrying the JU1580 haplotype in the 6 Mb region and in a few recombinants with either the N2 or distant haplotypes (***Figure 2C***). The JU1580 and N2 haplotypes show fewer fixed differences in RAD polymorphic sites between them (24 SNPs) than with the divergent haplotype group (63 and 57 SNPs, respectively). In addition, from our whole-genome JU1580 sequencing data, N2 and JU1580 display a very low level of molecular diversity

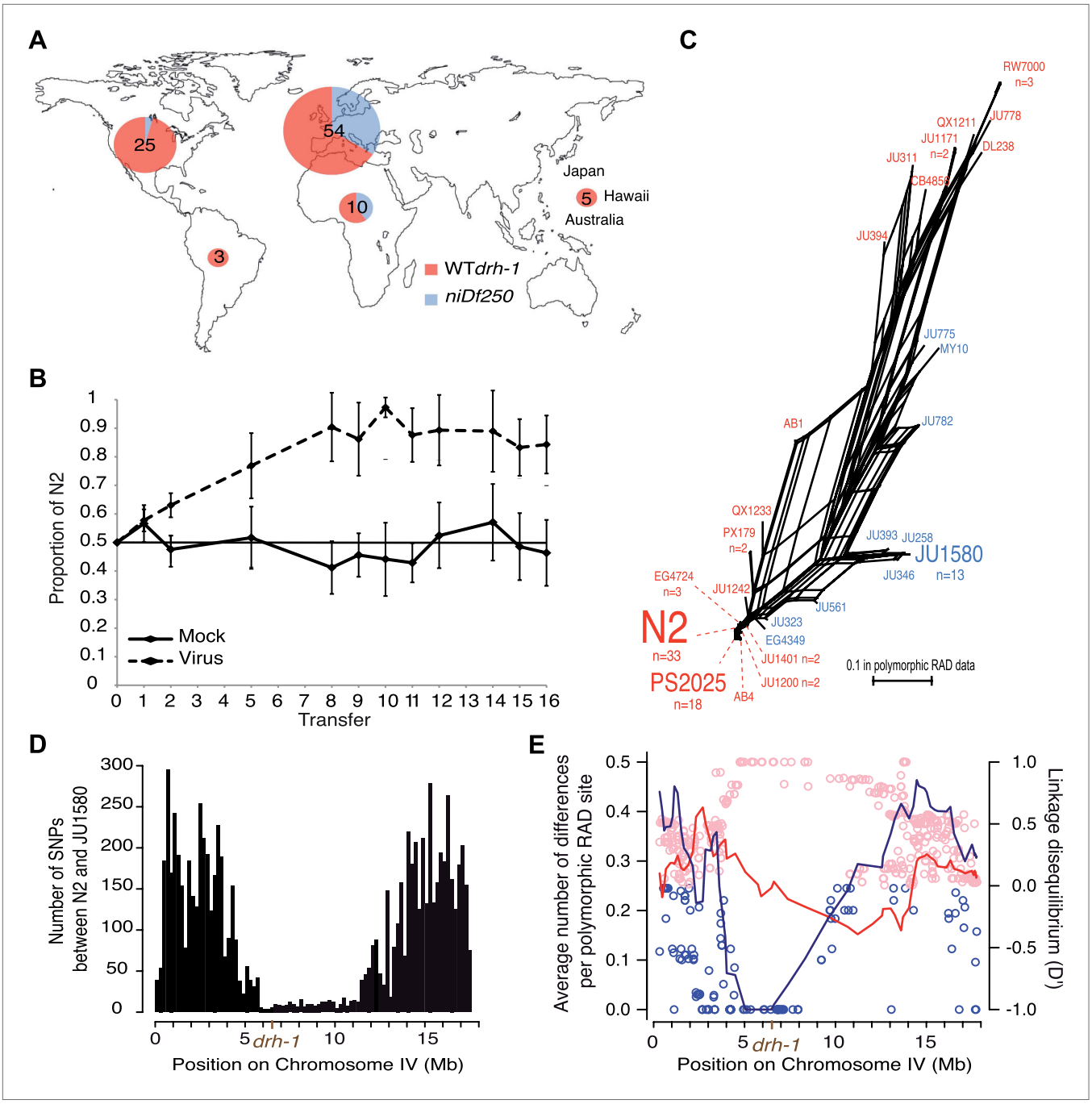


Figure 2. Geographic distribution and evolutionary genetic context of *drh-1* alleles. (**A**) Geographic distribution of *drh-1* alleles. The respective frequencies of the *niDf250* and N2 alleles of *drh-1* are represented for each world region in blue and red, respectively, based on genotyping of the 97 wild isolates. (**B**) Competition experiment between the N2 reference and the JU2196 introgression line. In the absence of the Orsay virus, the proportion of the N2 genotype remains close to 50% throughout the experiment (48.6 ± 5.3%). In the presence of the Orsay virus, the proportion of the N2 genotype increases and appears to stabilize around 90% after nine transfers (88.1 ± 4.6%). The presence of the virus has a significant effect (linear model, $p=1.6 \times 10^{-4}$). Error bars represents standard deviation. (**C**) Neighbor-network of the 97 isolates in the chromosome IV central region associated with Orsay virus sensitivity. Only one isolate per haplotype is represented; font size is relative to the number (n) of isolates sharing this haplotype. Haplotypes in blue or red carry the *niDf250* allele of *drh-1*, respectively. (**D**) Distribution of SNPs along chromosome IV between N2 and JU1580, based on JU1580 whole-genome sequencing. (**E**) Molecular diversity (left y axis scale) is plotted along chromosome IV for isolates carrying the *niDf250* or the N2 allele as blue or red lines, respectively. Linkage Disequilibrium *D'* values (right y axis scale) between polymorphic RAD sites along chromosome IV and the *niDf250* or N2 alleles of *drh-1* are represented with blue or red circles, respectively.

DOI: 10.7554/eLife.00994.006

Figure 2. Continued on next page

Figure 2. Continued

The following figure supplements are available for figure 2:

Figure supplement 1. Infection by the Orsay virus has an effect on progeny production of *drh-1* deleted strains and on longevity of the *drh-1(ok3495)* mutant.

DOI: 10.7554/eLife.00994.007

Figure supplement 2. Chromosome IV haplotypes for the 97 isolates (modified from Supplemental Figure 7 in ***Andersen et al., 2012***).

DOI: 10.7554/eLife.00994.008

in the central region of chromosome IV (***Figure 2D***). Moreover, we observe a strong decrease in molecular diversity between IV:4529464 to IV:6662701 in isolates carrying the *niDf250* allele (***Figure 2E***, lines), but not in those carrying the *drh-1* (N2) allele. Thus, the divergence between the N2 and JU1580 haplotypes in this region, including the *niDf250* deletion, appears recent relative to much of the species' genetic diversity. Furthermore, the *niDf250* allele is in high or even full linkage disequilibrium with a large region of chromosome IV (***Figure 2E***, dots). This lack of diversity and the high linkage disequilibrium around *niDf250* suggest a partial sweep of the haplotype linked to *niDf250*. They also imply that the sensitive *drh-1* allele may have spread by hitch-hiking with a favorable allele.

DRH-1 initiates an antiviral small RNA response

Having established the major role of *drh-1* allelic variation in natural variation of antiviral defense in *C. elegans*, we wished to understand the molecular mechanisms of DRH-1 action in the antiviral response. We previously showed that disruption of small RNA pathway genes such as *rde-1*, which encodes an Argonaute protein essential for RNAi in response to exogenous dsRNA (RDE-1), renders N2 animals as sensitive to the Orsay virus as JU1580 (***Félix et al., 2011***). DRH-1 interacts with the double-stranded RNA (dsRNA) binding protein RDE-4 and the dsRNA-specific endonuclease Dicer (DCR-1), both of which act upstream of RDE-1 in exogenous RNAi (***Tabara et al., 1999***; ***Parrish and Fire, 2001***; ***Barber et al., 2010***) However, DRH-1 is dispensable for exogenous RNAi (***Gu et al., 2009***). We therefore wondered if DRH-1 could act specifically to promote DCR-1 processing of long dsRNA that is produced during viral RNA replication. Indeed, we find that DCR-1 is required for viral resistance in the N2 strain (***Figure 3A***). We therefore postulated that the antiviral response involves siRNAs processed by DCR-1, which may initiate a cascade of events analogous to canonical *C. elegans* RNAi.

In *C. elegans*, the RNAi pathway is divided into primary and secondary steps. Long dsRNA is processed by DCR-1 to generate a primary siRNA duplex ~23 nucleotides (nt) in length with 5′ monophosphates and 2 nt 3′ overhangs. By examining libraries derived solely from small RNAs with 5′ monophosphates (5′ dependent libraries), we could interrogate the primary siRNA response specifically. In wild-type N2 animals, infection with virus leads to generation of predominantly 23 nt small RNAs with no first nucleotide bias, mapping both sense and antisense to the viral genome in equal proportions (***Figure 3B***). These small RNAs are likely primary DCR-1 products generated from the double-stranded intermediate of viral replication, consistent with the small RNA response against single-strand RNA viruses in insects (***Flynt et al., 2009***). Furthermore, overlapping RNAs mapping to complementary strands of the viral genome showed the 2 nt 3′ overhang characteristic of Dicer products ($p<10^{-9}$, χ^2 test against a uniform distribution of overhang length) (***Figure 3C***). The same viRNA pattern was observed in *rde-1* mutants lacking the primary siRNA Argonaute protein RDE-1 (***Figure 3D***). Thus *rde-1* mutants are proficient in primary viRNA formation despite being sensitive to viral infection (***Figure 3—figure supplement 1***) (***Félix et al., 2011***). In contrast, JU1580 animals showed a markedly different pattern (***Figure 3E***). First, a much higher proportion of small RNAs derived from the sense strand than the antisense strand (88% vs 49% in N2). Second, the size distribution of viRNAs was flattened, with a greatly reduced proportion of 23 nt RNAs and an increased proportion of 15–20 nt RNAs. These shorter viRNAs were not present in libraries prepared from uninfected JU1580 or N2 controls (***Figure 3—figure supplement 2A***) and appeared to be derived from regions of strong secondary structure within the virus (***Figure 3—figure supplement 3***). The *drh-1* mutant (in an N2 strain background) displayed an identical viRNA pattern to JU1580 (***Figure 3F***). Thus JU1580 and *drh-1* mutant strains are deficient in primary siRNA generation against the virus.

Primary siRNAs act upstream of an amplification step by triggering the synthesis of secondary siRNAs antisense to targeted RNAs (***Sijen et al., 2001***). Secondary siRNAs are synthesized by RNA-dependent RNA polymerases (RdRPs), and bind to a number of secondary siRNA-specific Argonaute

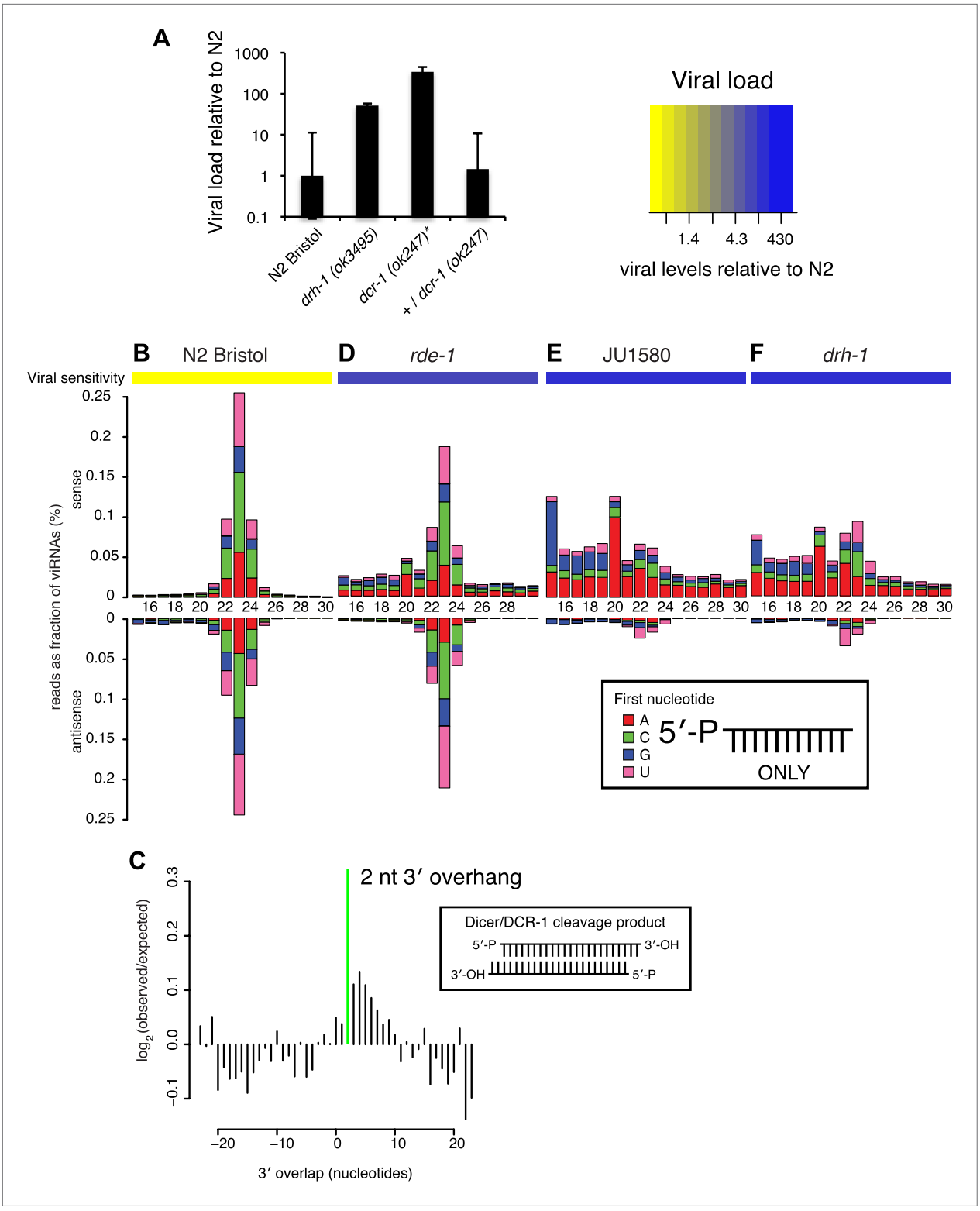


Figure 3. DRH-1 is required for the Orsay antiviral response and primary viRNA generation. (**A**) qRT-PCR analysis of viral load after 4 days of infection with the Orsay virus. *, *dcr-1* mutants are sterile, data shown are homozygous mutant animals from heterozygous mothers. (**B**) Primary viRNA populations in strains as indicated. 5′ dependent small RNA sequencing captures only primary siRNAs with a 5′ monophosphate. Data are grouped as sense or antisense and according to length and the identity of the first nucleotide. From the same samples viral load was measured by qRT-PCR of the Orsay virus RNA1 genome after four days of infection (heatmap, see also ***Figure 3A*** and ***Figure 3—figure supplement 1B***). (**C**) Analysis of phasing of 23 nt primary

Figure 3. Continued on next page

Figure 3. Continued

viRNAs generated in infected N2 animals. The x axis shows the length of the overhang in nucleotides, either 5′ (negative numbers) or 3′ (positive numbers), for each pair of sequences that map to overlapping regions on opposite strands. A value of 0 represents a pair of viRNAs with perfect complementarity that would form blunt ends. The y axis shows the number of times each particular overhang was observed relative to the number of times that such an overhang would be expected if overhangs were random. Green bar indicates the 2 nt 3′ overhang. (**D**–**F**) same as in (**B**).

DOI: 10.7554/eLife.00994.009

The following figure supplements are available for figure 3:

Figure supplement 1. Viral sensitivity in a number of mutants of small RNA pathway genes.

DOI: 10.7554/eLife.00994.010

Figure supplement 2. Additional small RNA sequencing controls.

DOI: 10.7554/eLife.00994.011

Figure supplement 3. Relationship between antiviral small RNAs and predicted secondary structure within the viral genome.

DOI: 10.7554/eLife.00994.012

proteins to bring about target silencing (***Yigit et al., 2006***). Secondary siRNAs have a modal length of 22 nt, are 5′ triphosphorylated, have a strong preference for a 5′ guanine (G), and are referred to as 22G siRNAs (22Gs). Small RNAs can be enzymatically treated prior to adaptor ligation to allow sequencing (5′ independent) of both primary and secondary siRNAs (***Pak and Fire, 2007***; ***Sijen et al., 2007***). N2 animals infected with Orsay virus showed a robust secondary 22G siRNA response, primarily antisense to the viral genome (***Figure 4A***, ***Figure 4—figure supplement 1***). In contrast, *rde-1* mutants lacked 22G siRNAs, consistent with a role for RDE-1 in initiating the secondary siRNA response (***Figure 4B***, ***Figure 4—figure supplement 1***). In addition, mutants lacking DRH-3 or the RdRP RRF-1 and a strain deficient in 12 worm Argonaute proteins (WAGO-1 through 12) that bind secondary siRNAs (MAGO-12) (***Yigit et al., 2006***) lacked 22G viRNAs, but still produced primary siRNAs (***Figure 4F,L,M***). The above have previously been implicated in secondary siRNA generation in other contexts (***Yigit et al., 2006***; ***Pak and Fire, 2007***; ***Sijen et al., 2007***) and *drh-3*, *rrf-1* and MAGO-12 mutant strains are also sensitive to viral infection similar to *rde-1* (***Figure 3—figure supplement 1***). Taken together these data show that a canonical secondary siRNA pathway is engaged to amplify the antiviral response.

drh-1 mutants and JU1580 animals displayed the same profile of sense siRNAs in 5′ independent libraries as in 5′ dependent libraries (***Figures 3E–F and 4C,E***). Mutants deficient in the adjacent gene *drh-2* have a viRNA profile similar to N2 (***Figure 4I***) and are as resistant to viral infection as N2 (***Figure 3—figure supplement 1***), implying that this gene is not involved in the antiviral siRNA pathway. Moreover transgenic JU1580 animals carrying the *drh-1* gene from N2 showed the same overall viRNA profile as N2 (***Figure 4D***, ***Figure 4—figure supplement 1***), further supporting the conclusion that *drh-1* deficiency is primarily responsible for the defective siRNA synthesis and virus sensitivity in JU1580. Importantly however, there were residual antisense 22G siRNAs present in both JU1580 and *drh-1* mutants. This suggests that the few DCR-1 products with the correct length in *drh-1* mutants can still be used to generate secondary siRNAs as in N2, implying that *drh-1* is not essential for the secondary siRNA pathway. Consistent with this interpretation, the reduced 23 nucleotide long primary siRNA products in JU1580 still displayed a 2 nucleotide 3′ overhang characteristic of DCR-1 activity (***Figure 4—figure supplement 2***).

Furthermore, residual 22G siRNAs are also present in mutants deficient in the DCR-1 accessory factor RDE-4, which is required for efficient DCR-1 activity in the exogenous RNAi pathway (***Figure 4J***; ***Tabara et al., 1999***; ***Parrish and Fire, 2001***). Additionally, *drh-1* mutants displayed no difference in endogenous 22G siRNAs mapping antisense to protein-coding genes, whilst, in agreement with previous data (***Gu et al., 2009***), *drh-3* mutants showed markedly reduced levels of endogenous siRNAs (***Figure 4—figure supplement 3***). Together with our observation that *drh-1* mutants are deficient in primary viRNA production, this suggests that DRH-1 acts early in the antiviral siRNA pathway and is not required for downstream amplification steps. These observations are in contrast to earlier work describing a role for DRH-1 downstream of secondary siRNA production in a flockhouse virus replicon model (***Lu et al., 2009***).

To further test this interpretation, we examined double mutant strains. The prominent 23 nt peak for sense and antisense viRNAs attributed to DCR-1 activity present in *drh-3* and *rde-1* single mutants was absent in both *drh-3*; *drh-1* and *drh-1; rde-1* double mutants (***Figure 4G,H***). This is consistent with the idea that DRH-1 acts upstream of DRH-3 and RDE-1 in the antiviral siRNA pathway and acts in

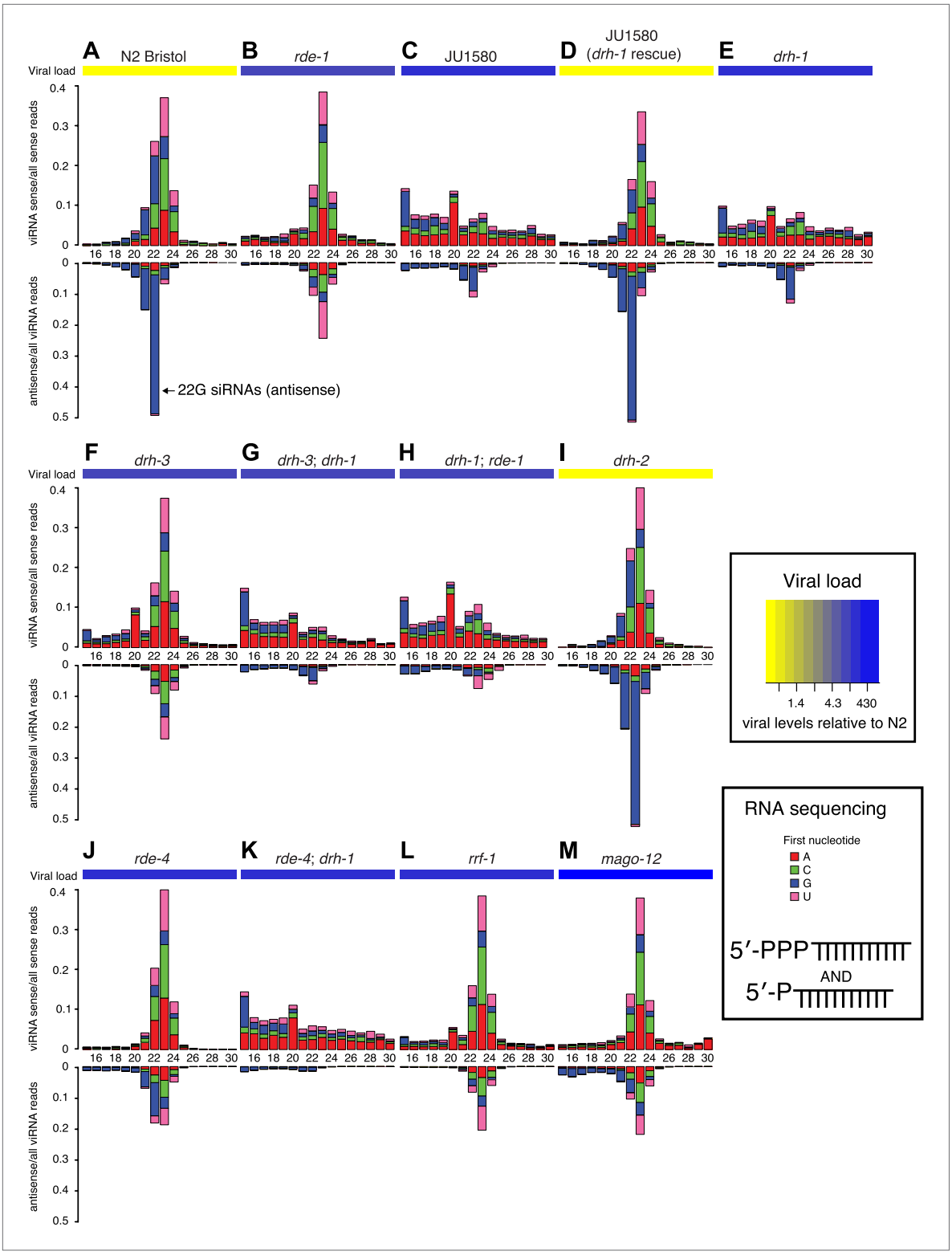


Figure 4. DRH-1 acts upstream of a 22G secondary siRNA pathway. (**A**–**M**) Primary and secondary viRNA populations in strains as indicated. 5′ independent small RNA sequencing captures 5′ primary siRNAs (5′ monophosphate) and secondary siRNAs (5′ triphosphate). Data are grouped as sense or antisense and according to length and the identity of the first nucleotide. From the same samples viral load was measured by RT-qPCR of the Orsay virus RNA1 genome after 4 days of infection (heatmap, see also ***Figure 3A*** and ***Figure 3—figure supplement 1***).

DOI: 10.7554/eLife.00994.013

Figure 4. Continued on next page

Figure 4. Continued

The following figure supplements are available for figure 4:

Figure supplement 1. Distribution of viRNAs along the Orsay genome.
DOI: 10.7554/eLife.00994.014

Figure supplement 2. Analysis of residual Dicer products in JU1580 mutants.
DOI: 10.7554/eLife.00994.015

Figure supplement 3. Analysis of 22G-RNAs mapping to endogenous loci.
DOI: 10.7554/eLife.00994.016

Figure supplement 4. *drh-1* mutants are not hypersensitive to RNAi.
DOI: 10.7554/eLife.00994.017

concert with DCR-1. The *rde-4; drh-1* double mutant showed a further reduction in primary 23 nucleotide long Dicer products compared to the *drh-1* single mutant (***Figure 3—figure supplement 2***). Furthermore, residual antisense 22G siRNAs in *drh-1* mutants (***Figure 4E***) were mostly absent in *rde-4*; *drh-1* double mutants (***Figure 4K***). These data support the conclusion that residual DCR-1 activity on viral dsRNA in *drh-1* mutants is dependent on RDE-4. Thus, the double mutant data confirms the position of DRH-1 as an upstream factor essential for the generation of robust levels of antiviral siRNA in response to infection.

Discussion

Overall our data support a key and unique role for the *C. elegans* RIG-I-like protein DRH-1 in primary siRNA synthesis by either guiding DCR-1 activity to the viral genome or assisting DCR-1 processing of the double-stranded viral RNA. We suggest that the physical interaction between DRH-1 and DCR-1, and the potential for DRH-1 to recognize the viral genome as foreign, possibly through its well-conserved RIG-I domain, may enable DRH-1 to recruit DCR-1 to the double-stranded replicating viral genome and instigate a hierarchical antiviral siRNA response (***Figure 5***). Our data also show that the role of DRH-1 in viral recognition is distinct from that of its paralogs. Given the sequence and domain similarities between DRH-1, DRH-2 and DRH-3, it will be of interest to determine the mode of RNA recognition by DRH-2 and DRH-3 in the future.

Our findings imply that parallel RNAi pathways are involved in recognition of viral infection, aberrant endogenous transcripts and exogenous RNAi (exo-RNAi) in an experimental setting. It is interesting therefore that DRH-1 is found in a protein complex with DCR-1 and RDE-4 even in the absence of infection (***Thivierge et al., 2012***). This might suggest that this complex limits the availability of DCR-1 for exo-RNAi. However, *drh-1* and N2 worms respond equally to exo-RNAi, suggesting that this is not the case (***Figure 4—figure supplement 4***). The constitutive nature of the DCR/DRH-1 complex may allow cells to respond much more rapidly to the presence of viral replication intermediates. Recent observations from Rui Lu and colleagues are in agreement with our findings (***Guo et al., 2013***).

Antiviral small RNAs generated by Dicer are an evolutionarily conserved mechanism for fighting infection by positive-strand RNA viruses (***Aliyari and Ding, 2009***). However, there must be some mechanism to distinguish between the viral genome and cellular RNAs. *Drosophila* and plants have a dedicated Dicer enzyme responsible for viral dsRNA recognition. We suggest that in *C. elegans* DRH-1 may perform this function. The absence of a role for *drh-1* in the endogenous small RNA pathway (***Figure 4—figure supplement 3***; ***Gu et al., 2009***) supports the idea that it encodes a viral-specific recognition factor. In mammals, the C-terminal domain of the DRH-1 ortholog RIG-I is able to recognize the 5′ triphosphate on viral genomes in the context of the double-stranded replication intermediate (***Hornung et al., 2006***; ***Pichlmair et al., 2006***; ***Rehwinkel et al., 2010***). The C-terminal domain is conserved in DRH-1, thus suggesting that DRH-1 might use a recognition mechanism analogous to RIG-I to recruit DCR-1 to the Orsay virus RNA. In support of this, 5′ RACE experiments from total RNA from infected animals using Orsay virus specific primers fail to detect product unless the 5′ triphosphate is removed prior to adaptor ligation (***Figure 4—figure supplement 3***), suggesting that the majority of Orsay genomic RNA molecules are indeed phosphorylated at the 5′ end.

It is interesting that the recognition function of DRH-1 may be conserved with mammalian RIG-I whilst the effector pathways are apparently distinct. It will therefore be intriguing to examine whether any part of the function of DRH-1 in antiviral RNAi might be conserved in mammals. As yet, although small RNA responses have been analyzed in mammalian cells infected with viruses (***Parameswaran et al., 2010***) it is not clear whether these have a significant role in the defense of viral infection or

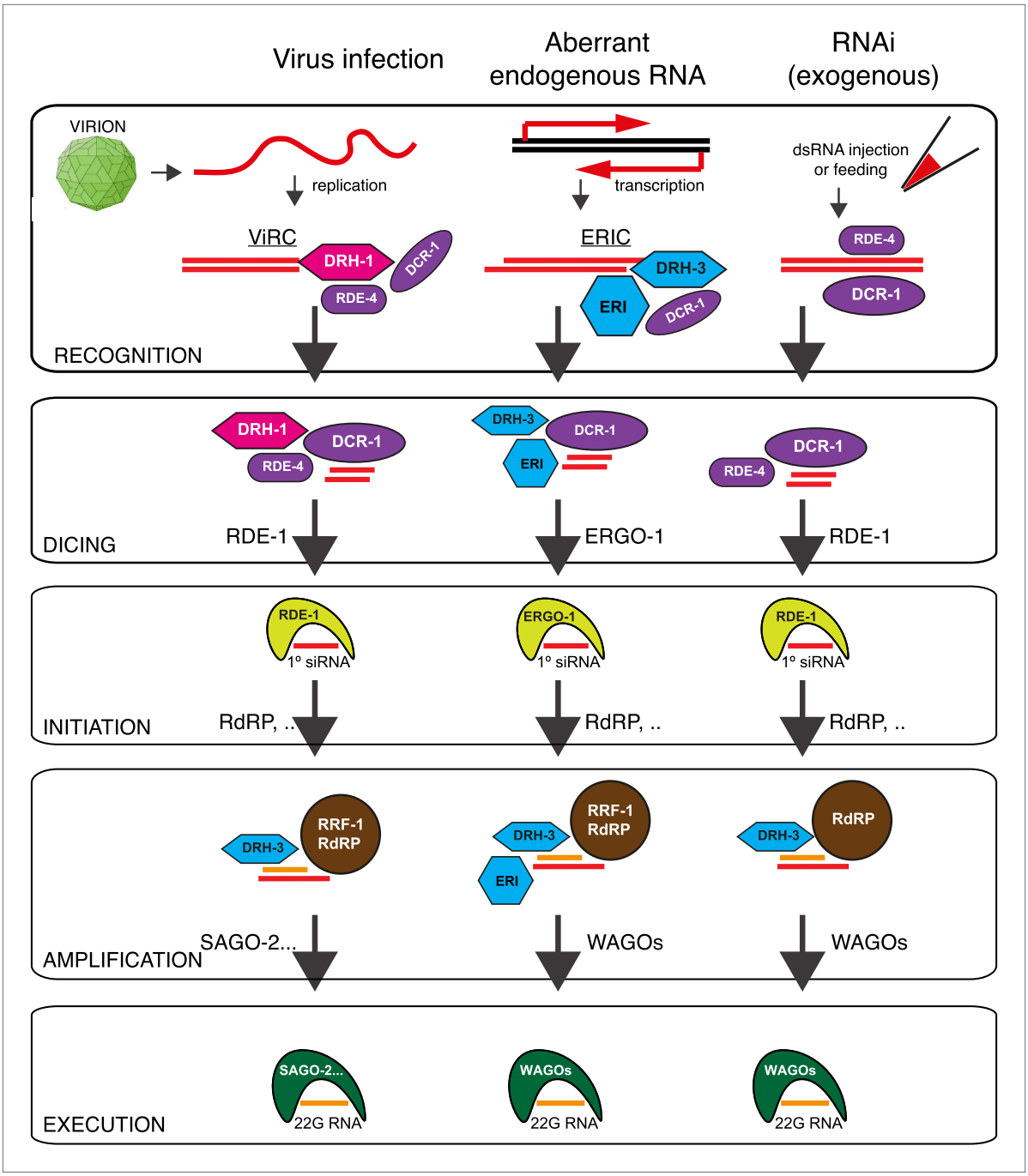


Figure 5. Model: DRH-1 triggers a hierarchical antiviral RNAi pathway. Upon infection of the N2 *C. elegans* strain by the Orsay virus, DRH-1 recruits DCR-1 and its partner RDE-4 to the viral dsRNA replication intermediate. DCR-1 cleaves the viral genome into 23 nt viRNA duplexes with a 2 nt 3′ overhang. Duplex viRNAs are incorporated into the Argonaute protein RDE-1 and one strand is lost to give rise to primary viral siRNAs (primary viRNAs). Primary viRNAs and RDE-1 recruit an RdRP complex to the viral genome to synthesize secondary viral siRNAs, which act to silence viral transcripts or inhibit virus replication. The antiviral RNAi pathway is dependent on the SAGO-2 secondary Argonaute protein (***Figure 3—figure supplement 1C***). The antiviral RNAi pathway has parallels to the exogenous RNAi pathway and the endogenous RNAi pathway thought to recognize aberrant endogenous transcripts (***Gu et al., 2009***). A complex of DRH-1, DCR-1 and RDE-4 has previously been observed in whole animal lysates (***Tabara et al., 2002***; ***Duchaine et al., 2006***; ***Thivierge et al., 2012***). We refer to this complex as the Viral Recognition Complex (ViRC). ERI, other ERI factors. ERIC, ERI Complex.

DOI: 10.7554/eLife.00994.018

whether they display any similarities to antiviral RNAi pathway in plants or nematodes. Examining whether cells deficient in RIG-I show alterations in small RNA pathways upon viral infection may help to address this question. Equally, it will be interesting to identify whether DRH-1 has a signaling role in *C. elegans* upon viral infection. Such studies might indicate whether the gene expression or the RNAi function is the ancestral role of RIG-I-like genes.

RIG-I and MDA-5 respond to viral infection with changes in cytoplasmic localization that result in the activation of the interferon response (**Nakhaei et al., 2009**). In addition, *RIG-I* and *MDA-5* are themselves interferon-induced genes. It will therefore be of great interest to explore the behavior of DRH-1 upon viral infection. A recent genome-wide analysis of gene expression upon Orsay virus infection in *C. elegans* did not detect any significant changes in *drh-1* transcript levels (**Sarkies et al., 2013**). Furthermore, a GFP-DRH-1 fusion protein did not show marked alterations in expression levels or subcellular localization upon infection with the Orsay virus (our unpublished observations). It will be important to analyze the behavior of the endogenous DRH-1 protein upon infection in future studies.

Our discovery of an inactivating deletion in *drh-1* carried by many wild isolates of *C. elegans* is consistent with the rapid evolution under strong selection known to characterize proteins involved in immunity, including antiviral RNAi defense (**Obbard et al., 2006**; **Vasseur et al., 2011**). Yet it is puzzling that a derived allele with deleterious consequences in the presence of viral infection is found at intermediate frequency. One possible explanation might be the low natural occurrence of infecting viruses, meaning that the *drh-1* deletion is effectively neutral. In support of this argument we have not been able to detect similar intestinal symptoms of viral infection in our extensive *C. elegans* sampling (**Félix and Duveau, 2012**), including in other years of sampling in the Orsay orchard. Alternatively the *drh-1* deletion may have hitch-hiked to fixation with a closely linked beneficial mutation. Such a phenomenon is likely to be common in *C. elegans* natural populations due to the low effective outcrossing rate, which results in high linkage disequilibrium, especially when associated with a positive selective sweep (**Cutter et al., 2009**). Indeed we find that the natural *drh-1* deletion allele is in high linkage disequilibrium with the surrounding region of chromosome IV (**Figure 2E**). A related possibility is that the *drh-1* deletion itself might have a positive fitness effect in some natural conditions. Although we did not detect a fitness advantage for the region surrounding the *drh-1* deletion under laboratory conditions, we cannot test all possible natural conditions, thus a beneficial effect for the *drh-1* deletion or its surrounding region remains a possibility. Interestingly, RIG-I appears to have been lost in several clades, including chickens (**Zou et al., 2009**; **Barber et al., 2010**; **Table 1**), which might explain the increased sensitivity of chickens to avian influenza virus when compared to ducks (**Barber et al., 2010**).

In conclusion, we found that the RIG-I domain has an ancient role in antiviral immunity outside of mammals, yet that this broad conservation in animals is compatible with recurrent losses in insects, ducks, and down to some *C. elegans* isolates. Our results further indicate that the conserved biochemical activity of RIG-I is viral recognition, whereas downstream effector pathways, such as RNAi and interferon responses, may differ between *C. elegans* and mammals.

Materials and methods

Genetics

C. elegans were grown under standard conditions at 20°C unless otherwise indicated. The wild-type strain was var. Bristol N2 (**Brenner, 1974**). All strains used are listed in **Supplementary file 1A**.

Virus filtrate preparation

Virus filtrate was prepared as described previously (**Félix et al., 2011**).

Genome-wide association mapping in wild isolates

Choice and infection of the 97 natural isolates of *C. elegans*

We used 97 natural isolates that have been partly sequenced in a previous study using RAD-sequencing (**Andersen et al., 2012**). We first submitted each isolate to a bleaching treatment that eliminates horizontally transmitted symbionts such as the Orsay virus (**Félix et al., 2011**). For each isolate, we infected two 55-mm plates containing five young adults, in triplicate. Cultures were incubated with the Orsay virus at 23°C for 7 days. Maintenance over more than 4 days after infection was performed by transferring a piece of agar (approximately 0.1 cm^3) every 2–3 days to a new plate with food. For each infection batch, JU1580 was infected as a control. At 7 days post-infection, nematodes from two plates were collected in M9 and RNA was extracted as described previously (**Félix et al., 2011**).

Table 1. Evolution of Dicer and RIG-I family proteins

				helicase + RIG-I structures in species	Dicer
Cnidaria			*Nematostella vectensis*	2	1
Bilateria	Protostomia	Polyzoa			
		Platyzoa	*Schmidtea mediterranea*	0	1
		Kryptochozoa			
		Mollusca	*Aplysia californica*	0	1
		Annelida	*Platynereis dumerilii*	0	1
		Ecdysozoa	*Drosophila melanogaster*	0	2
			Trichinella spiralis	2	1
			Caenorhabditis elegans	3	1
	Deuterostomia		*Branchiostoma floridae*	2	1
			Meleagris gallopavo	3	1
			Taeniopygia guttata	3	1
			Gallus gallus	2	1
			Homo sapiens	3	1

Presence of Dicer and RIG-I family proteins in selected animals. Data were obtained from Pfam (version 26.0) (***Finn et al., 2010***) (pfam.sanger.ac.uk). RIG-I family proteins were identified by having both helicase domains and the RIG-I C-terminal domain (Pfam: PF11648). Available sequence data is sparse for some clades and absence of data might not be sufficient evidence for absence of genes.

DOI: 10.7554/eLife.00994.019

RT-qPCR for GWAS

cDNA was generated from 1 µg total RNA with random primers using Superscript III (Life Technologies, Foster City, CA). cDNA was diluted to 1:10 for RT-qPCR analysis. RT-qPCR was performed using LightCycler 480 SYBR Green I Master (Roche, Mannheim, Germany). The amplification was performed on a LightCycler 480 Real Time PCR System (Roche). Each sample was normalized to *eft-2*, and then viral RNA1 (primers GW194 and GW195 in the 97) or RNA2 (oTB17 and oTB18 in the recombinant analysis) levels were compared to the level present in a reference RNA extract obtained from infected JU1580 animals.

Association mapping

Association mapping was performed using the EMMA package with the default kinship matrix (***Kang et al., 2008***). The mapped trait was the log-transformed mean value of the qPCR on the RNA1 of the Orsay virus.

Mapping using recombinant F2-derived families

Production and infection of the F2-derived families

JU1580 hermaphrodites were crossed to N2 males. 11 heterozygous larvae (stage L4) of the F1 cross progeny were picked singly in 35-mm plates and allowed to self. 110 recombinant F2 larvae (stage L4) were picked as individuals into 35-mm Petri dishes. F2 animals were cultured at 20°C for 24 hr and then inoculated with 10 µl of Orsay virus filtrate. After 4 days of culture at 20°C, the F2 animals and their F3/F4 progeny were resuspended into M9 and washed three times in 1 ml of M9. Animals were pelleted in approximately 20 µl of M9 after the last wash. 5 N2 and 5 JU1580 animals were treated in the same way than the recombinant F2 animals to serve as controls.

RT-qPCR of F2 families

To measure the viral load in the infected F2 families, 5 µl of nematode pellets were added to 45 µl of lysis solution (with 1:100 DNase I) from the Power SYBR Green Cells-to-Ct kit (Ambion, Austin, TX). The lysis mixtures were freeze-thawed 10 times using liquid nitrogen and a hot water bath, and then vortexed for 30 min in 96-well plates. 5 µl of Stop Solution (Ambion) were added to each lysis mixture to

complete the RNA extraction step. The cDNA synthesis and the qPCRs were then performed using the Power SYBR Green Cells-to-Ct kit (Ambion). We used 2 µl of cDNA (equivalent of cDNA contained in 0.09 µl of nematode pellet) as a template for the qPCR. The level of viral RNA was normalized to *gapdh*.

Sequencing

DNA from JU1580 or pools of 20 N2 × JU1580 recombinant lines was isolated and DNA libraries were prepared for high-throughput sequencing using the Nextera DNA sample prep kit from Illumina. Sequencing was 150 paired end sequencing was performed on a Illumina MiSeq instrument. The JU1580 genomic reads are accessible at the NCBI Short Read Archive under the accession number: SRS369862.

Sequencing analysis

Processing of Illumina reads, mapping and variation detection was performed using the CLC Genomics Workbench 5 (version 5.5.1). N2 reference genome and annotation used was WormBase release WS220 (www.wormbase.org). SNP and indel analysis was performed using the R statistical environment.

Introgression of the candidate region into the N2 genomic background

Production and infection of the recombinant animals

N2 males were crossed to JU1580 hermaphrodites, and the F1 males outcrossed to N2 hermaphrodites. As *niDf250* had already been identified in the genome sequence analysis, F2 animals were genotyped by PCR for the presence of its JU1580 allele in the 6 Mb central region of chromosome IV with primers oTB40 and oTB43 (***Supplementary file 1B***). From one heterozygous animal, 20 homozygous *niDf250* F3 animals were selected and crossed to N2 males to start another cycle of introgression. We repeated this cycle twice and in the final cycle 20 homozygous *niDf250* F3 animals were isolated. We used a pyrosequencing genotyping method (PyroMark Q96 ID instrument; Biotage) to select among these 20 animals those with the full 6 Mb central region of chromosome IV from JU1580 into the N2 genetic background, and only this region. We genotyped six SNPs on chromosome IV (one at either end of the chromosome, one at either end of the 6 Mb region and two within this region), three SNPs on chromosomes I, III, V, X and two on II. We thus obtained an introgressed line, called JU2170. To obtain recombinants within the 6 Mb central region of chromosome IV for fine mapping, JU2170 hermaphrodites were crossed to N2 males. From 20 F1 heterozygous hermaphrodites, we isolated 300 F2 individuals and selected for recombinants between the two SNPs at either end of the region (IV: 3,877,431 and IV: 11,083,410). We thus obtained five recombinants. We further narrowed down the recombination break point by genotyping inside the candidate region (***Figure 1—figure supplement 2***).

Pyrosequencing

SNPs were genotyped by pyrosequencing, using a PyroMark Q96 ID instrument from Biotage (Uppsala, Sweden). The PSQ Assay Design software was used to design pyrosequencing primers. To biotinylate one strand in the PCR reaction, we added a universal tag on one primer, as described (***Aydin et al., 2006***). For each strain, 10 adult animals were mixed with 10 µl of worm lysis buffer (***Fay and Bender, 2008***; ***Félix et al., 2011***) containing proteinase K at 100 µg/ml. After lysis at 60°C for 1 hr and proteinase K inactivation at 95°C for 15 min, we added 1 µl of worm lysate to 50 µl of PCR mix composed of: 0.25 µl of GoTaq DNA polymerase (Promega, Madison, Wisconsin), 5 µl of dNTP at 2 mM, 10 µl of 5x GoTaq buffer, 0.17 µl of non-biotinylated primer at 10 mM, and 0.87 µl of corresponding universal biotinylated primers at 10 mM, and 1 µl of the reverse primer. Single-stranded DNA was then purified and the pyrosequencing reaction was performed following the manufacturer's indications. Primers are listed in ***Supplementary file 1B***.

drh-1 expression

Non-synchronized animals were cultured in a 55-mm plate and collected in M9 just before starvation. RNA extraction and RT-qPCR were performed as described previously (***Félix et al., 2011***).

drh-1 rescue in the JU1580 strain

JU1580 animals were transformed as described (***Mello and Fire, 1995***) with the fosmid WRM0640dC01 that contains the entire length of the *drh-1* gene and its operon CEOP4647. The injection mix contained

10 ng/μl of fosmid DNA, 5 ng/μl of the co-marker transgene *myo-3::gfp::unc-54*, 85 ng/μl of 1 kb DNA ladder (Invitrogen), 20 mM potassium phosphate pH 7.5, and 3 mM potassium citrate pH 7.5. The transgene was then integrated via X-ray irradiation as described (***Fire, 1986***). We controlled that the transgenic copy of *drh-1* was transcriptionally active in transformed animals by qRT-PCR on a portion of the mRNA that is deleted in the wild JU1580 strain.

Viral load in mutants in the RNAi pathway and in *drh-1* rescued JU1580

Infection of strains of interest

For all strains mentioned in ***Figures 3 and 4***, as well as for the strain PD8753 that carries a balanced mutation in *dcr-1*, one or two young adults were inoculated with 20 μl of viral filtrate for 4 days at 20°C (or 15°C for *mago-12* animals) in 55-mm plates in five biological replicates.

RT-qPCR on mutant strains

4 days after infection, all animals except those from the PD8753 strain were collected in M9; RNA extraction and RT-qPCR were performed as described previously (2). Aliquots of RNA were kept apart for small RNA libraries (see below). For the infected progeny of PD8753 parents, 16 *dcr-1* and 16 *dcr-1/+* adults per replicate were selected under the fluorescent microscope (i.e., the balancer chromosome carries a *gfp* reporter) and lysed individually in 10 μl of lysis solution (with 1:100 DNase I) from the Power SYBR Green Cells-to-Ct kit (Ambion). The lysis mixtures were pooled for each replicates and the qRT-PCR was performed as indicated above for the F2 families.

Small RNA sequencing

Library preparation and sequencing

Preparation of RNA used for small RNA libraries is described above. For 5′ independent libraries, 3–5 μg of total RNA was pre-treated with 5′ polyphosphatase (Epicenter) following the manufacturer's instructions. Small RNA libraries were generated from either polyphosphatase treated or total RNA (3–5 μg) using the TruSeq Small RNA kit (Illumina) following the manufacturer's instructions. Small RNA libraries were sequenced using the Illumina MiSeq. Small RNA sequence data were submitted to the Gene Expression Omnibus (GEO) under accession number GSE41693.

Sequencing analysis

Small RNA libraries were sequenced using the Illumina MiSeq. Fastq files generated by the machine had adaptors removed using the program Cutadapt v1, and were converted into Fasta files using a custom Perl script. For alignment to the viral genome, Fasta files were trimmed to leave only reads of between 15 and 30 nucleotides using a custom Perl script and were aligned using Bowtie (version 0.12.7) to RNA1 and RNA2 of the Orsay virus genome (***Félix et al., 2011***), reporting only the best single alignment with up to one mismatch allowed. Sam files from Bowtie were converted into bam files using the SAMtools utility (***Li and Durbin, 2009***) and bam files were converted into bed files using the BedTools utility (***Quinlan and Hall, 2010***). Bed files were read into the R environment and plots of the length and first nucleotide as well as read distribution along RNA2 were generated using custom scripts written in R. The level of virus as measured by qRT-PCR was observed to correlate non-linearly with the total level of sense small RNAs in different N2 wild-type samples, thus data were normalized to all viral siRNAs or all sense viRNAs to illustrate differences in read distribution between different samples. Analysis of the overlap to identify potential DCR-1 signatures was carried out by selecting reads of the same length coming from opposite strands on the viral genome, which overlapped by at least one nucleotide. The number of reads overlapping by every possible length either 3′ or 5′ was compared to that expected were the entire distribution uniform using a χ^2 test. Only one overlap was statistically significantly enriched with a cut-off of $p<0.05$ as indicated in the text. For alignment to the *C. elegans* genome to analyze any potential changes in 22G levels, fasta files were trimmed to leave only reads of between 18 and 30 nucleotides, and the fasta files were collapsed using the FastX Toolkit. Reads matching to microRNA precursors downloaded from miRbase (***Kozomara and Griffiths-Jones, 2011***) were removed using a custom Perl script, and the remaining reads were aligned to the Ce6 genome (WS190) using Bowtie, reporting the best match with no mismatches (parameters—best–k 1–v 0). After converting bam files to bed files as above, 22Gs mapping antisense to the Ce6/WS190 UCSC annotations of genes (sangerGene.txt) downloaded from the

UCSC genome browser website, were then selected from the alignments using a custom Perl script. Genes with abundant 22G reads in N2 wild-type worms were further selected by setting a cut-off of at least five reads in at least one of the N2 wild-type samples and duplicate or overlapping annotations were removed. Comparison of mutants to N2 wild-type was carried out after normalizing to the total number of aligned reads in each library. Secondary structure analysis of the viral genome was performed using RNA-fold (***McCaskill, 1990***) on a 50 bp window sliding by increments of 20 bp. A Z-score for the secondary structure strength of each window was calculated by comparing the ensemble free energy mean ensemble free energy for 100 random shuffles of the 50 bp window.

5′ RACE analysis of the Orsay virus

Total RNA isolated from infected animals as described above was ligated to the 5′ adaptor from the Illumina Truseq small RNA kit either with or without prior treatment with 5′ polyphosphatase. Only monophosphorylated 5′ ends will be able to ligate to the adaptor, as for the small RNA sequencing. Ligated products were then reverse transcribed using a primer specific for RNA1. The resulting cDNA was then analysed by standard Taq PCR and gel electrophoresis using either primers designed to amplify within the 5′ end of the viral genome as a positive control or primers to amplify from the adaptor into the 5′ end, thus determining whether the adaptor was able to ligate efficiently to the 5′ end (***Figure 4—figure supplement 3B***).

Molecular evolution analysis of the 6 Mb central region of chromosome IV

For each of the 97 isolates, 97 SNPs included in the 6 Mb central region of chromosome IV associated with Orsay virus sensitivity were extracted from RAD-sequencing data (***Andersen et al., 2012***). DNAsp allowed us to classify all isolates in 28 different haplotypes for this region. The neighbor-net network (***Bryant and Moulton, 2004***) was then drawn using the SplitTree software (***Librado and Rozas, 2009***). The average number of differences per polymorphic RAD site along chromosome IV was calculated using DNAsp (***Huson and Bryant, 2006***) using a sliding window of 25 SNPs every 10 SNPs. The linkage disequilibrium (*D′*) between *drh-1* alleles and polymorphic RAD sites on chromosome IV was calculated by DNAsp.

Progeny and longevity assays

Progeny production

For each strain, we seeded ten 55-mm NGM plates with five L4 animals on each. 5 of the 10 plates were infected with the Orsay virus. Before starvation was reached, we transferred a piece of agar (4 × 4 mm at the surface, 1 cm to the bottom of the plate) from each of the five plates of each treatment (to maximize infection probability) to a single 90 mm NGM plate. 40 L4 animals (F2 progeny from the infected animals) were then isolated for each treatment. Each scored individual was transferred into a new plate 24, 36, 48, 60, 72, 96 and 120 hr after the L4 stage, and progeny number was scored 48 hr after each transfer (most were then at the L4 and adult animals). To ease scoring, some plates were cooled to 4°C after 48 hr and scored within 2 days. N2 and the *drh-1* mutant in the N2 background (RB2519) were assayed in parallel. JU1580 and JU1580 rescue (SX2377) were independently blind tested in parallel.

Longevity assay

From the same starting populations, we in addition isolated for the survival assay ten animals on nine 55 mm diameter NGM plates (n = 90). Each pool of 10 animals was then transferred to a new plate every 24 hr. Survival was assessed every day by scoring the capacity of each animal to react to stimulation (shaking plates and touching them with a pick). Dead animals were removed from the plate. We also measured the lifespan of each parent from the progeny scoring (n = 40).

Infection rates

Each F2 population was fixed and prepared for FISH as previously described (***Kowalinski et al., 2011***) using custom Stellaris (Biosearch Technologies) probes for the Orsay virus RNA1 molecule (***Supplementary file 1B***) labeled with Quasar 670 Dye. Standard fluorescence microscopy was performed using an upright Zeiss AxioImager M1 equipped with a Pixis 1024B camera (Princeton instruments) and a Lumen 200 metal arc lamp (Prior Scientific). A L4 larva was scored as infected when we could see a signal inside at least one intestinal cell. The proportion of infected L4 individuals was the following: 35.0% (n = 109) of JU1580; 3.4% (n = 121) of SX2377; 53.0% (n = 103) of RB2519; 9.0% (n = 100) of N2 animals.

Statistical analysis

The lmer function from the R package « lme4 » was used to determine the effect of the treatment on the dynamics of progeny production by comparing with an ANOVA the Akaike Information Criterion of a model involving both time, treatment and their interaction to that of a null model involving only time. For the longevity assay, we used the survdiff function (log rank test) of the R package « survival » as to test the effect of the treatment.

Competition assay

Competition

The N2 and JU2196 genotypes were competed in the presence or absence of the virus. JU2196 is a near isogenic line with the 4.4 Mb central part of chromosome IV from JU1580 into the N2 background (***Figure 1—figure supplement 2A***). Both strains were cleaned and roughly synchronized by bleaching. Once the bleached embryos had developed to adulthood, we let them lay embryos for 1 hr and half to further synchronize the populations. After letting the embryos hatch overnight, we started the competition experiment with L2 hermaphrodite larvae in a 1:1 ratio (20 plates with 10 N2 and 10 JU2196 animals on each), adding 50 µl of Orsay virus preparation on half of these plates and 50 µl of M9 solution in the other. Populations were harvested 5 days later with M9 solution under sterile conditions in microfuge tubes. After 2 min of centrifugation at 3000 rpm, the supernatant was removed and a fraction of the pellet containing 100–400 animals (generally 2 µl) was transferred to a new NGM plate. After this initial transfer, transfers were repeated approximately every 36 hr, before starvation. We did not observe a high occurrence of males in this experiment, thus it is unlikely that the tested chromosome IV region recombined.

Genotyping

At each transfer, 2 µl of the pellet was mixed with 18 µl of worm lysis buffer containing proteinase K at 100 µg/ml. The product of the worm lysis was then used as a PCR template using the pyrosequencing primers IV_6124501 Forward and Reverse (***Supplementary file 1B***). We quantified the proportion of the *drh-1*(N2) allele using the quantitative option of the pyrosequencer (PyroMark Q96 ID instrument; Biotage). To measure the accuracy of this quantification method, a standard curve was performed with different known proportions of alleles (known proportions of genomic DNA). The correlation between observed and real allele frequencies exceeded 0.983 and the average standard deviation calculated from five replicates of observed frequencies was 6%.

Statistical analysis

The lmer function from the R package « lme4 » was used to determine the effect of the treatment on the genotype frequency by comparing with an ANOVA the Akaike Information Criterion of a model involving both time, treatment and their interaction to that of a null model involving only time.

RNAi

unc-22 and empty vector RNAi bacteria were grown for 6 hr at 37°C. *unc-22* bacteria were then serially diluted with the empty vector bacteria and seeded onto NGM agar plates containing IPTG (1 mM) and carbenicillin (25 µg/ml). After drying overnight, N2, *drh-1* or *eri-1* worms were added and then grown at 20°C for 4 days. Each strain was tested in triplicate at each dilution and 15 animals selected at random from each plate were scored for either twitching or paralysis phenotypes.

Acknowledgements

We thank M Tanguy and J Rehwinkel for critical insights into viral biology, C Bradshaw and G Allen for Bioinformatics support and S Moss, I Nuez and A Richaud for laboratory management. We are grateful to C Franz and the D Wang laboratory for viral filtrate preparations, to E Andersen for help with the whole-genome association, and to M De Bono for reagents. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Some deletion mutations were provided by the International *C. elegans* Gene Knockout Consortium which is funded by the National Institute of Health, the Canadian Institute for Health Research, Genome Canada, Genome BC, and the Michael Smith Foundation. Fosmids were provided by D Moerman laboratory. AA was supported by a Herchel-Smith postdoctoral fellowship. PS was supported by a research fellowship at Gonville and Caius College, Cambridge.

Additional information

Funding

Funder	Grant reference number	Author
Wellcome Trust	092096	Jérémie Le Pen, Eric A Miska
Cancer Research UK	RG57329	Eric A Miska
European Research Council	RG58558	Eric A Miska
Centre National de la Recherche Scientifique		Marie-Anne Félix, Tony Bélicard
Agence Nationale pour la Recherche	ANR 11 BSV3 01301	Marie-Anne Félix, Lise Frézal
Coup d'Elan de la Fondation Bettencourt-Schueller		Marie-Anne Félix
Ecole Normale Supérieure		Marie-Anne Félix

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

AA, TB, JLP, PS, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; LF, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; NJL, Acquisition of data, Contributed unpublished essential data or reagents; M-AF, EAM, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Additional files

Supplementary files

• Supplementary file 1. (**A**) Table 1: Strains used in this study. (**B**) Table 2: Oligonucleotides used in this study.

DOI: 10.7554/eLife.00994.020

Major datasets

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset ID and/or URL	Database, license, and accessibility information
Sarkies P, Miska EA	2013	Analysis of small RNA response to viral infection in *C. elegans*	GSE41693; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41693	Publicly available at GEO (http://www.ncbi.nlm.nih.gov/geo/).
Sarkies P, Miska EA	2013	Sequencing of a *C. elegans* wild isolate JU1580	SAMN01766984; http://www.ncbi.nlm.nih.gov/biosample/?term=SAMN01766984	Publicly available at the NCBI BioSample database (http://www.ncbi.nlm.nih.gov/biosample).

References

Aliyari R, Ding S-W. 2009. RNA-based viral immunity initiated by the Dicer family of host immune receptors. *Immunol Rev* **227**:176–88. doi: 10.1111/j.1600-065X.2008.00722.x.

Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, et al. 2012. Chromosome-scale selective sweeps shape *Caenorhabditis elegzans* genomic diversity. *Nat Genet* **44**:285–90. doi: 10.1038/ng.1050.

Aydin A, Toliat MR, Bähring S, Becker C, Nürnberg P. 2006. New universal primers facilitate Pyrosequencing. *Electrophoresis* **27**:394–7. doi: 10.1002/elps.200500467.

Barber MRW, Aldridge JR, Webster RG, Magor KE. 2010. Association of RIG-I with innate immunity of ducks to influenza. *Proc Natl Acad Sci USA* **107**:5913–8. doi: 10.1073/pnas.1001755107.

Bouché N, Lauressergues D, Gasciolli V, Vaucheret H. 2006. An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J* **25**:3347–56. doi: 10.1038/sj.emboj.7601217.
Brenner S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics* **77**:71–94.
Bryant D, Moulton V. 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* **21**:255–65. doi: 10.1093/molbev/msh018.
Cutter AD, Dey A, Murray RL. 2009. Evolution of the *Caenorhabditis elegans* genome. *Mol Biol Evol* **26**:1199–234. doi: 10.1093/molbev/msp048.
Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. 2006. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* **313**:68–71. doi: 10.1126/science.1128214.
Diaz-Pendon JA, Li F, Li W-X, Ding S-W. 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* **19**:2053–63. doi: 10.1105/tpc.106.047449.
Ding S-W. 2010. RNA-based antiviral immunity. *Nat Rev Immunol* **10**:632–44. doi: 10.1038/nri2824.
Ding S-W, Voinnet O. 2007. Antiviral immunity directed by small RNAs. *Cell* **130**:413–26. doi: 10.1016/j.cell.2007.07.039.
Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D, Pang K, et al. 2006. Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124**:343–54. doi: 10.1016/j.cell.2005.11.036.
Fay D, Bender A. 2008. SNPs: introduction and two-point mapping. *WormBook* 1–10. doi: 10.1895/wormbook.1.93.2.
Félix M-A, Ashe A, Piffaretti J, Wu G, Nuez I, Bélicard T, et al. 2011. Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLOS Biol* **9**:e1000586. doi: 10.1371/journal.pbio.1000586.
Félix M-A, Duveau F. 2012. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biology* **10**:59. doi: 10.1186/1741-7007-10-59.
Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, et al. 2010. The Pfam protein families database. *Nucleic Acids Res* **38**:D211–22. doi: 10.1093/nar/gkp985.
Fire A. 1986. Integrative transformation of *Caenorhabditis elegans*. *EMBO J* **5**:2673–80.
Flynt A, Liu N, Martin R, Lai EC. 2009. Dicing of viral replication intermediates during silencing of latent Drosophila viruses. *Proc Natl Acad Sci USA* **106**:5270–5. doi: 10.1073/pnas.0813412106.
Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, Ellacott GA, et al. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep* **7**:1168–75. doi: 10.1038/sj.embor.7400837.
Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler J-L. 2006. Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. *Nat Immunol* **7**:590–7. doi: 10.1038/ni1335.
Gitlin L, Barchet W, Gilfillan S, Cella M, Beutler B, Flavell RA, et al. 2006. Essential role of mda-5 in type I IFN responses to polyriboinosinic:polyribocytidylic acid and encephalomyocarditis picornavirus. *Proc Natl Acad Sci USA* **103**:8459–64. doi: 10.1073/pnas.0603082103.
Gu W, Shirayama M, Conte D, Vasale J, Batista PJ, Claycomb JM, et al. 2009. Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol Cell* **36**:231–44. doi: 10.1016/j.molcel.2009.09.020.
Guo X, Zhang R, Wang J, Ding S-W, Lu R. 2013. Homologous RIG-I–like helicase proteins direct RNAi- mediated antiviral immunity in *C. elegans* bydistinct mechanisms. *PNAS* doi: 10.1073/pnas.1307453110.
Hornung V, Ellegast J, Kim S, Brzózka K, Jung A, Kato H, et al. 2006. 5′-Triphosphate RNA is the ligand for RIG-I. *Science* **314**:994–7. doi: 10.1126/science.1132505.
Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**:254–67. doi: 10.1093/molbev/msj030.
Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. 2008. Efficient control of population structure in model organism association mapping. *Genetics* **178**:1709–23. doi: 10.1534/genetics.107.080101.
Knight SW, Bass BL. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**:2269–71. doi: 10.1126/science.1062039.
Kowalinski E, Lunardi T, McCarthy AA, Louber J, Brunel J, Grigorov B, et al. 2011. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147**:423–35. doi: 10.1016/j.cell.2011.09.039.
Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**:D152–7. doi: 10.1093/nar/gkq1027.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–60. doi: 10.1093/bioinformatics/btp324.
Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451–2. doi: 10.1093/bioinformatics/btp187.
Lu R, Maduro M, Li F, Li HW, Broitman-Maduro G, Li WX, et al. 2005. Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* **436**:1040–3. doi: 10.1038/nature03870.
Lu R, Yigit E, Li W-X, Ding S-W. 2009. An RIG-I-Like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in *Caenorhabditis elegans*. *PLOS Pathog* **5**:e1000286. doi: 10.1371/journal.ppat.1000286.
McCaskill JS. 1990. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers* **29**:1105–19. doi: 10.1002/bip.360290621.
Mello C, Fire A. 1995. DNA transformation. *Methods Cell Biol* **48**:451–82. doi: 10.1016/S0091-679X(08)61399-0.

Nakhaei P, Genin P, Civas A, Hiscott J. 2009. RIG-I-like receptors: sensing and responding to RNA virus infection. *Semin Immunol* **21**:215–22. doi: 10.1016/j.smim.2009.05.001.
Obbard DJ, Jiggins FM, Halligan DL, Little TJ. 2006. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol* **16**:580–5. doi: 10.1016/j.cub.2006.01.065.
Pak J, Fire A. 2007. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315**:241–4. doi: 10.1126/science.1132839.
Parameswaran P, Sklan E, Wilkins C, Burgon T, Samuel MA, Lu R, et al. 2010. Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLOS Pathog* **6**:e1000764. doi: 10.1371/journal.ppat.1000764.
Parrish S, Fire A. 2001. Distinct roles for RDE-1 and RDE-4 during RNA interference in *Caenorhabditis elegans*. *Rna* **7**:1397–402. doi: 10.1017.S1355838201011074.
Pichlmair A, Schulz O, Tan CP, Näslund TI, Liljeström P, Weber F, et al. 2006. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. *Science* **314**:997–1001. doi: 10.1126/science.1132998.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–2. doi: 10.1093/bioinformatics/btq033.
Rehwinkel J, Reis e Sousa C. 2010. RIGorous detection: exposing virus through RNA sensing. *Science* **327**:284–6. doi: 10.1126/science.1185068.
Rehwinkel J, Tan CP, Goubau D, Schulz O, Pichlmair A, Bier K, et al. 2010. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* **140**:397–408. doi: 10.1016/j.cell.2010.01.020.
Sarkies P, Ashe A, Le Pen J, McKie MA, Miska EA. 2013. Competition between virus-derived and endogenous small RNAs regulates gene expression in *Caenorhabditis elegans*. *Genome Res* **23**:1258–70. doi: 10.1101/gr.153296.112.
Schlee M. 2013. Master sensors of pathogenic RNA - RIG-I like receptors. *Immunobiology* doi: 10.1016/j.imbio.2013.06.007.
Schott DH, Cureton DK, Whelan SP, Hunter CP. 2005. An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* **102**:18420–4. doi: 10.1073/pnas.0507123102.
Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, et al. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**:465–76. doi: 10.1016/S0092-8674(01)00576-1.
Sijen T, Steiner FA, Thijssen KL, Plasterk RHA. 2007. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **315**:244–7. doi: 10.1126/science.1136699.
Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, et al. 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLOS Biol* **1**:E45. doi: 10.1371/journal.pbio.0000045.
Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, et al. 1999. The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**:123–32. doi: 10.1016/S0092-8674(00)81644-X.
Tabara H, Yigit E, Siomi H, Mello CC. 2002. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* **109**:861–71. doi: 10.1016/S0092-8674(02)00793-6.
Thivierge C, Makil N, Flamand M, Vasale JJ, Mello CC, Wohlschlegel J, et al. 2012. Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. *Nat Struct Mol Biol* **19**:90–7. doi: 10.1038/nsmb.2186.
van Rij RP, Saleh M-C, Berry B, Foo C, Houk A, Antoniewski C, et al. 2006. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes Dev* **20**:2985–95. doi: 10.1101/gad.1482006.
Vasseur E, Patin E, Laval G, Pajon S, Fornarino S, Crouau-Roy B, et al. 2011. The selective footprints of viral pressures at the human RIG-I-like receptor family. *Hum Mol Genet* **20**:4462–74. doi: 10.1093/hmg/ddr377.
Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, Machaca K. 2005. RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* **436**:1044–7. doi: 10.1038/nature03957.
Yigit E, Batista PJ, Bei Y, Pang KM, Chen C-CG, Tolia NH, et al. 2006. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**:747–57. doi: 10.1016/j.cell.2006.09.033.
Yoneyama M, Fujita T. 2010. Recognition of viral nucleic acids in innate immunity. *Rev Med Virol* **20**:4–22. doi: 10.1002/rmv.633.
Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, et al. 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5**:730–7. doi: 10.1038/ni1087.
Zou J, Chang M, Nie P, Secombes CJ. 2009. Origin and evolution of the RIG-I like RNA helicase gene family. *BMC Evol Biol* **9**:85. doi: 10.1186/1471-2148-9-85.

Supplementary Files

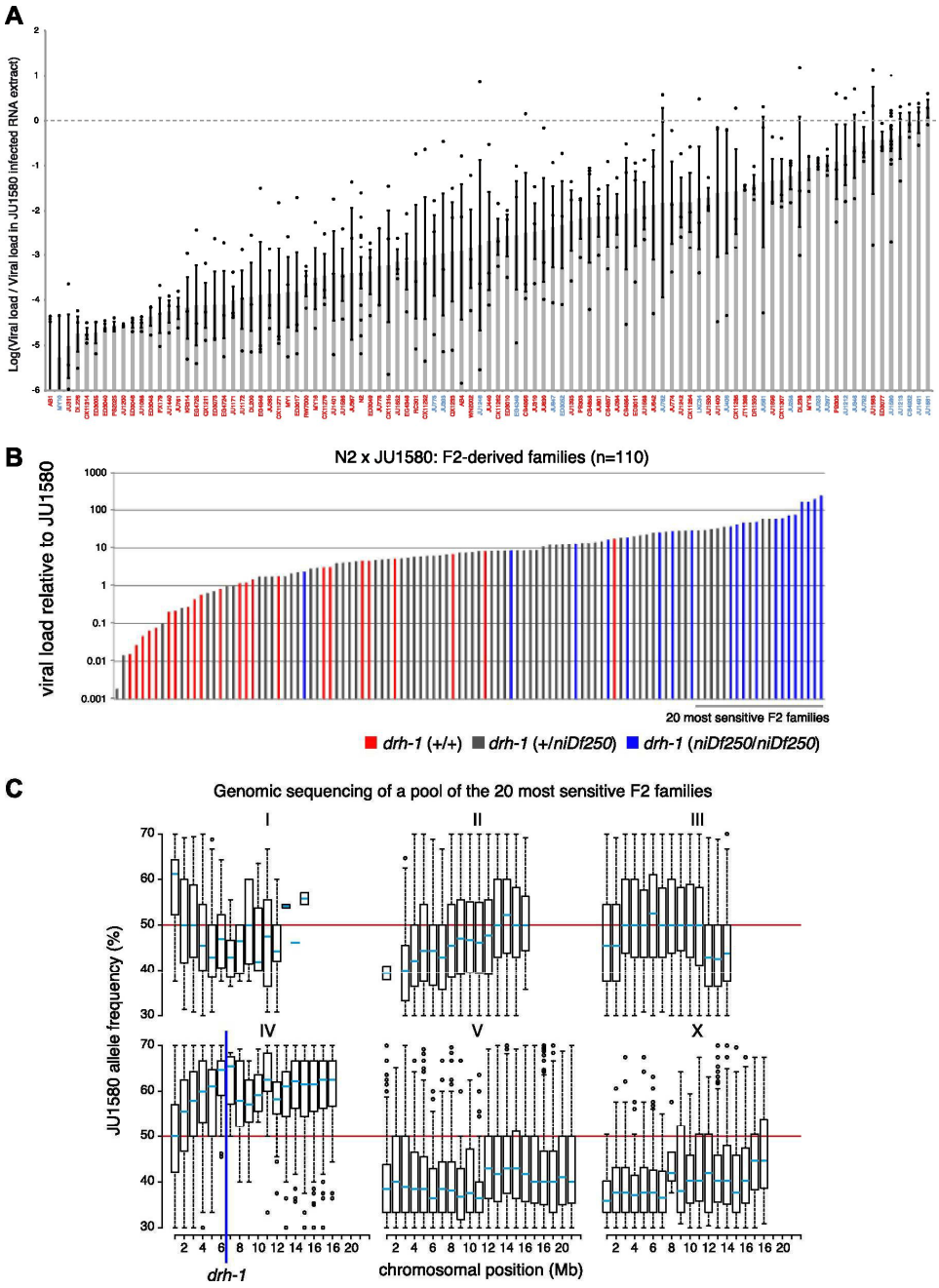
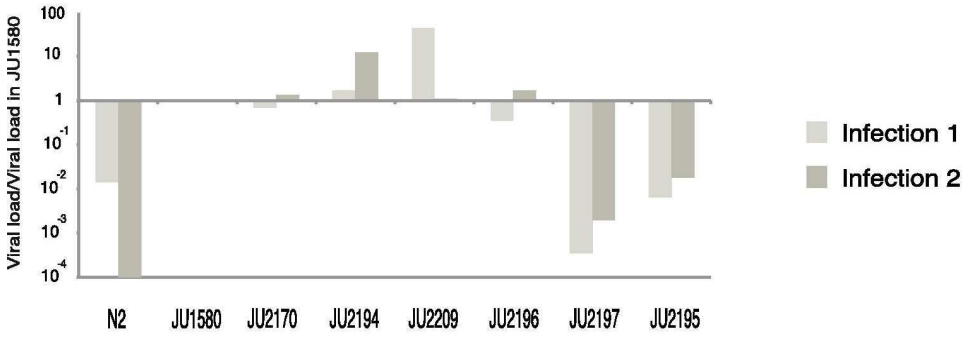


Figure 1—figure supplement 1. Variation in the ability of the Orsay virus to replicate. in *C. elegans*. (A) For each of the 97 wild isolates listed on the horizontal axis, the mean of the logarithm and standard error of the RT-qPCR values on the viral RNA1 assayed 7 days post infection at 23°C are reported. For a given strain, the value of each replicate is represented as a dot. Isolates labeled in blue or red carry the *niDf250* or the N2 allele of *drh-1*, respectively. (B) To investigate whether the viral susceptibility of JU1580 was linked to drh-1, we crossed N2 and JU1580 and allowed the F1 progeny to self-fertilize. Each line thus generated will carry a different combination of N2 and JU1580 SNPs allowing the separation of the drh-1 mutation from any unlinked additional differences in genetic background. After two generations of self-fertilization we then infected these lines and assayed both for sensitivity to infection and for the presence of the JU1580 drh-1 deletion. We saw good correlation between increased viral sensitivity and the *drh-1* deletion (p=2.9×10–8, Wilcoxon test). (C) Distribution of JU1580 SNPs in recombinant lines from B. A pool of 20 sensitive recombinant lines was selected from a total of 110 independent F2 lines and subjected to high-throughput sequencing. Only average SNP frequencies between 20% and 70% were chosen to exclude false SNP calls. This showed that viral sensitivity was linked to chromosome IV. Red lines indicate an average SNP frequency of 50%. The blue line indicates position of the *niDf250* deletion.

A

Parents		*mfIR(IV;JU1580>N2)*						Position on Chr IV (bp)	Genetic position (cM)
N2	JU1580	JU2170	JU2194	JU2209	JU2196	JU2197	JU2195		
A	G	A	A	A	A	A	A	2279105	-8.39
T	C	C	C	C	C	nd	nd	3329219	-4.72
T	G	G	G	G	G	G	T	5393585	1.65
G	A	A	A	A	A	A	G	6124501	3.01
T	G	nd	nd	nd	nd	G	T	6544139	3.20
T	A	nd	nd	nd	nd	T	T	***6567528***	3.20
T	C	nd	nd	nd	C	T	T	***6597185***	3.20
WT *drh-1*	*niDf250*	*niDf250*	*niDf250*	*niDf250*	*niDf250*	WT *drh-1*	WT *drh-1*	***6607636..6607793***	3.21
A	T	nd	nd	nd	T	A	A	***6619249***	3.21
A	T	nd	nd	nd	T	A	A	***6676736***	3.22
G	T	nd	nd	nd	G	G	G	6699208	3.22
C	T	nd	nd	nd	nd	nd	C	7048084	3.28
C	T	T	T	T	C	C	T	7607965	3.36
T	C	C	C	T	T	T	C	8263671	3.70
C	T	T	T	C	nd	C	T	10211735	4.56
T	A	A	T	T	T	T	A	11083410	4.83
A	T	A	A	A	A	A	A	15829314	14.42



B

JU1580 variations on chromosome IV:6,567,528-6,676,736 bp		
Variation type	**functional consequence**	**count**
non-synonymous SNVs	missense	1
	nonsense	0
	frameshift	0
	splice disruption	0
other SNVs		19
structural variations	altered coding sequence	1

SNV = single-nucleotide variants. Variants are as compared to *C. elegans* reference WS220 (www.wormbase.org).

Figure 1—figure supplement 2. Genotype and sensitivity to the Orsay virus of recombinants in the chromosome IV region. (A) Genotype at 17 loci (16 SNPs and the *drh-1* allele) along chromosome IV of the six introgressed and recombinant lines and their parents (N2 and JU1580). Chromosomal segments from N2 or JU1580 are represented in red or blue, respectively. The chromosomal positions of loci in the new candidate region are in bold italic. The viral load of each strain over that in JU1580 is represented below as in Figure 1B, for two replicate infections. (B) Summary of genetic differences between N2 and JU1580 in a 155 kb region on chromosome IV (6,567,528-6,676,736) based on resequencing of JU1580. Only differences predicted to result in altered function were considered.

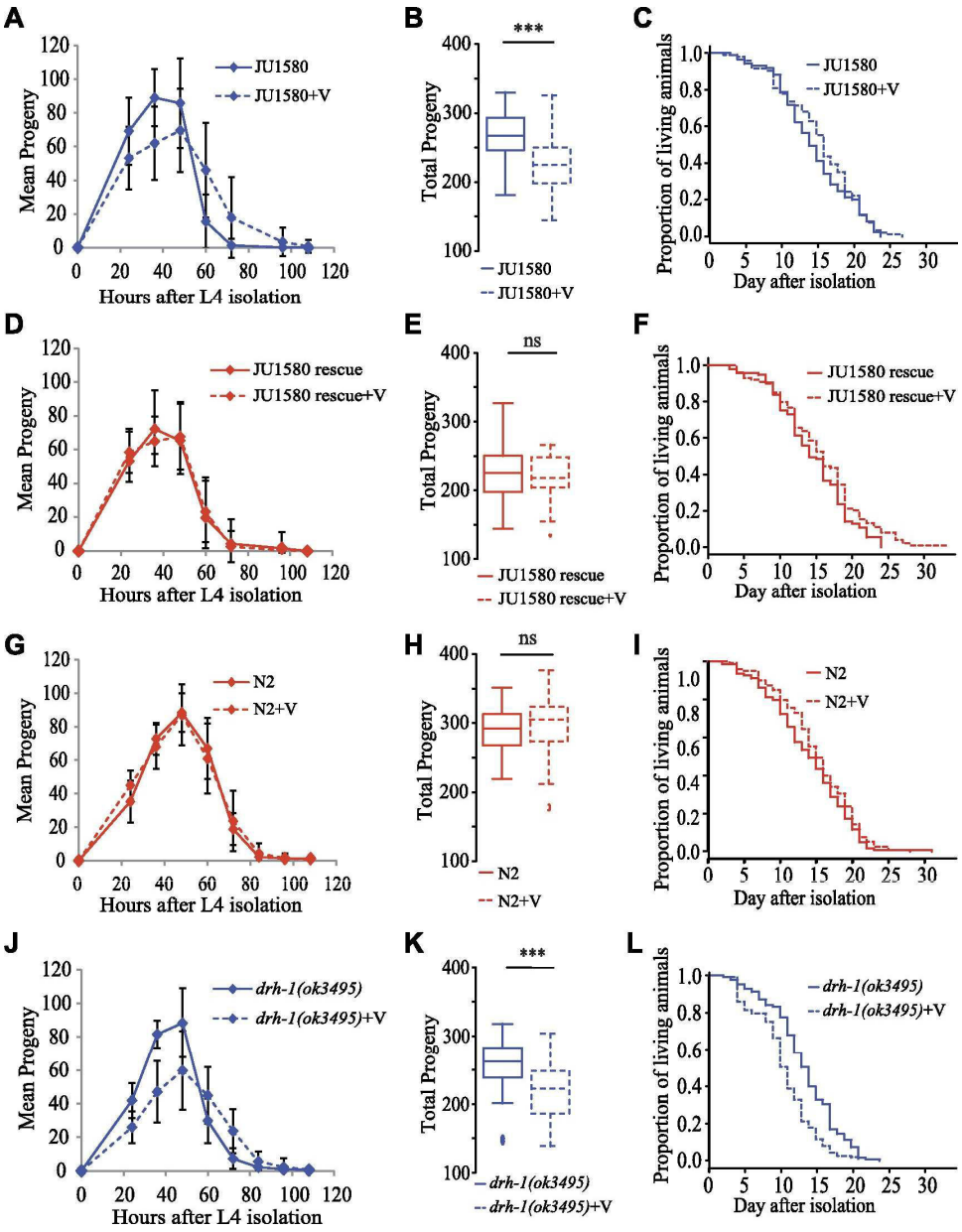


Figure 2—figure supplement 1. Infection by the Orsay virus has an effect on progeny production of *drh-1* deleted strains and on longevity of the *drh-1(ok3495)* mutant.
(A) Dynamics of progeny production of JU1580 in the absence or presence of the Orsay virus (n = 40 animals). (B) Total progeny production in the same experiment. (C) Survival curves in the absence or presence of virus (n = 130 animals). (D–F) Idem with the SX2377 strain (JU1580 rescue). (G–I) Idem with the N2 strain. (J–L) Idem with the RB2519 strain (the *rde-1(ok3495)* mutant). JU1580 animals and *drh-1(ok3495)* mutants show a significant delay in progeny production in the presence of the virus (linear model, p=0.049 and p=7.5 × 10–4, respectively), as well as a decrease in total progeny production (Wilcoxon rank test, p=1.6 × 10–4 and p=1.2 × 10–4, respectively). The N2 and SX2377 strains with the intact drh-1 gene do not show this viral sensitivity. Concerning longevity, the *drh-1(ok3495)* mutant has a significantly reduced lifespan in the presence of the virus (logrank test, p=5.2 × 10–4), but the JU1580 isolate does not (p=0.85). Strains only carrying deleted versions of drh-1 are represented in blue while strains carrying the N2 *drh-1* are represented in red. ***: p<0.001.

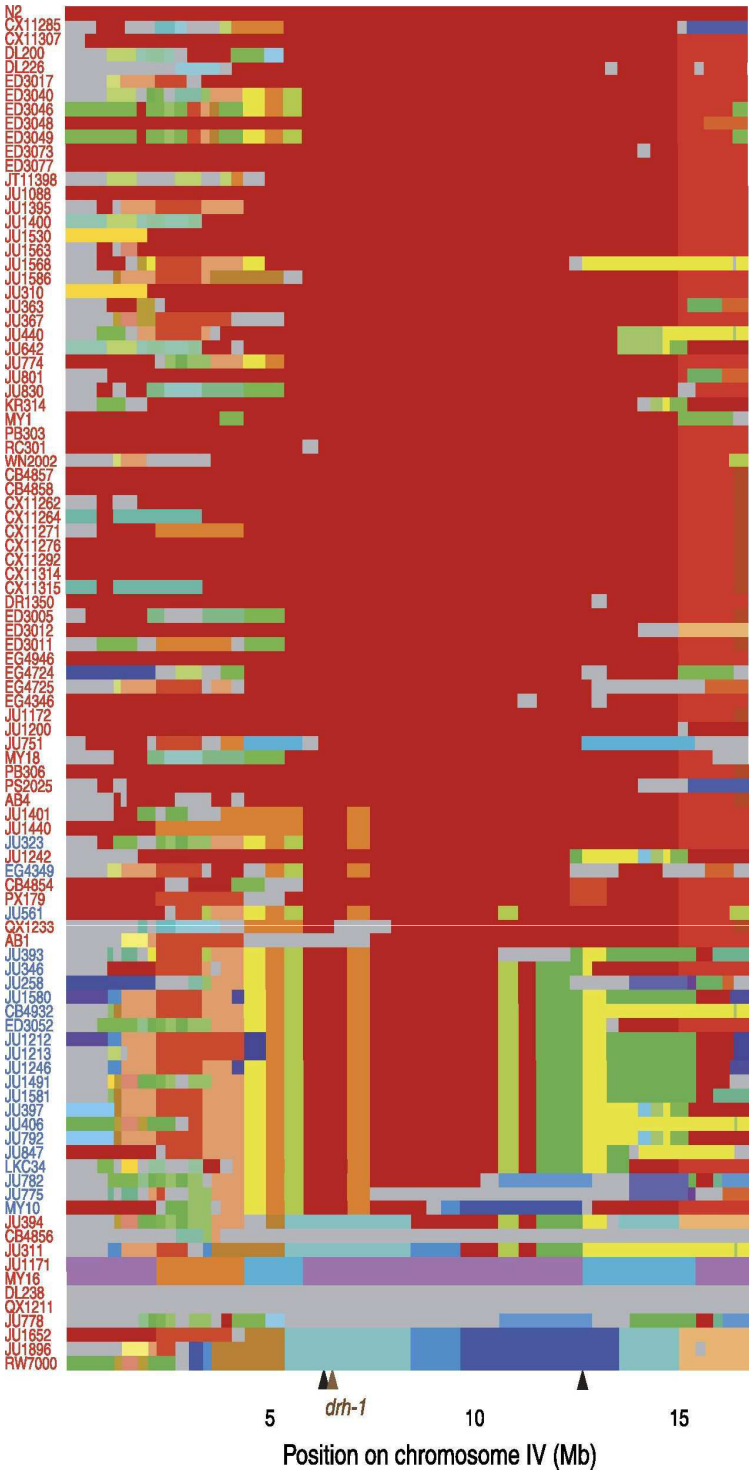


Figure 2—figure supplement 2. Chromosome IV haplotypes for the 97 isolates (modified from Supplemental Figure 7 in Andersen et al., 2012). Each row represents one of the 96 isolates, ordered from the less divergent (top) to the most divergent (bottom) from N2 for the region associated with Orsay virus sensitivity (IV:6,388,961 to IV:12,408,993; between black arrows). All clones that share a given region of a chromosome are shown with the same color in that region (N2 is in red). Haplotypes unique to a single isolate are colored in gray. Isolates in blue or red carry the *niDf250* or N2 allele of *drh-1*, respectively.

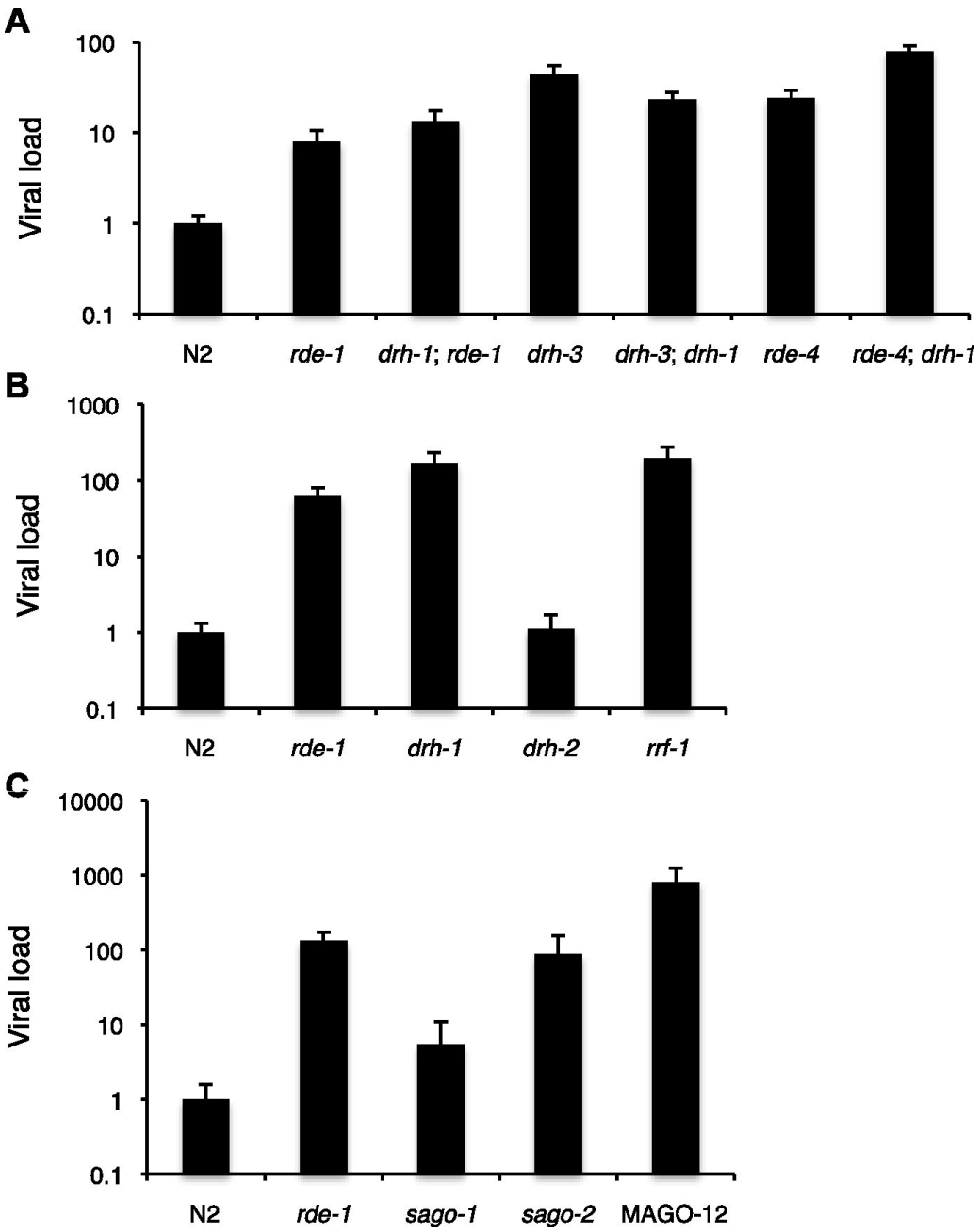


Figure 3—figure supplement 1. Viral sensitivity in a number of mutants of small RNA pathway genes. (A–C) RT-qPCR analysis of viral levels after 4 days of infection with the Orsay virus. Different panels refer to different sets of experiments using different stocks of virus. Note that sago-2 appears to be highly sensitive to viral infection.

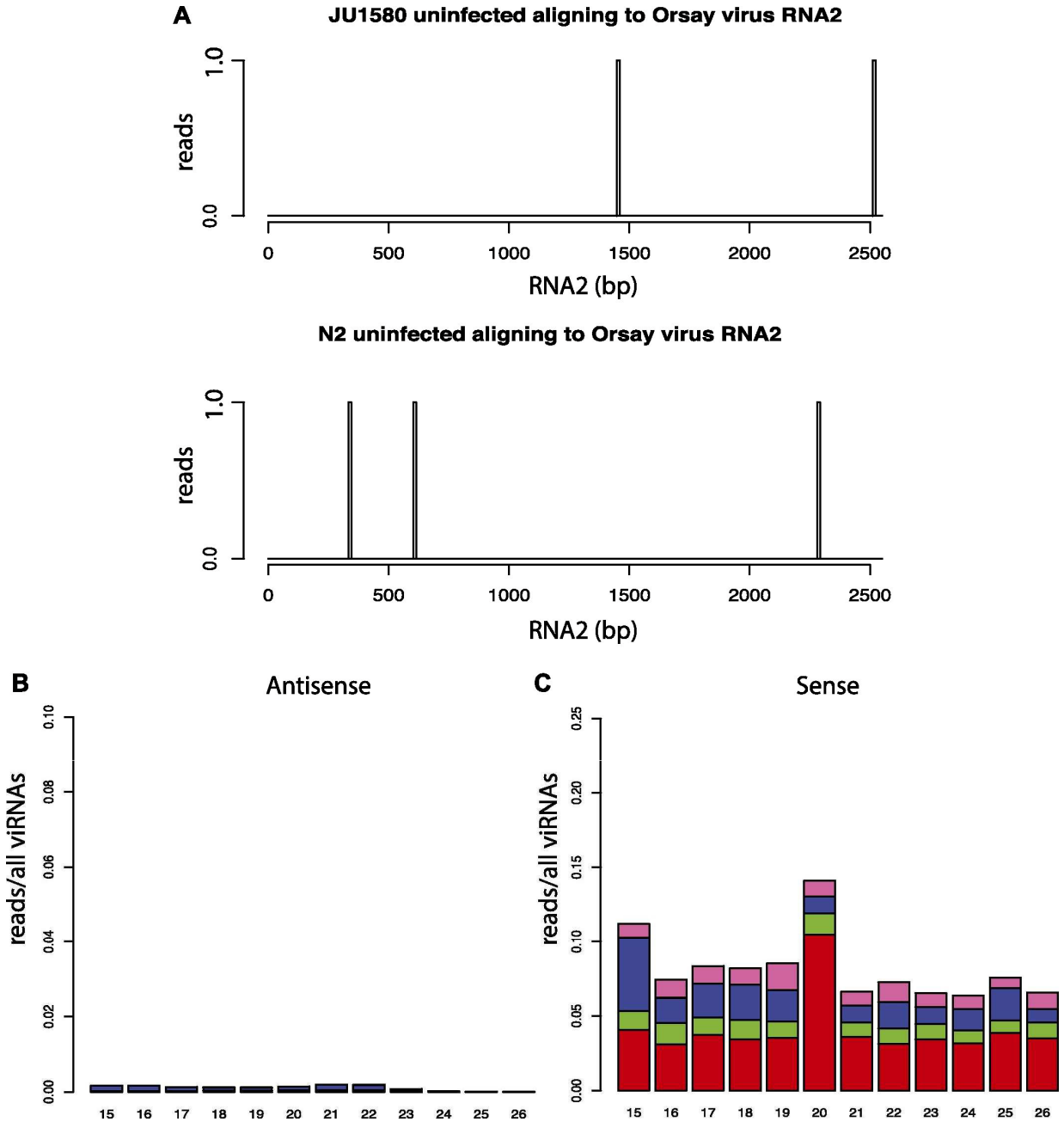


Figure 3—figure supplement 2. Additional small RNA sequencing controls.
(A) Virtually no small RNAs in libraries prepared from uninfected animals align to the viral genome. All reads aligning to RNA2 of the Orsay genome with up to one mismatch are shown. (B and C) 5′ dependent small RNA sequencing of infected *drh-1; rde-4* double mutants (compare to Figure 4K).

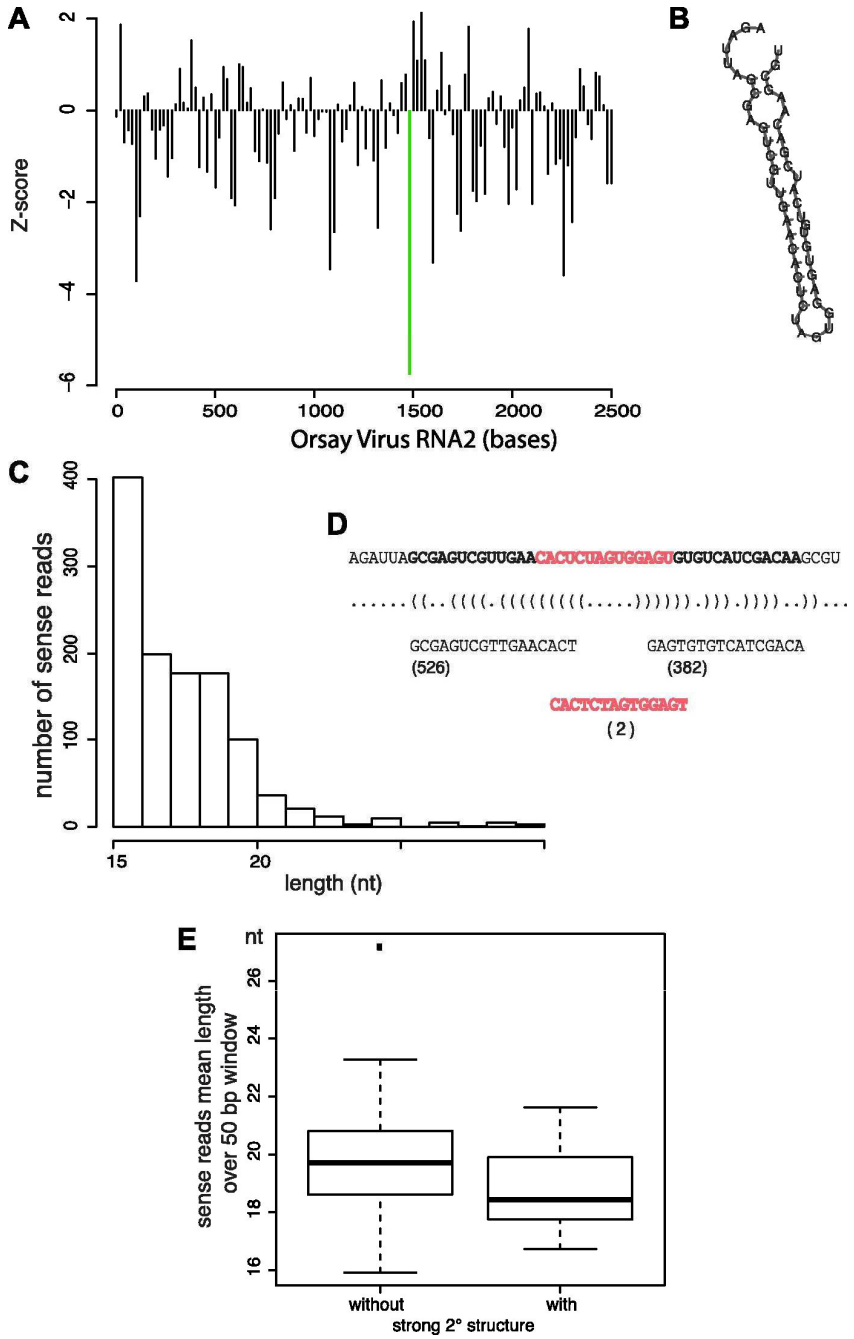


Figure 3—figure supplement 3. Relationship between antiviral small RNAs and predicted secondary structure within the viral genome. (A) Profile of the strength of predicted secondary structures for each 50 bp sliding window in 20 bp steps across the viral genome, as measured by the Z-score of the RNA-fold free energy compared to 100 random shuffles of the sequence. Thus negative values represent a stronger structure (more negative free energy) than random. The region with the strongest predicted structure is highlighted by a green bar. (B) Predicted minimum free energy structure of the region highlighted in (A). (C) Distribution of read lengths for JU1580 small RNAs mapping to the viral genome anywhere within this region. (D) Detailed analysis of the reads mapping within the selected region indicating the number of reads that come from the predicted arms and loop of the hairpin shown in (B). (E) Boxplot of read lengths for JU1580 small RNAs mapping either to regions with weak predicted secondary structures, or to regions with strong predicted secondary structures. The difference between unstructured or structured regions is statistically significant (p=0.03, Wilcoxon unpaired test).

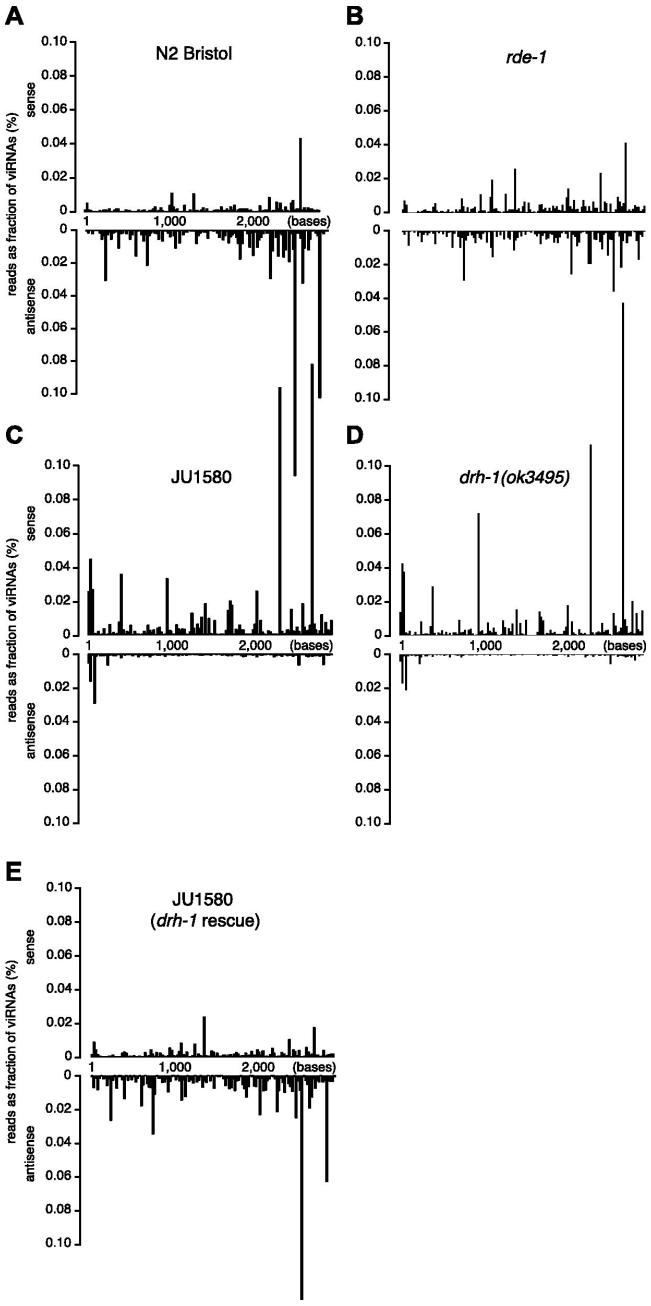


Figure 4—figure supplement 1. Distribution of viRNAs along the Orsay genome. (A–E) Small RNAs (primary and secondary siRNAs from 5′ independent sequencing) were mapped to Orsay RNA2 genomic sequence in strains indicated. JU1580 (*drh-1* rescue) in panel (E) refers to SX2375.

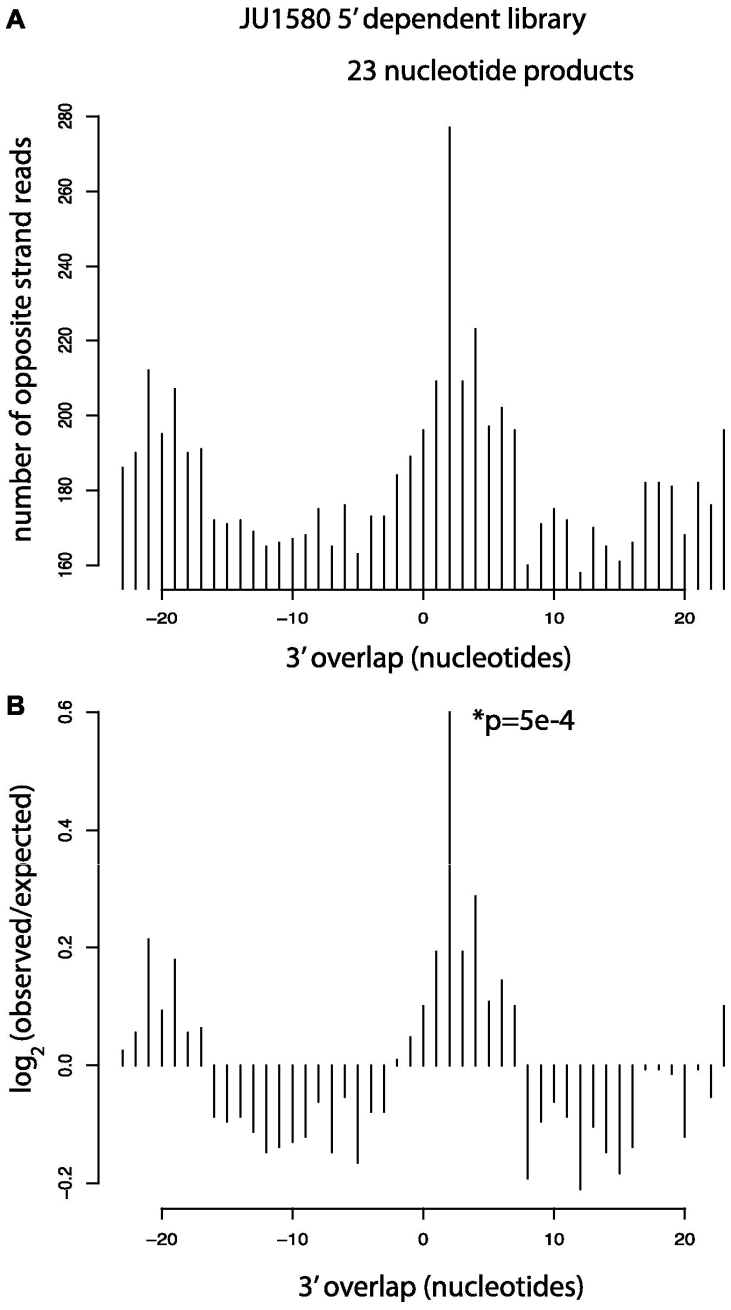


Figure 4—figure supplement 2. Analysis of residual Dicer products in JU1580 mutants. (A) Number of 23 nucleotide long reads showing 3´ overlap as indicated on the x axis. (B) log2 observed/expected showing enrichment for a 2 nucleotide 3´ overlap. The p value is for a $\chi 2$ test against a uniform distribution.

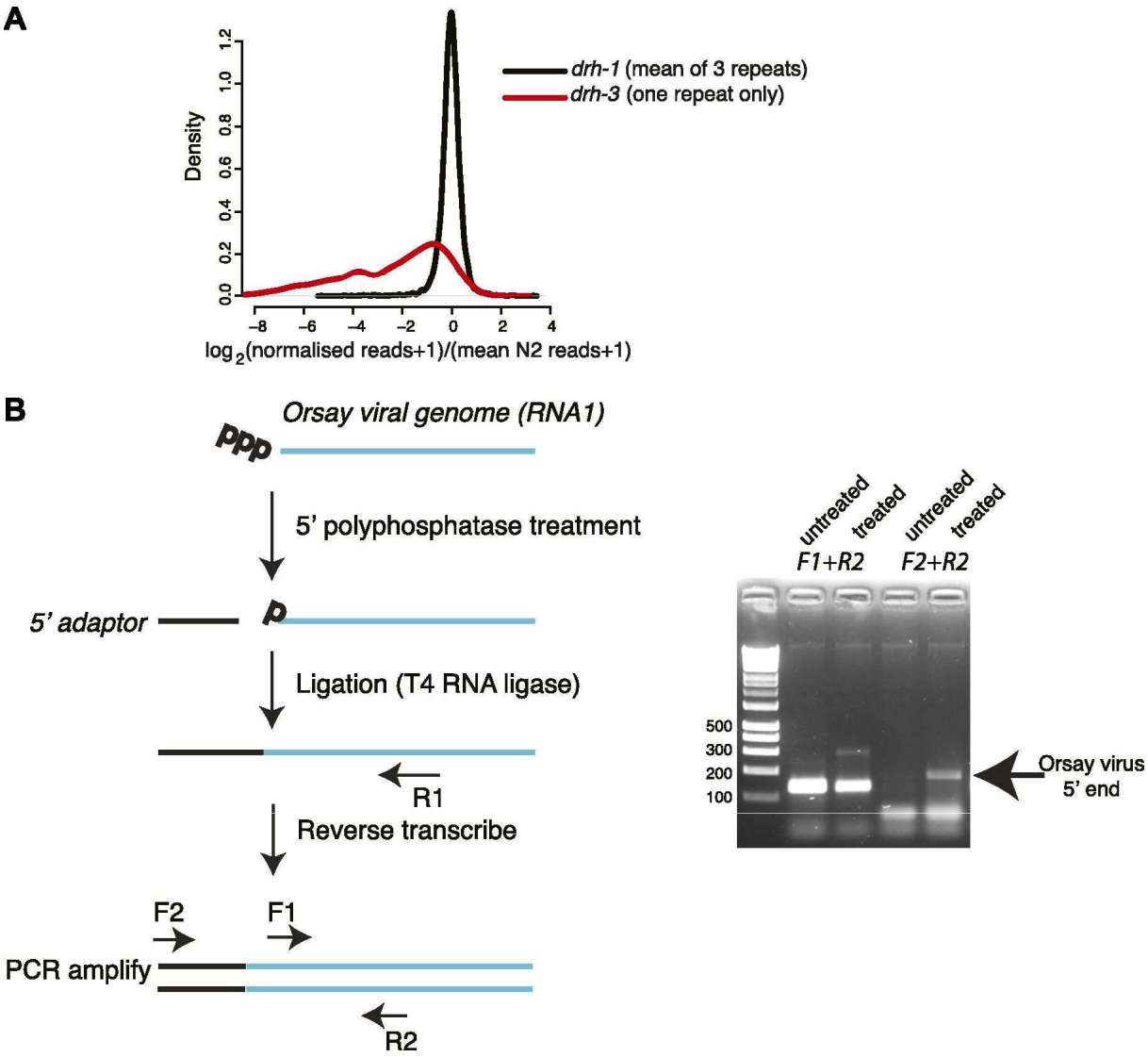


Figure 4—figure supplement 3. Analysis of 22G-RNAs mapping to endogenous loci. (A). Endogenous 22G secondary siRNAs mapping to genes in *drh-1* and *drh-3* mutants compared to wild-type (N2) animals. All 22G siRNAs mapping to non-overlapping protein-coding genes for which at least 10 reads could be detected in at least one repeat of N2 sequencing are shown. Reads for each protein-coding gene were scaled so that each library was directly comparable by multiplying the number of reads by the ratio of the total size of the library to the size of the largest N2 library. 1 read was then added to allow the log2 of the number of reads to be taken. The density plot shows the distribution of differences of normalized reads averaged over 3 repeats of each of *drh-1* mutants and N2. The distribution of differences of drh-3 to N2 obtained using the same methods is shown for comparison. (B) 5´ RACE analysis of the Orsay viral genome. The left-hand panel shows the method used to asses whether the 5´ end of the virus carries a triphosphate or not. Primer pairs used for PCR are shown—F1 is a positive control annealing within the viral sequence and F2 is adaptor specific. The right-hand panel shows amplification with the primer pairs as indicated in the left-hand panel.

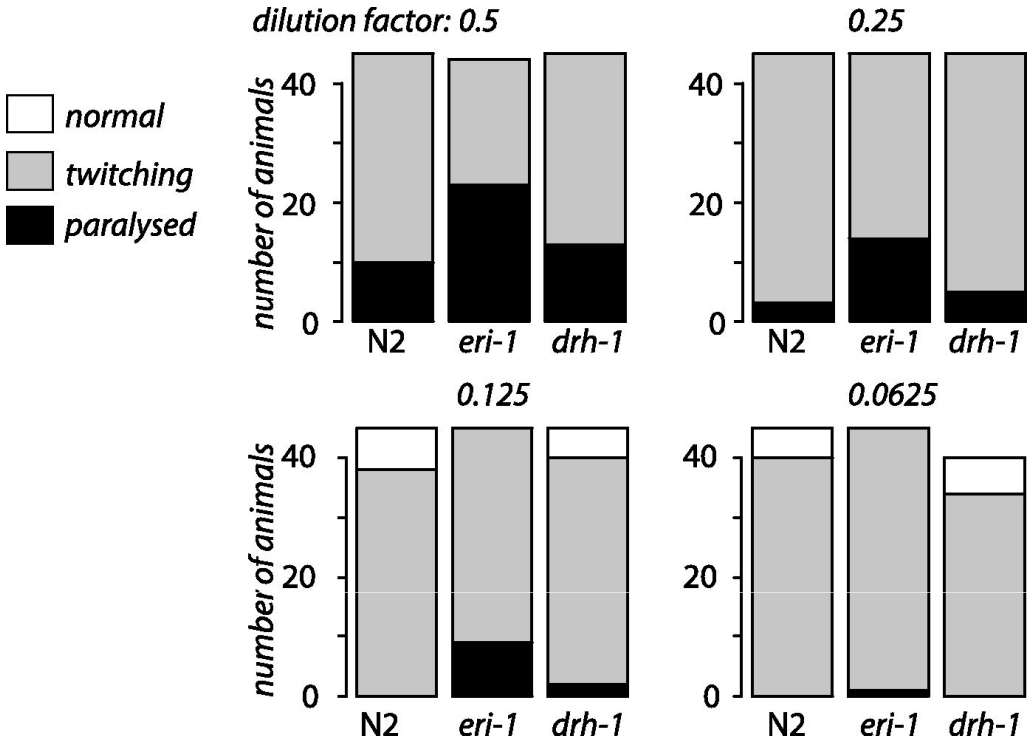


Figure 4—figure supplement 4. *drh-1* mutants are not hypersensitive to RNAi. The bar chart shows the number of L4 stage animals displaying either paralysis, twitching or no phenotype after feeding on RNAi *E. coli* for four days is shown for serial dilutions of *unc-22* RNAi *E. coli* with E. coli expressing an empty vector control. No twitching was observed in any strain when the empty vector RNAi alone was used. We observed enhanced RNAi in *eri-1* vs N2 ($p<1 \times 10–4$, Fisher's exact test) for every dilution but no difference N2 vs *drh-1* ($p<0.1$, Fisher's exact test) at any dilution.

DISCUSSION

Deux principaux résultats ressortent de mes travaux de thèse. Dans un premier temps, ils ont permis d'identifier et caractériser les premières infections virales naturelles du genre *Caenorhabditis*. Dans un second temps j'ai pu identifier les bases génétiques de la variabilité dans la capacité du virus d'Orsay à se répliquer dans son hôte. Je discuterai ici de l'apport de ces travaux dans la compréhension des interactions hôte-parasite et dans la compréhension de la défense antivirale chez *C. elegans*.

I. Les premières infections virales naturelles de *Caenorhabditis* : implications

I.A. Un nouveau groupe de virus proche des nodavirus

I.A.1. Structure des virus de Caenorhabditis

Les virus d'Orsay et de Santeuil possèdent une organisation génomique similaire aux *Nodaviridae* (groupe IV dans la classification de Baltimore) retrouvés chez les insectes et les poissons. Comme chez les nodavirus, le génome des deux virus de *Caenorhabditis* est composé de deux ARN simple brin positif. L'ARN1 code pour une Polymérase à ARN dépendante de l'ARN (RDRP) et l'ARN2 code pour les protéines de la capside (ORFα). Ils diffèrent des nodavirus connus en deux points : i) aucun ARN3 subgénomique de l'ARN1 codant pour une protéine inhibitrice de la voie des siARN antiviraux comme pour le Flock House virus (FHV) n'a été détecté et ii) un nouveau cadre de lecture inédit, nommé ORFδ, a été identifié sur l'ARN2 (Félix et al., 2011)Figure 3A). Cet ORFδ ne possède pas d'homologue connu et il ne semble pas coder pour une protéine inhibitrice de l'ARN interférence (Guo et Lu, 2013). Des études ultérieures sur le virus d'Orsay ont montré que le produit de cet ORFδ se retrouve toujours fusionné à la protéine de capside α par un phénomène de décalage de cadre de lecture provoqué par la présence d'une séquence « glissante » conservée entre les virus d'Orsay et de Santeuil située juste après le codon stop de l'ORFα suivie, 8 nucléotides plus loin, d'une structure en tige-boucle (Jiang et al., 2014). Cette séquence peut provoquer dans certains cas le décalage du ribosome lors de la traduction qui va déterminer un nouveau cadre de lecture pour la suite de la traduction de l'ARN2 en ignorant le codon stop du cadre de lecture précédent. De plus, elle apparaît conservée entre les virus, suggérant une fonction essentielle de celle-ci. La protéine δ est très probablement un élément de la capside dont la fonction n'a

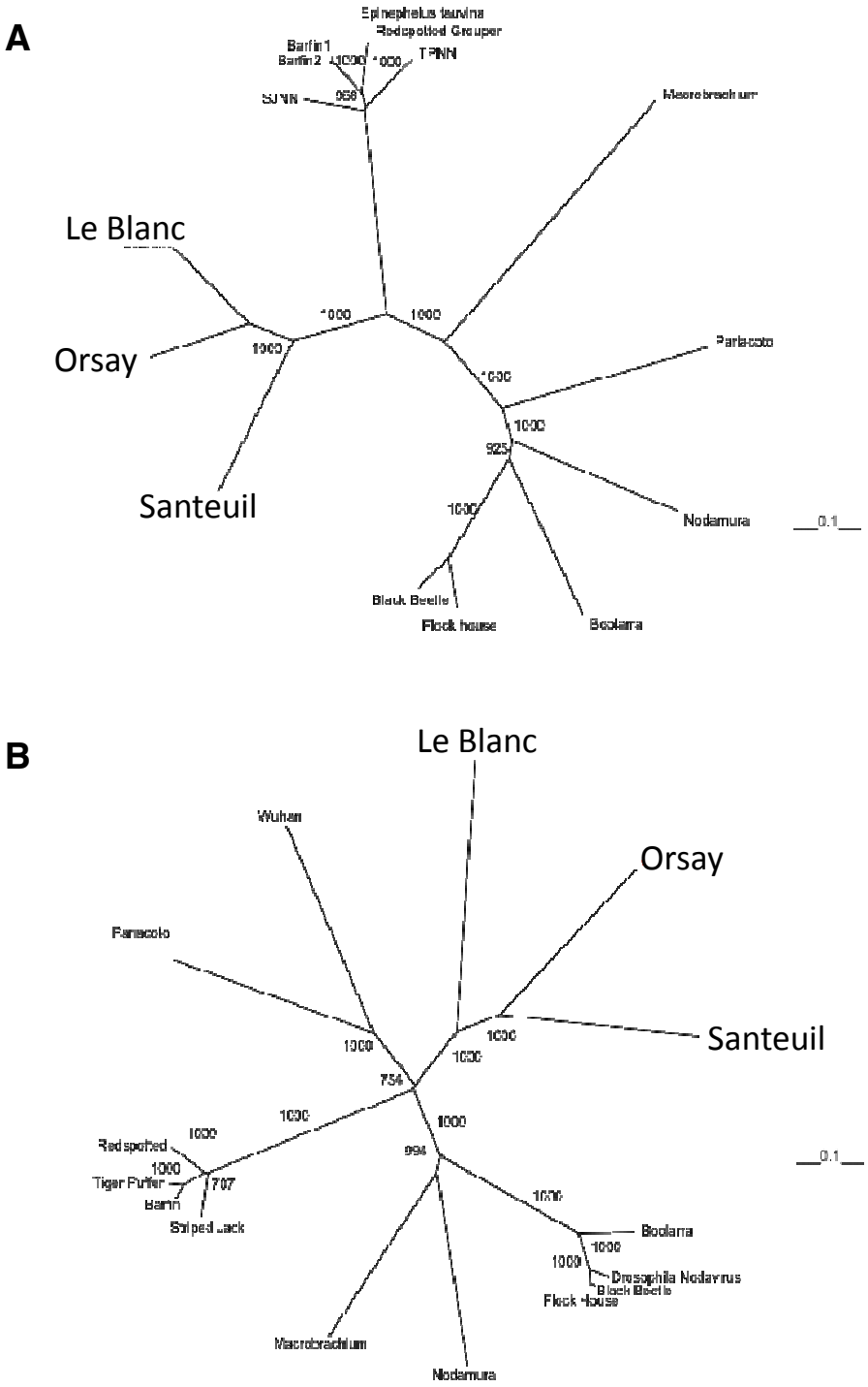


Figure 29. Analyse phylogénétique des virus de *Caenorhabditis* des Nodaviridae connus selon la méthode « neighbour-joining). A: Analyse à partir des séquences protéiques prédites des RDRP. **B:** Analyse à partir des séquences protéiques prédites de la capside La longueur des branches est proportionnelle au nombre de substitutions selon l'échelle indiquée. Figure obtenue suite à des communications personnelles avec le laboratoire de David Wang.

pas encore été identifiée. Une fonction envisageable pourrait être la fixation à un récepteur membranaire de la cellule intestinale, permettant ainsi au virus d'entrer dans la cellule intestinale. Même si l'on sait que la protéine δ seule n'a pas d'effet suppresseur de l'ARN interférence de l'hôte, on ne peut pas exclure un tel rôle pour la fusion capside α – domaine δ. L'introduction d'un codon stop dans la partie δ de la protéine capside α – protéine δ abolit la capacité d'infection du virus d'Orsay (David Wang, communication personnelle), suggérant un rôle essentiel de cette partie.

I.A.2. Relation phylogénétique avec les nodavirus

Une analyse phylogénétique réalisée à partir des séquences protéiques de ces virus révèle que les deux virus de *Caenorhabditis* sont plus proches entre eux qu'avec les autres nodavirus, suggérant probablement un nouveau groupe de nodavirus (Félix et al., 2011)Figures 3B et 3C). Depuis l'identification de ces deux premiers virus, un troisième virus de *Caenorhabditis briggsae* du même type a été séquencé et nommé virus du Blanc (Franz et al., 2012). Il possède la même organisation que les deux autres virus et ne semble pas phylogénétiquement plus proche du virus de Santeuil que du virus d'Orsay : la RDRP prédite du virus du Blanc est plus proche de la RDRP prédite du virus d'Orsay d'un point de vue des acides aminés tandis que la séquence protéique prédite de la capside du virus du Blanc est plus éloignée des deux autres virus qu'ils ne le sont entre eux (Figure 29).

I.A.3. Phénotypes liés à l'infection

I.A.3.α. Une désorganisation des cellules intestinales

Les infections virales ont été initialement identifiées par la présence de phénotypes anormaux des cellules intestinales chez JU1580 (*C. elegans* porteur du virus d'Orsay) et JU1264 (*C. briggsae* porteur du virus de Santeuil). Ces deux virus, ainsi que celui du Blanc, provoquent les même phénotypes quoiqu'un peu plus marqués chez les *C. briggsae* infectés par leurs virus : une densité anormalement faible de granules intestinaux, une élongation ou une disparition des noyaux, un cytoplasme très fluide et, dans certains cas, une fusion de plusieurs cellules intestinales (Félix et al., 2011)Figure 1). Cependant les mécanismes moléculaires et cellulaires mis en place lors de ces infections virales et aboutissant à ces phénotypes doivent être encore élucidés.

Des marquages par immunofluorescence de la capside et de la RDRP ont permis de localiser ces protéines virales durant une infection. Les protéines RDRP sont présentes dans le cytoplasme tandis que les protéines de la capside se concentrent au niveau de la

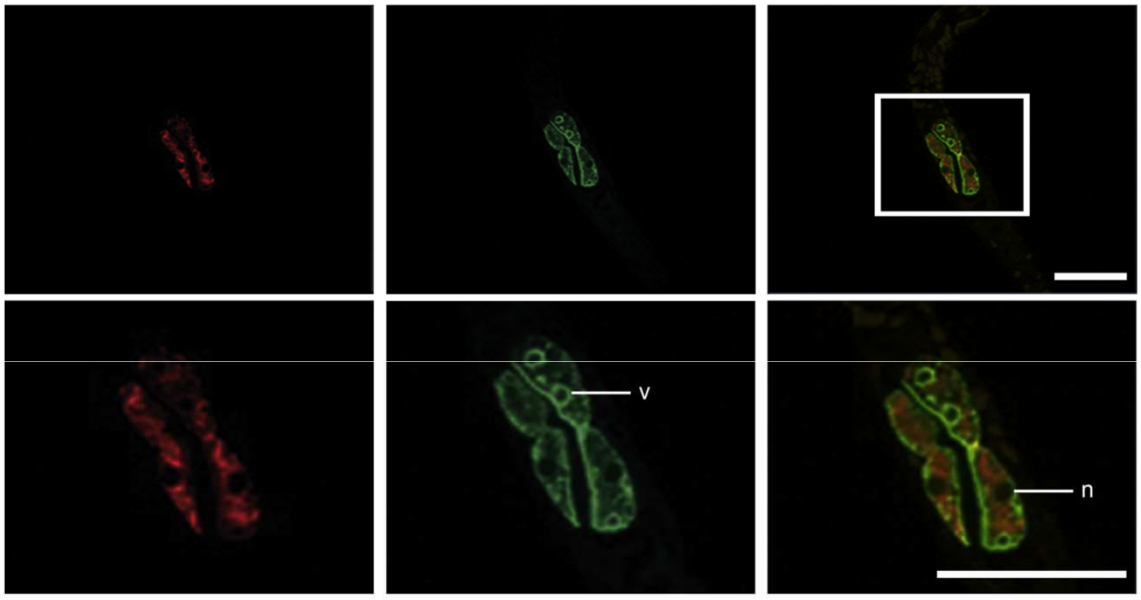


Figure 30. Localisation subcellulaire des protéines codées par le virus d'Orsay en microcsopie à confoncal lors d'une infection de JU1580. Ces protéines sont marquées par des anticorps anti-RDRP (rouge) et anti-capside (vert). n: noyau; v: vésicule. La barre d'échelle représente 20 µm. Images issues de Franz et al. (2014).

membrane apicale de la cellule intestinale, de larges vésicules cytoplasmiques et semblent s'agréger en certains points du cytoplasme (Franz et al., 2014) (Figure 30). Les signaux de la RDRP et de la capside α ne se chevauchent jamais, suggérant une séparation stricte des différentes étapes du cycle viral. Alors que les mécanismes moléculaires permettant la réplication de ces virus restent encore inconnus, la réplication semble s'établir au niveau du cytoplasme. Les capsides, quant à elles, présentent une affinité plus importante pour les membranes. De plus, les filaments intermédiaires, habituellement abondants sous la membrane plasmique, semblent exclus des zones contenant les capsides virales, suggérant une réorganisation du cytosquelette cortical probablement nécessaire à la sortie du virion. Ce type de réarrangement a été montré lors de l'infection de *C. elegans* par la microsporidie *Nematocida parisii*, seul parasite intracellulaire découvert avant les virus (Estes et al., 2011).

Il peut arriver qu'un même animal possède six cellules infectées sur les vingt que compte l'intestin (Franz et al., 2014). La plupart des cellules infectées, soit 98,7 % d'entre elles, appartiennent à la partie la plus antérieure de l'intestin. Ceci peut être dû au fait qu'elles sont les premières cellules exposées au virus et que la probabilité d'infection est plus importante pour ces cellules. On sait également que la partie antérieure et la partie postérieure de l'intestin ne participent pas aux mêmes fonctions digestives. Il est possible que les cellules antérieures possèdent des caractéristiques différentes des cellules postérieures et qu'elles soient plus perméables au virus, soit par voie passive, soit par la présence d'un récepteur membranaire auquel se fixe la capside du virus.
La dynamique d'infection à l'échelle de l'intestin n'est pas clairement définie. Lorsque plusieurs cellules sont infectées, ces cellules ont tendance à être adjacentes, suggérant une capacité du virus de passer d'une cellule à une autre. Néanmoins distinguer le passage latéral du virus à une cellule adjacente d'une nouvelle infection provenant directement du lumen n'est pas trivial (Figure 31).

I.A.3.β. Impact des virus sur les traits d'histoire de vie de leurs hôtes

Le phénotype observé lors d'une infection peut parfois être spectaculaire, notamment chez *C. briggsae*, avec une perte de structure des cellules intestinales et une fusion de celles-ci (Félix et al., 2011)Figure 1I). L'intestin réalisant des nombreuses fonctions essentielles chez ces *Caenorhabditis*, la perte ou la dégradation de cet organe devrait diminuer drastiquement la viabilité des organismes infectés. Pourtant, en conditions de laboratoire, la démographie des populations infectées ne semble pas, à première vue, très différente de celle des populations non infectées. L'infection des souches originellement sensibles au virus (JU1580 pour le virus d'Orsay et JU1264 pour le virus de Santeuil) retarde le pic maximal de ponte de manière significative mais n'affecte pas particulièrement leur durée de vie (Félix et al., 2011); Figure S3; Ashe, Bélicard, Le Pen,

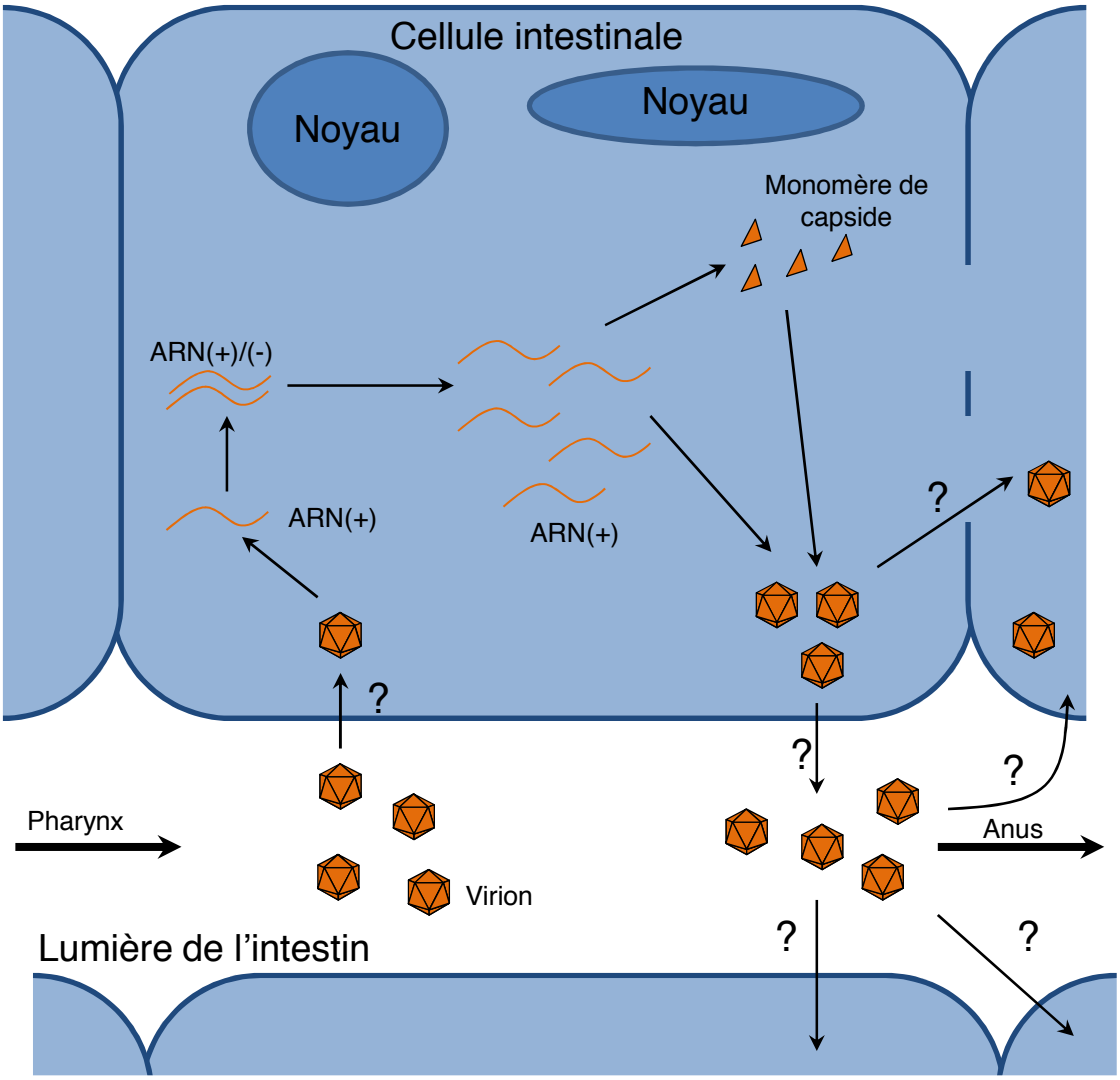


Figure 31. Cycle de vie des virus de *Caenorhabditis* tel qu'il est connu aujourd'hui. Le virion pénètre dans l'animal par le tube digestif. Par un processus encore indéfini, il entre dans une cellule intestinale depuis le lumen. Les ARN simple brin sont libérés dans le cytoplasme et la RDRP est produite à partir de l'ARN1. Cette RDRP permet l'amplification des ARN1 et 2 du virus probablement dans le cytoplasme. Les ARN2 sont alors traduits en protéines de capside. Ces protéines de capsides se retrouvent au niveau de la membrane apicale de la cellule ou au niveau de larges vésicules non-identifiées. Le génome est empaqueté dans la capside et le virion ainsi formé quitte la cellule par un processus non-lytic indéterminé. On ne sait pour l'instant pas si le virion peut être transmis à une cellule adjacente ou si une nouvelle infection a lieu à partir de virions présents dans la lumière de l'intestin. Une fois dans la lumières, les virions sont déféqués et se retrouvent dans le milieu extérieur, pouvant ainsi établir une nouvelle infection s'il sont ingérés par un autre animal.

Sarkies et al., 2013; Figure 2 - figure supplement 1). Ce phénomène n'a pas encore d'explication physiologique. On sait néanmoins que les cellules antérieures constituent un lieu privilégié de la sécrétion de protéines chargées de digérer les aliments dans la lumière de l'intestin. On peut penser qu'une infection de cette partie de l'intestin perturbe l'assimilation de nutriments en quantité suffisante pour la production d'ovocytes, retardant ainsi la production maximale de descendants. Cette perturbation peut être liée à plusieurs phénomènes non exclusifs : i) le virus utilise les ressources de l'hôte pour sa propre multiplication, réduisant ainsi les ressources disponibles pour le fonctionnement de l'organisme hôte ; ii) l'infection par le virus endommage les cellules qui ne peuvent plus exercer leur rôle dans la digestion ; iii) une défense antivirale, même si inefficace, se met en place dans ces cellules et consomme des ressources initialement dédiées au fonctionnement de la cellule. Une dernière hypothèse pourrait venir de la présence de virus dans la gonade somatique (Félix et al., 2011)Figure 4E) qui perturberait directement le processus d'ovogenèse et ainsi réduirait son efficacité. Néanmoins, d'autres expériences de FISH ont été réalisées par la suite et il ne semble pas que l'ARN1 ni l'ARN2 des virus se retrouvent dans la gonade somatique (Franz et al., 2014), suggérant soit l'existence d'une barrière empêchant le virus d'entrer dans les cellules germinales, soit l'existence d'une réponse antivirale très efficace à l'infection de ces cellules.

Un report du pic de ponte peut avoir un effet important dans la nature puisqu'un hermaphrodite infecté doit survivre plus longtemps pour pondre la même quantité d'œufs qu'un hermaphrodite non infecté. Une expérience de compétition entre une souche de *C. elegans* sensible au virus d'Orsay et une autre résistante révèle que la souche sensible est nettement désavantagée en présence de virus alors qu'elle se maintient à une fréquence de 50% dans la population en absence de celui-ci (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2B).

I.B. Les *Caenorhabditis* et leur virus : un modèle pour l'évolution des interactions animal-virus

Depuis une vingtaine d'années, *C. elegans* est devenu un modèle d'étude du système immunitaire inné. Certains parasites de l'Homme ou de plantes sont capables d'induire un phénotype chez *C. elegans*. Bien évidemment, la plupart du temps, ces phénotypes ne peuvent pas être identiques à ceux observés lors de l'infection de leur hôte original. En général, les bactéries pathogènes pour l'Homme colonisent l'intestin du nématode mais ne pénètrent pas à l'intérieur des cellules intestinales, suggérant soit une défense efficace

de l'hôte, soit une inefficacité du parasite à pénétrer dans un hôte qui n'est pas son hôte naturel. Parmi ces parasites, seul un nombre limité de bactéries sont capables d'établir une colonisation persistante de la lumière de l'intestin (*Salmonella enterica, Serratia marcescens, Enterococcus faecalis, Pseudomonas aeruginosa* ou encore *Staphylococcus aureus*) (Sifri et al., 2005). Le problème de ces études réside dans le fait que les mécanismes de défense révélés ne sont pas le résultat de l'évolution entre l'hôte et ce parasite, puisque *C. elegans* n'a probablement jamais rencontré ce parasite auparavant.
La découverte d'un parasite naturel, tel que le virus d'Orsay pour *C. elegans*, offre de nouvelles opportunités pour comprendre l'évolution des interactions entre un hôte et son parasite. Alors que l'évolution des interactions hôte-virus est très étudiée pour les systèmes bactéries-bactériophages ou les systèmes plantes-phytovirus, il n'existe pas, à l'heure actuelle, de modèle simple d'utilisation permettant ces études chez les animaux. Certains virus infectant naturellement *D. melanogaster* ont été identifiés. Cependant, la plupart de ces virus sont asymptomatiques (Drosophila X Virus, Nora Virus) ou induisent des phénotypes dans des conditions rarement rencontrées dans la nature (létalité pour une exposition à une forte dose de CO_2 pour le Drosophila C Virus ou le Sigma virus). De plus, dans la plupart des cas, ces virus nécessitent d'être injectés dans leur hôte, ce qui est peu pratique. L'intérêt du système *Caenorhabditis elegans*-virus d'Orsay réside dans plusieurs points:

- Les symptômes, identifiables au microscope, nous permettent de détecter facilement une infection dans une population.
- L'hermaphroditisme de *C. elegans* permet des études sur des génotypes homozygotes contrôlés à la mutation près.
- L'infection est simple : ajout d'une solution virale sur la population devant être étudiée et transfert horizontal du virus.
- L'infection peut être supprimée en javellisant une population de *Caenorhabditis* adultes et en récupérant ainsi des descendants non infectés.
- Le temps de génération faible permet la mise en place d'expérience de coévolution au laboratoire.
- La souche contenant le virus peut être congelée, permettant ainsi de garder une trace lors d'expériences d'évolution expérimentale à la fois du virus et de son hôte.
- Le génome de *C. elegans* est assemblé et très bien annoté. De plus, de nombreux mutants sont disponibles et l'utilisation de l'ARN interférence est très développée.

Ces caractéristiques en font un système très pratique pour l'étude de l'évolution des interactions hôte-virus en laboratoire. Comme annoncé précédemment, une meilleure compréhension du fonctionnement du virus serait utile pour un meilleur développement du système. D'autres problèmes, parfois limitants, nous empêchent d'exploiter le

système à son plein potentiel. Par exemple, nous n'avons pour le moment pas moyen de contrôler la quantité de virus initialement utilisé pour inoculer les animaux si ce n'est par la dilution des préparations virales, peu pratique à effectuer.

Lise Frézal, en post-doctorat au laboratoire, a isolé récemment (2013) un nouveau variant du virus d'Orsay à Ivry, qui semble plus virulent que le premier variant isolé. Ceci ouvre de nombreuses portes vers l'étude des interactions entre génotype de l'hôte. Il serait notamment intéressant de voir si chacun des 97 isolats naturels de *C. elegans* réagit de manière similaire à l'infection par ces deux variants. Si cela n'est pas le cas, il serait alors possible d'identifier les bases de des interactions entre les génotypes des hôtes et ceux des virus. L'obtention de nouveaux variants serait également idéale pour mener des expériences de co-évolution au laboratoire.

Contrairement au virus d'Orsay, de nombreux variants du virus de Santeuil et du virus du Blanc ont été isolés et séquencés par Lise Frézal, rendant le type d'expérience précédemment envisagé possible. Pour le moment, Gautier Brésard, étudiant en thèse, a identifié un isolat de *C. briggsae*, HK104, résistant au virus du Blanc mais sensible au virus de Santeuil, suggérant une résistance ou une sensibilité spécifique à l'un de ces deux virus. Ses travaux actuels consistent à identifier la base génétique de cette résistance spécifique à l'aide de lignées recombinantes.

I.C. Un rôle majeur des petits ARN dans la lutte contre les infections virales

I.C.1. La voie siARN amorcée par DRH-1 constitue une réponse antivirale chez C. elegans*.*

En ce qui concerne les virus, seul le Vesicular Stomatitis Virus était capable d'infecter *C. elegans* (infection limitée à des cellules embryonnaires, l'infection de l'organisme étant impossible) (Schott et al., 2005; Wilkins et al., 2005). Ces études, associées à celles effectuées sur un transgène du FHV (Lu et al., 2005; Lu et al., 2009), ont été capables d'identifier le mécanisme d'ARN interférence comme une réponse antivirale. Nous avons pu confirmer cette implication de l'ARN interférence dans la réponse antivirale dans les deux études menées lors de ma thèse (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Félix et al., 2011). Nous avons également pu confirmer le rôle de *drh-1* dans cette voie. (Lu et al., 2009) ont montré que la quantité de réplicons du FHV est augmentée et aussi imortante dans les animaux simple mutants pour *rde-1*, *rde-4* et *drh-1*, suggérant un rôle clé de ces protéines dans la réponse immuniatire. Toujours d'après cet étude, il est

apparu cependant que la quantité de petits ARN viraux etait affectée dans le mutant *rde-4* mais pas dans les deux autres mutants, suggérant une intervention des protéines DRH-1 et RDE-1 en aval de la production des petits ARN antiviraux. Néanmoins, lorsque nous observons le profil de production de petits ARN dans notre étude, il apparait que DRH-1 agit en amont de cette production d'ARN antiviraux, probablement comme récepteur intracellulaire du génome viral et déclencheur de la voie des petits ARN antiviraux (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 4), en accord avec son appartenance au complexe incluant la protéine DCR-1. Cette contradiction peut s'expliquer par le fait qu'il n'y a pas d'abolition totale de la production d'ARN antiviraux dans le mutant *drh-1* (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 4). On peut imaginer que l'absence de protéines DRH-1 fonctionnelles permet une réplication infinie du réplicon mais aussi une transcription importante du transgène dans les cellules de l'animal après un choc thermique. On peut ainsi penser qu'un mécanisme indépendant de *drh-1*, et donc non spécifique au virus, de production de petits ARN se met alors en place.

Ces résultats lèvent une partie du voile concernant le rôle de l'ARN interférence chez *C. elegans*. En laboratoire, il est courant de nourrir les nématodes avec des bactéries produisant des ARN doubles brins spécifiques d'un gène dont on souhaite éteindre l'expression (Fire et al., 1998b). L'intérêt de cette méthode est indiscutable en laboratoire. En revanche, l'intérêt de la voie des petits ARN exogènes (l'ARN servant de patron n'est pas synthétisé par l'organisme lui-même) dans la nature n'est pas élucidé chez *C. elegans*. De manière surprenante, il apparaît que la voie exogène de l'extinction de l'expression des gènes du nématode par ARN interférence est fonctionnelle dans le mutant *drh-1*. De même, une mutation dans le gène *drh-1* n'abolit pas la production de microARN, suggérant que *drh-1* n'est pas impliqué dans la voie des ARN endogènes (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 4 - figure supplement 3). La réponse siARN déclenchée par *drh-1* est par conséquent spécifique des infections virales (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 5). Néanmoins, ces trois voies des petits ARN partagent des éléments communs et la seule différence entre la voie des ARN exogènes et celle des petits ARN viraux réside dans l'intervention de DRH-1 dans cette dernière. Cette spécificité provient probablement de la capacité de la protéine DRH-1 à reconnaître certains ARN. Elle pourrait être due à une structure particulière de l'ARN viral et/ou à la présence de triphosphate en extrémité 5' du virus (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 4 - figure supplement 3). Même si pour le moment, l'interaction entre DRH-1 et les ARN viraux à l'échelle moléculaire n'est pas totalement comprise, on sait désormais que cette voie peut servir à amorcer une réponse spécifique contre les virus à ARN chez *C. elegans*.

La souche JU1580 présente également la particularité d'être incompétente pour la voie exogène des petits ARN (Félix et al., 2011). Ainsi, il est impossible d'éteindre

l'expression d'un gène du soma en la nourrissant avec des bactéries produisant des ARN doubles brins correspondant à ce gène. La base génétique de cet incompétence n'a pas été identifiée; on sait seulement que l'injection d'une copie fonctionnelle de *drh-1* dans la souche JU1580 la rend résistante au virus (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1G), suggérant que l'incompétence de cette souche pour la voie des ARN exogènes est due à un facteur de cette voie qui n'est pas commun avec la voie des ARN viraux.

Les premières études menées sur l'ARN interférence par Craig Mello et Andrew Fire ont montré que l'injection d'un ARN double brin spécifique d'un gène particulier de *C. elegans* permettait l'extinction de ce gène dans l'individu injecté mais aussi dans la génération suivante (Fire et al., 1998b). Pour certains gènes, mais pas d'autres, cette extinction peut s'étendre sur plusieurs générations (Vastenhouw et al., 2006; Alcazar et al., 2008). Le signal qui pourrait être transmis à la génération suivante n'a pas encore été identifié. On sait cependant que la transmission de cette extinction ne requiert pas la machinerie nécessaire à la production des siARN primaires par la descendance (Grishok et al., 2000). En revanche, les protéines impliquées dans la production des siARN secondaires et celles impliquées dans l'extinction transcriptionnelle de l'expression des gènes sont indispensables à son maintien. Ces dernières, les protéines NRDE (Nuclear RNAi DEficiency) (Guang et al., 2008), sont des argonautes qui ciblent par l'intermédiaire des siARN secondaires des ARNm en cours de production dans le noyau. Elles sont capables de recruter des méthyltransférases de la lysine 3 de l'histone 9 (H3K9me) qui vont déposer des marques inhibant localement la transcription (Burton et al., 2011b). Il ne semble pas que ces méthylations soient transmises à la descendance puisque, même si elles sont transmises lors de divisions cellulaires, elles sont réinitialisées dans la lignée germinale. Une transmission des petits ARN depuis le soma vers la lignée germinale pourrait expliquer ce phénomène de transmission de l'extinction. La raison pour laquelle cette extinction multi-générationnelle de l'effet de l'ARN interférence opère avec certains gènes et pas d'autres n'est pour l'instant pas claire.
On peut donc se demander si les petits ARN antiviraux se transmettent dans la descendance des animaux mis en présence du virus d'Orsay. Autrement dit, l'immunité antivirale peut-elle se transmettre à la descendance ? (Rechavi et al., 2011) ont utilisé le système artificiel du FHV intégré au génome de *C. elegans* et exprimé sous promoteur de choc thermique pour tenter de répondre à cette question. Ce transgène est marqué par la GFP, dont l'expression témoigne directement de la capacité du réplicon produit à se répliquer. Comme observé précédemment, l'extinction du réplicon peut se transmettre en absence de la machinerie qui produit les siARN primaires dans la descendance. En revanche,

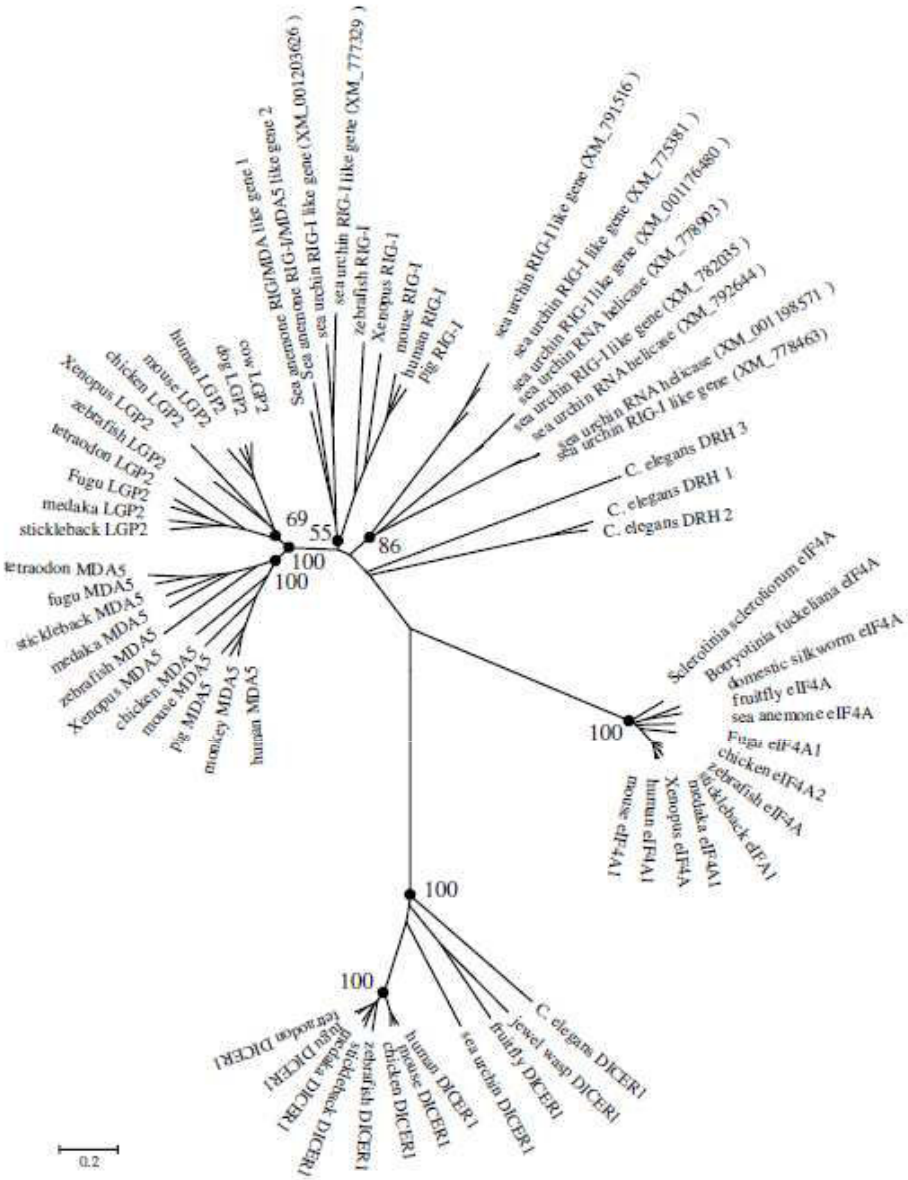


Figure 32. Arbre phylogénétique des domaines DExD/H hélicases à ARN. Cet arbre a été réalisé à partir de l'alignement séquences protéiques connues ou prédites des homologues de RIG-I, MDA-5, LGP2, eIF4A et DICER en utilisant la méthode neighbour-joining. La longueur des branches est proportionnelle au nombre de substitutions selon l'échelle indiquée. Issu de Zou et al. (2009).

cette extinction est supprimée lorque la descendance est incapable de produire des siARN secondaires. La faiblesse de cette étude est que les séquences virales sont intégrées au génome de son hôte et donc présentes dans la lignée germinale. Il est par conséquent difficile de différencier ce qui se passe lors cette infection virale artificielle de ce qui se passe lors de l'extinction d'un gène endogène du nématode. Une étude plus récente, cette fois utilisant le virus d'Orsay, indique que la transmission de la réponse antivirale a bien lieu entre un parent infecté et sa descendance immédiate (Sterken et al., 2014). Les conditions d'infection de cette étude sont trèsn différentes des nôtres : elles se font en milieu liquide et dans des conditions (concentration élevée de virus?) qui permettent d'infecter la souche standard de laboratoire N2, résistante dans la grande majorité de nos expériences, à un niveau similaire à JU1580. Leur expérience consiste ensuite à infecter N2 avec le virus d'Orsay puis de récupérer les descendants après avoir éliminé ou non toute trace de virus par javellisation des œufs. Le virus d'Orsay semble se répliquer moins dans les descendants dont les parents ont été initiallement infectés comparé à une infection dite « naïve ». Ces résultats diffèrent de ceux du laboratoire d'Eric Miska (communication personnelle) qui suggèrent qu'au contraire, la descendance de vers N2 infectés montre le même niveau d'infectabilité que des vers naïfs et ne contient pas de siARNs viraux. Les conditions de reproducibilité des expériences de Sterken et al. restent à évaluer.

I.C.2. Les RIG-I-Like Receptors chez les mammifères set les DRH chez C. elegans: *un même combat, des armes différentes.*

Les protéines DRH (DRH-1, DRH-3 et potentiellement DRH-2) chez *C. elegans* et les RIG-I-Like Receptor (RLR) des mammifères, dont font partie RIG-I, MDA5 et LGP2, partagent des domaines similaires: un domaine RIG-I C-terminal et un domaine Dead box hélicase très conservé (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1F) (Figure 32). Ces domaines permettent la reconnaissance des ARN présents dans le cytoplasme et l'engagement dans des réponses spécifiques. Les protéines RIG-I et MDA-5 reconnaissent à eux deux une multitude d'ARN viraux et déclenchent la production d'interférons et de cytokines par l'intermédiaire de deux domaines CARD (Caspase Activation et Recruitment Domain) N terminaux. La protéine LGP2 ne possède pas quant à elle de domaine CARD et est considérée comme un inhibiteur de la réponse antivirale déclenchée par les protéines RIG-I et MDA-5. Les protéines DRH-1 et DRH-3 ne possèdent pas non plus de domaines CARD et sont clairement impliquées dans la production de petits ARN : DRH-1 déclenche la production d'ARN impliqués spécifiquement dans la réponse antivirale (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 5) tandis que DRH-3 permet la production de petits ARN secondaires, quelle que

soit l'origine de l'ARN double brin initial. Le domaine RIG-I est tellement conservé entre les RLR et DRH-1 et DRH-2 qu'il est possible d'interchanger ce domaine entre les différentes protéines sans en altérer la fonction (Guo et al., 2013).

L'existence même de la protéine DRH-2 est pour le moment sujette à discussion. Le gène *drh-2* est considéré comme un pseudogène dans la base de données Wormbase. Il serait un paralogue récent de *drh-1* chez *C. elegans* puisque *C. briggsae* et *C. brenneri* ne possèdent qu'un seul homologue de *drh-1*. Il se situe entre *drh-1* et le promoteur de l'opéron auquel ils appartiennent (contenant également le gène *nstp-2*). Cependant, (Duchaine et al., 2006) auraient réussi à purifier des peptides provenant de *drh-2*. D'après (Tabara et al., 2002) , *drh-2* pourrait être transcrit en une protéine présentant 74% d'identité avec DRH-1 dans le cas d'un épissage alternatif particulier ou dans le cas où le codon d'initiation de la traduction est différent d'une méthionine. Dans cette même étude, les auteurs ont tenté d'éteindre l'expression de ces gènes par ARNi. Des expériences d'ARN interférence contre *drh-1*, *drh-2* et les deux en même temps ne provoquent pas de phénotype développemental mais semblent induire une perte de la compétence pour l'ARN interférence contre un second gène. L'implication de *drh-1* et de *drh-2* dans la voie des ARN exogènes a ensuite été réfutée par des études menées sur les souches mutantes pour ces gènes (Lu et al., 2009; Ashe, Bélicard, Le Pen, Sarkies et al., 2013). Cependant, Lu et al. ont observé une diminution de la quantité d'ARN1 et d'ARN3 provenant d'un FHV artificiel exprimé par le ver après un choc thermique dans le mutant *drh-2*, suggérant une inhibition de la voie des petits ARN antiviraux par la protéine DRH-2. Nous n'avons observé aucun changement concernant l'intensité de l'infection virale ou la production de petits ARN antiviraux chez ce mutant chez *C. elegans* (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 4). Un bon moyen de vérifier si *drh-2* est exprimé et s'il possède un rôle dans la réponse antivirale ou dans l'ARNi serait d'obtenir un double mutant *drh-2; drh-1*. L'obtention de ce double mutant à partir des deux simples mutants n'est pas envisageable au regard de la proximité physique des deux gènes. L'utilisation de nouvelles méthodes d'édition du génome, comme le CRISPR/Cas, permettrait d'obtenir une telle souche.

Une des surprises de cette étude vient du fait que nous faisons face à une sorte de fusion de deux réponses antivirales considérées jusque là comme distinctes. Ainsi, *C. elegans* a conservé la voie des petits ARN comme réponse antivirale, observée précédemment chez les plantes ou les arthropodes. Néanmoins, le récepteur amorçant la production de petits ARN chez *C. elegans* est un homologue des récepteurs qui déclenchent la réponse antivirale chez les mammifères. dans une grande partie des Eumétazoaires, notamment chez les Cnidaires et tous les Deutérostomiens, mais semblent avoir été perdues chez les insectes et les mollusques (Ashe, Bélicard, Le Pen, Sarkies et al., 2013. Table 1). Contrairement à ce qui est observé chez les autres Eumétazoaires étudiés, l'homologue de RIG-I aurait donc disparu chez les insectes au profit de l'apparition d'une nouvelle

protéine DICER, DCR-2, spécialisée dans la réponse antivirale. On sait que les protéines RIG-I et MDA5 déclenchent les mêmes voies chez tous les vertébrés (Chen et al., 2013; Langevin et al., 2013). L'absence d'interferon chez la plupart des autres Eumétazoaires suggère que ces protéines RIG-I induisent d'autres voies. Il serait intéressant de savoir si, comme chez *C. elegans*, ces protéines sont impliquées dans la reconnaissance des ARN viraux et amorcent une réponse par les petits ARN. De même, certaines espèces sont dépourvues de protéines contenant à la fois un domaine hélicase et RIG-I et ne possèdent qu'une seule protéine DICER, suggérant que de nouvelles voies impliquées dans les réponses antivirales restent à découvrir.

DRH-1 est clairement incapable d'engager une réponse similaire à celle des mammifères. En revanche, l'existence d'une réponse antivirale par les petits ARN chez les mammifères est sujette à débat. Il semblerait que les cellules embryonnaires soient capables de monter une telle réponse (Li et al., 2013; Maillard et al., 2013) mais pas les cellules différenciées (Backes et al., 2014).

La partie N terminal de DRH-1 ne possède pas de domaine CARD mais il n'est pas impossible que cette protéine joue un autre rôle lors d'une infection virale. On sait par exemple que la protéine DICER-2 de la drosophile induit également la production du ligand vago qui déclenche une réponse antivirale via la voie Jak/STAT (Deddouche et al., 2008).

Nous avons mis en évidence ici l'importance de la découverte des virus dans la compréhension du fonctionnement de la réponse antivirale chez *C. elegans*. De manière surprenante, le gène *drh-1*, impliqué dans la reconnaissance spécifique des ARN viraux, semble avoir perdu sa fonction dans certaines souches naturelles de *C. elegans* alors qu'il s'agit d'un gène hautement conservé chez les eumétazoaires. La deuxième partie du projet a été d'essayer de comprendre l'histoire évolutive de l'interaction *C. elegans* - virus d'Orsay en tentant de retracer l'histoire de la délétion *niDf250*.

II. Histoire évolutive de l'interaction *Caenorhabditis* – Virus naturels

II.A. Les virus sont spécifiques à l'échelle de l'espèce

Caenorhabditis briggsae et *C. elegans* se retrouvent dans les mêmes substrats en France et plus particulièrement à Orsay (Essonne) et à Santeuil (Vexin) (Félix et Duveau, 2012). Ces deux espèces se côtoient et interagissent vraisemblablement avec les mêmes espèces dans leur milieu naturel. Il s'avère que les virus découverts sont exclusifs à l'espèce dans laquelle ils ont été découverts (il en est de même pour le virus du Blanc), suggérant une adaptation des virus à leur hôte (Félix et al., 2011). La base génétique et/ou physiologique de cette barrière spécifique n'a pas encore été identifiée. Nous avons montré qu'une souche de *C. elegans* déficiente dans la production de petits ARN antiviraux contre le virus d'Orsay est tout de même résistante au virus de Santeuil (Félix et al., 2011). On peut supposer que l'incompatibilité se fait avant ou au moment de l'entrée du virus dans la cellule.

II.B. La capacité de réplication du virus d'Orsay varie entre souches sauvages de *C. elegans*

Les souches standards de laboratoire N2 (*C. elegans*) et AF16 (*C. briggsae*) ne sont pas sensibles aux virus d'Orsay et de Santeuil, respectivement (Félix et al., 2011), suggérant une variation intraspécifique de la capacité du virus à se répliquer. L'infection de 97 isolats naturels représentant la diversité mondiale connue de *C. elegans* confirme cette variation (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1 - figure supplement 1A). Le patron de variation apparemment continu suggère l'implication de plusieurs variations génétiques dans la capacité à empêcher la réplication du virus. L'analyse pangénomique sur ce caractère nous a révélé une seule région candidate de 6 Mb au centre du chromosome IV (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1A). La garnde taille de la région est directement liée au fait que les parties centrales des chromosomes ont un fort déséquilibre de liaison dans les populations naturelles de *C. elegans*. Une cartographie plus fine de la région par des croisements de laboratoire a révélé la présence d'un polymorphisme du gène *drh-1* - une délétion de 159 pb, nommée *niDf250* - perturbant le fonctionnement de la protéine DRH-1 (Ashe, Bélicard, Le Pen,

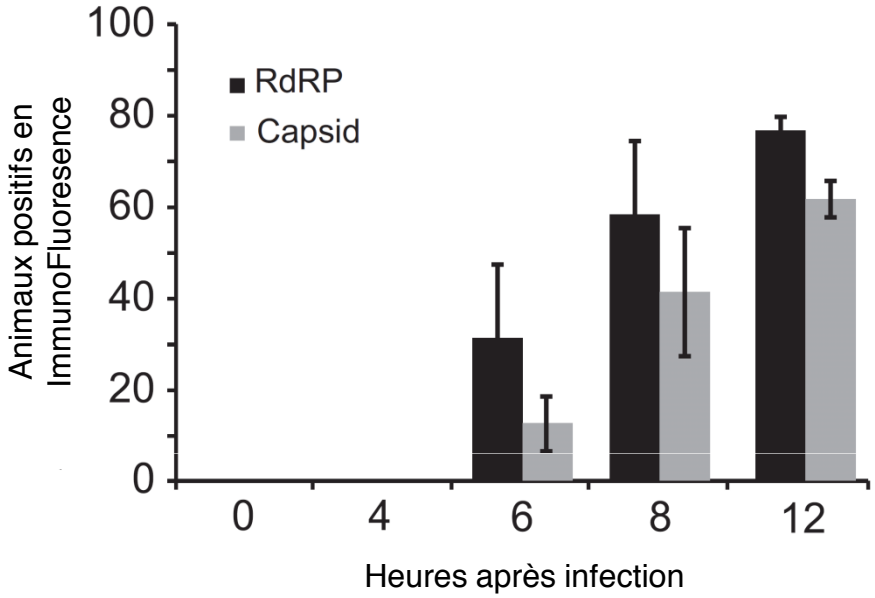


Figure 33. Suivie de l'infection du virus d'Orsay à partir d'infections effectuées au stade L1. Les animaux infectés sont déficients pour l'ARN interférence (mutant *rde-1*). Image issue de Franz et al. (2014).

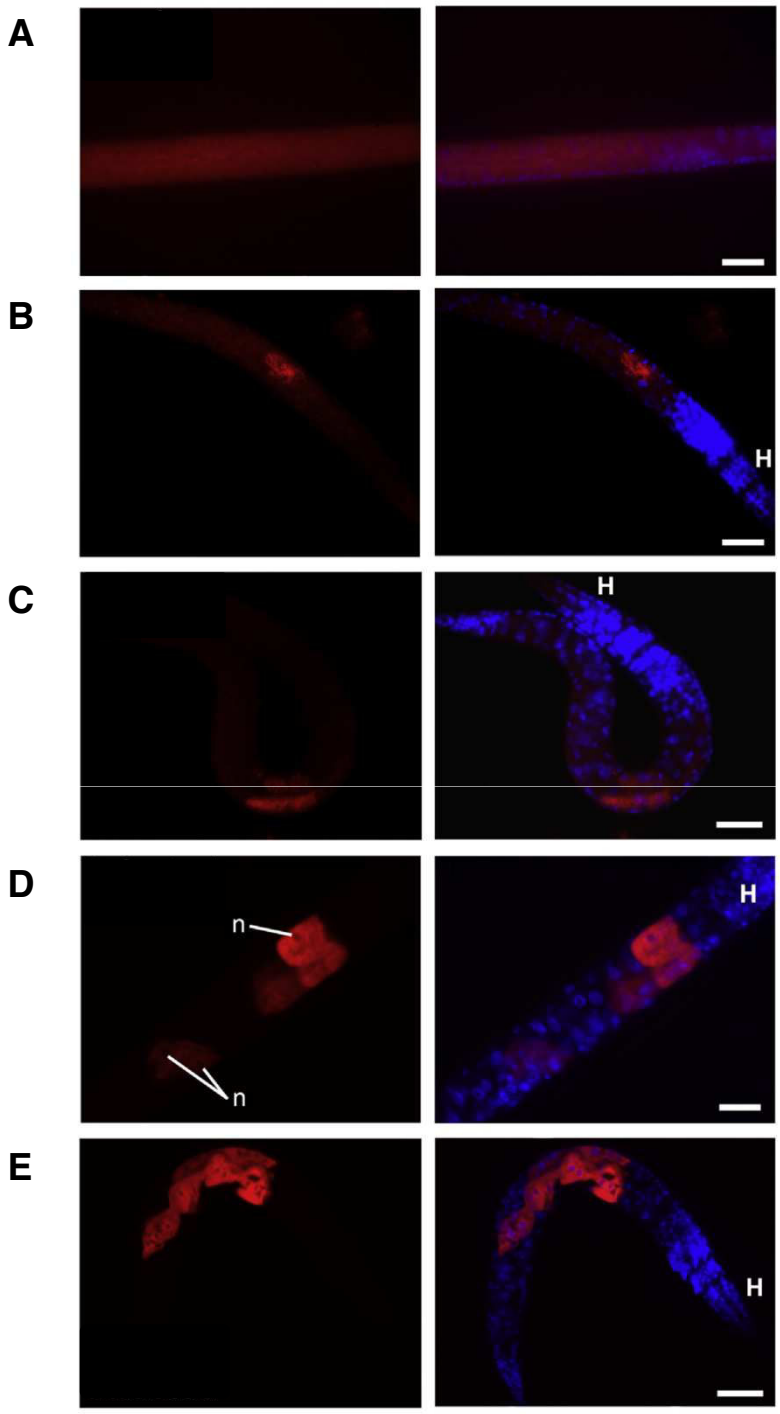


Figure 34 Détection du RNA1 du virus d'Orsay (rouge) par FISH dans des mutants *rde-1*. **A:** En absence de virus. **B-E:** Infectés par le virus d'Orsay. Les noyaux (n) sont marqués par DAPI (bleu). La barre d'échelle représente 20 µm. Images issues de Franz et al. (2014).

Sarkies et al., 2013; Figure 1 et figure supplement 2). Une délétion non-sens dans le gène *drh-1* dans le fond génétique N2, résistant, rend la réplication du virus possible à un niveau similaire à celui mesuré dans JU1580. Ces résultats suggèrent qu'un seul locus pourrait être impliqué dans la différence phénotypique entre N2 et JU1580 (voir aussi l'expérience en supplement 2B). La continuité de sensibilité observée dans les 97 souches naturelles peut s'expliquer de différentes façons, non exclusives:

- Erreur de mesure: Cette continuité peut être expliquée en partie par des erreurs de mesure, en particulier liées à la technique de RT-PCR utilisée pour mesurer la charge virale. On note d'amples variations entre réplicats, qu'ils soient dans des infections indépendantes (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1 - figure supplement 1A) ou pour la même infection.
- Variation biologique: Une partie de la variation observée entre réplicats semble directement imputable à la variation de la propagation du virus d'Orsay entre les réplicats d'infection. En effet, si l'on compare les trois réplicats d'infection réalisés pour chaque souche, on constate qu'il n'est pas rare d'observer un réplicat rapportant une charge virale très basse, voir nulle, alors que les deux autres rapportent une charge virale élevée, et ce indépendamment du lot d'infections (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1 - figure supplement 1A). Ceci est probablement dû à la variation biologique du processus d'infection elle-même. Une infection peut échouer dans une souche sensible tandis qu'il est parfois possible de détecter l'infection d'une souche capable d'une réponse antivirale. C'est par exemple le cas pour la souche N2 dans laquelle on peut certaines fois détecter la présence de virus par RT-PCR (Félix et al., 2011). Même pour les socuhes les plus sensibles, le virus d'Orsay n'infecte généralement que 50% de la population une semaine après l'infection (Félix et al., 2011). Le taux de transmission de ce virus est donc assez faible et il n'est pas surprenant d'observer l'absence d'amorçage de sa propagation ou sa disparition de certaines populations.

Participation d'autres loci: Enfin, il apparaît que plusieurs loci participent à la sensibilité au virus d'Orsay dans les populations naturelles et que nous n'ayons détecté dans cette étude que celui responsable de la différence de sensibilité entre JU1580 et N2 (parents des lignées recombinantes utilisées pour restreindre la région d'intérêt) (voir section IID). Nous avons ici supposé que le nombre d'individus infectés a atteint un plateau sept jours (trois générations de nématodes) après l'infection de quatre larves L4. Ceci est un choix arbitraire car jusqu'à présent, aucune expérience n'a vérifié cette assomption. Une meilleure compréhension de la dynamique de l'infection de la population serait nécessaire pour confirmer ce choix. Néanmoins, une invalidation de celui-ci ne remettrait

pas en cause l'étude effectuée puisque les études d'associations pangénomiques permettent d'étudier le caractère de son choix, pourvu qu'il varie. Le suivi d'une infection de L1 a cependant démontré qu'au bout de 12 heures après l'infection, environ 60% des individus sont infectés (Figure 33), pourcentage qui est similaire à celui retrouvé une semaine après infection. Cette expérience n'est effectuée que sur une seule génération et la quantité de virus déposée n'est pas totalement maîtrisée mais cela nous donne une indication sur la vitesse du processus d'infection dans une population. Ainsi, une mesure de la charge virale au bout d'une semaine ne permet pas d'étudier un seul aspect de l'infection mais d'envisager la capacité générale du virus à infecter et à se maintenir dans la population. Cette mesure prend donc en compte à la fois la capacité du virus à infecter les individus, sa capacité à se multiplier et sa capacité à se transmettre dans la population.

Nous utilisons dorénavant au laboratoire le FISH afin de détecter la présence de virus dans les populations de *Caenorhabditis* infectés (Figure 34). En plus d'être moins onéreuse que la mesure de la charge virale par RT-qPCR, elle permet également d'évaluer la prévalence de l'infection dans une population ainsi que d'avoir une meilleure idée de la quantité de cellules intestinales infectées. Ces deux derniers phénotypes peuvent avoir leur importance dans la compréhension du phénomène d'infection. Par exemple, la seule utilisation de la RT-qPCR ne permet pas de savoir si la charge virale observée est due au fait que le virus se réplique faiblement dans tous les individus ou s'il se réplique fortement dans un faible nombre d'individus. Grâce à cette technique, nous comptons en apprendre plus sur la dynamique de l'infection et notamment sur la capacité de transmission du virus en fonction du génotype de l'hôte.

II.C. Histoire évolutive de *niDf250*

II.C.1. Répartition de l'allèle **niDf250** *dans les populations naturelles*

On s'attend à ce que le processus de course aux armements entre l'hôte et son parasite aboutisse à l'évolution vers la résistance de l'hôte au parasite contemporain et que, par conséquent, l'allèle conférant une résistance soit dérivé. Par exemple, chez *Drosophila melanogaster*, l'insertion d'un transposon dans le gène CHKov-1 est responsable d'une augmentation de la résistance au Sigma virus dans les populations naturelles (Magwire et al., 2011). Par la suite, une duplication de ce nouvel allèle a permis une résistance encore plus marquée à ce virus. De manière surprenante, dans notre étude, l'allèle conférant une résistance au virus d'Orsay est ancestral par rapport à celui qui cause la sensibilité (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1F). Mise à part une conversion génique

postérieure à la délétion, il est difficilement imaginable ici que *C. elegans* ait évolué vers une sensibilité au virus d'Orsay. Pourtant, une collecte intensive de *C. elegans* montre que la délétion *niDf250* est très majoritaire (moyenne de 87,5% sur trois années de collecte, données non publiées) sur le site d'origine du virus d'Orsay. De même, 22 isolats sur les 97 possèdent la délétion *niDf250* identifiée dans la souche JU1580. La délétion *niDf250* se retrouve dans la population mondiale à une fréquence intermédiaire (23%), en majorité en Europe (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2A).

II.C.2. Des traces de balayage sélectif récent autour de niDf250

Comment expliquer une fréquence si importante pour un allèle conférant la sensibilité à un parasite ? Une première hypothèse pourrait provenir de la faible prévalence du virus d'Orsay dans la nature. En effet, ce dernier n'a été trouvé que deux fois, initalement en 2008 à Orsay et en 2013 à Ivry, parmi de multiples échantillonages dans de multiples localités (Lise Frézal et Marie-Anne Félix, en préparation). On peut alors penser que cette délétion se propage simplement sous régime de neutralité. L'étude de la région détectée par l'étude pangénomique révèle la présence de deux haplotypes majeurs très peu différents : celui incluant l'allèle fonctionnel (haplotype N2) et celui incluant l'allèle *niDf250* (haplotype JU1580) (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2C et Figure 2 - figure supplement 2). Il s'avère que le polymorphisme est nul sur une portion de 3Mb des haplotypes JU1580 contenant *niDf250* (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2E). Cette absence de polymorphisme associée à un fort taux de déséquilibre de liaison suggère un balayage sélectif récent de cette région. La présence de cee balayage sélectif amène trois hypothèses quant à la fréquence observée de la délétion:

- Le gène *drh-1* fonctionnel est délétère en absence de virus dans certaines conditions. Il est possible que la protéine DRH-1 induise des modifications physiologiques indésirables ou interfère avec la voie des petits ARN exogènes ou endogènes. Une délétion dans ce gène aurait alors été sélectionnée et se serait propagée dans les populations en emportant toute une région du chromosome IV.
- L'infection par le virus d'Orsay apporte dans certaines conditions.un avantage sélectif aux souches qui y sont sensibles
- L'allèle *niDf250* a été sélectionné passivement car il est en déséquilibre de liaison avec un allèle sélectionné positivement dans la région avec laquelle il ségrège.

II.C.3. Quelle pression de sélection pourrait avoir conduit à un balayage sélectif dans cette région ?

Afin de tester (dans les conditions de laboratoire) la première hypothèse, nous avons

réalisé une expérience de compétition au laboratoire entre une souche possédant la région liée à l'allèle *niDf250* dans un fond génétique N2 et une souche possédant l'haplotype N2 (la souche N2 donc). Au bout de 12 générations, les deux génotypes se maintiennent en proportion dans la population, ne validant pas l'hypothèse selon laquelle l'haplotype liée à *niDf250* conférerait un avantage dans les conditions testées. En revanche, en présence de virus, la fréquence du génotype portant l'allèle *niDf250* diminue, confirmant l'effet délétère de la délétion en présence de virus dans le milieu (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2B). La limite évidente de cette expérience de compétition est qu'elle a été effectuée en conditions de laboratoire. Il est envisageable que la délétion *niDf250* et/ou que la sensibilité au virus d'Orsay confèrent un avantage dans certaines conditions naturelles, expliquant ainsi pourquoi cette délétion est prépondérante localement.

On constate même que pour les expériences de compétition entre souches sensible et résistante, environ 10% d'animaux sensibles persistent dans la population après 12 générations (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 2B). Il n'est pas étonnant qu'une petite proportion d'animaux sensibles persiste dans la population. Il est vraisemblable qu'une augmentation de la proportion d'individus résistants dans la population a un effet négatif sur la fitness du virus. Le virus pourrait ainsi avoir disparu bien avant que l'allèle de résistance ne se fixe dans la population d'hôte, permettant ainsi une persistance du génotype sensible.

Certains parasites peuvent être plus virulents sur *C. elegans* et induisent une perte de fitness beaucoup plus importante, résultant le plus souvent d'une diminution de la survie avant la fin de la période de ponte des hermaphrodites (Tableau 2). C'est le cas par exemple de la bactérie *Leucobacter* Verde2 qui tue son hôte en moyenne un jour après l'infection ou encore de *Pseudomonas aeruginosa* qui peut tuer ce dernier en quatre heures (Sifri et al., 2005). Certains parasites intracellulaires protègent leurs hôtes contre d'autres parasites intracellulaires. C'est le cas de certaines *Wolbachia* qui protègent contre l'infection par le Drosophila C Virus (Osborne et al., 2009). De la même façon, il est possible que le virus d'Orsay puisse protéger les individus sensibles contre d'autres parasites. Un traitement supplémentaire de l'expérience de compétition a été d'infecter les souches de *C. elegans* par une souche de microsporidie qui affecte plus fortement *C. elegans* que le virus (Figure 35). Nous n'avons détecté aucun effet de la microsporidie sur la dynamique des fréquences des génotypes résistant versus sensible lors de la compétition. Il ne semble donc pas que la délétion *niDf250* ni l'infection par le virus protègent *C. elegans* contre l'infection par les microsporidies. Cette expérience préliminaire mérite probablement d'être réitérée en utilisant une souche de microsporidie plus virulente que celle utilisée et des conditions plus favorables au développement des microsporidies.

Une étude récente montre que lors d'une infection de JU1580 par le virus d'Orsay, une

panoplie de gènes impliquée dans la réponse à l'infection par *Pseudomonas aeruginosa* et *Microbacterium nematophilum* sont surexprimés (Sarkies et al., 2013). D'après le modèle établi lors de cette étude, JU1580 est incapable de reconnaître les ARN doubles brins produits lors de la réplication du virus d'Orsay en absence de protéines DRH-1 fonctionnelles (Figure 36). Cette réplication produit une grande quantité d'ARN simples brins positifs viraux possédant des structures identiques aux microARN supprimant l'expression de gènes antibactériens. Une partie des protéines DCR est donc accaparée par les ARN simples brins viraux au lieu de participer à la production de microARN. Les gènes antibactériens ne sont donc plus réprimés et une réponse se met en place. Afin de vérifier si cette surexpression se traduisait par un réel gain de fitness pour JU1580, nous avons tenté de comparer la longévité de cette souche infectée par *Pseudomonas aeruginosa* préalablement infectée ou non par le virus d'Orsay. Nous n'avons pas pu constater d'effet au laboratoire (Figure 37).

Afin de voir si cette région confère véritablement un avantage sélectif, une expérience intéressante, mais assez complexe à mettre en place, serait de mettre en compétition la souche qui possède cette région avec la souche qui possède l'haplotype N2 dans un environnement récapitulant au mieux les conditions de vie naturelles (microcosme ?) de *C. elegans* en présence ou non du virus d'Orsay. Il n'est cependant pas certain que les conditions favorisant *niDf250* ou l'haplotype qui lui est associé puissent être identifiées un jour tant les possibilités de conditions différentes à tester sont nombreuses.

II.D. Le polymorphisme de *drh-1* n'explique pas totalement la sensibilité au virus d'Orsay

Ces travaux sont les premiers qui se basent principalement sur l'étude d'association pangénomique (EAP) pour identifier la base génétique d'une variation de phénotype chez *C. elegans*. Le grand avantage de ce type d'association est de pouvoir tirer les informations directement de la diversité des populations naturelles et ainsi de prendre en considération toute la diversité naturelle connue (contrairement à l'analyse QTL qui ne prend en compte que le polymorphisme entre les parents).

Grâce à cette analyse, nous avons pu mettre en évidence le polymorphisme expliquant en grande partie la sensibilité au virus d'Orsay dans les populations naturelles de *C. elegans*. Cependant, l'EAP ne permet pas déterminer précisément l'architecture génétique de la sensibilité au virus d'Orsay. Même si nous avons été capables d'identifier ce locus majeur, il n'est pas exclu que la résistance ou la sensibilité au virus implique de nombreux autres loci mineurs indétectables par l'étude d'association.

Ainsi, même si la corrélation entre la présence de l'allèle *niDf250* et une plus grande

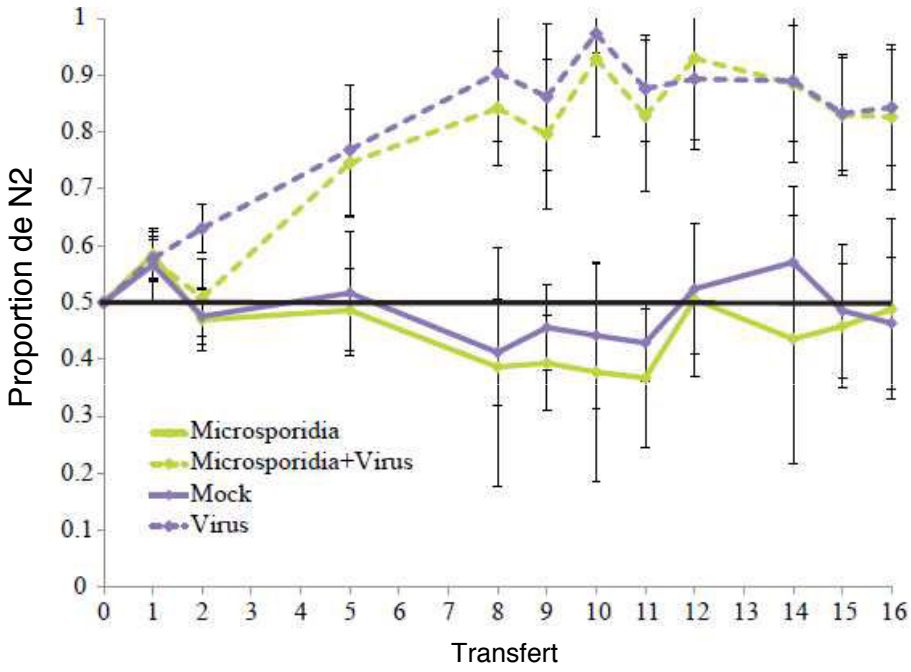


Figure 35. Expérience de compétition entre N2 et une souche possédant l'haplotype associé à *niDf250* en absence ou en présence du virus d'Orsay avec l'ajout ou non de microsporidies. L'ajout de microsporidie ne modifie pas les proportions des souches que ce soit avec ou sans virus.

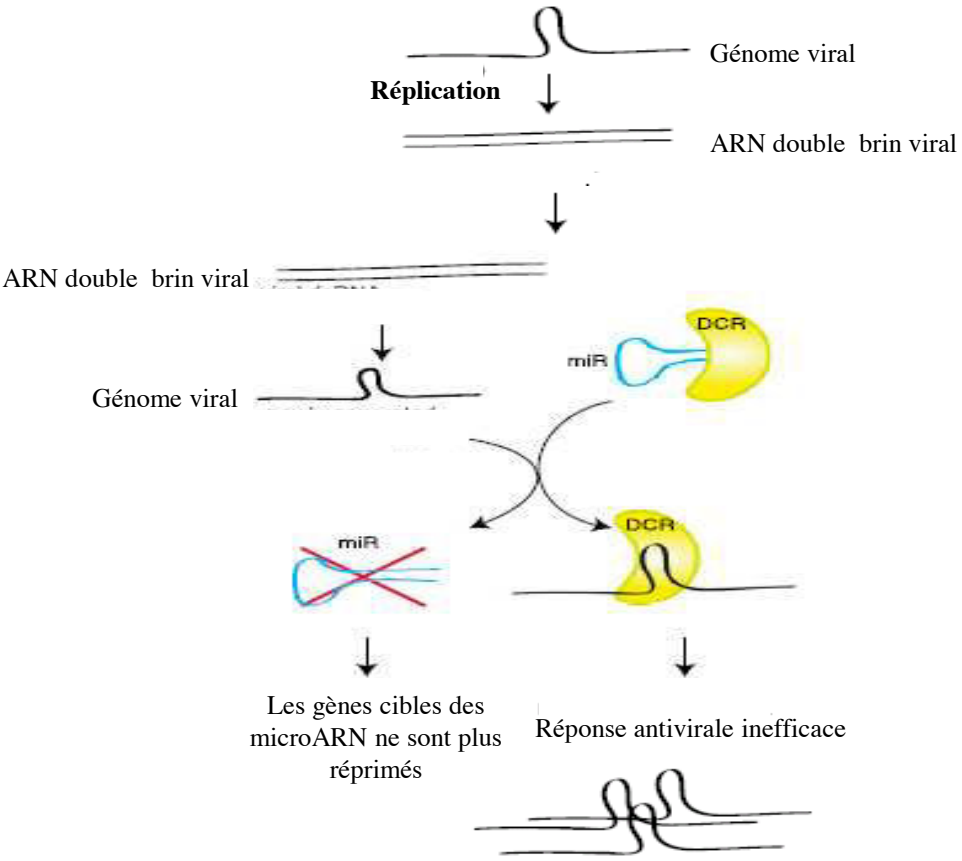


Figure 36. Modèle de compétition entre la voie des microARN et des siARN antiviraux chez JU1580. L'unique protéine DCR de *C. elegans* est présente en quantité limitée dans la cellule. Lors de la présence d'ARN double brin viral, la quantité de DCR allouée à la fabrication des microARN diminue. Les gènes habituellement réprimés par ces microARN ne le sont plus. Ces gènes peuvent notamment être impliqués dans la réponse aux bactéries. D'après Sarkies et al. (2013).

sensibilité au virus d'Orsay est forte, nous avons pu remarquer que quelques isolats parmi les 97 testés ne la suivent pas. En effet, le virus se réplique facilement dans quelques isolats possédant l'allèle fonctionnel de *drh-1* (Ashe, Bélicard, Le Pen, Sarkies et al., 2013; Figure 1 - figure supplement 1). Il est possible que ces souches soient déficientes pour un autre des acteurs de l'ARN interférence ou pour d'autres éléments impliqués dans une réponse antivirale encore non identifiée. De façon plus surprenante, il existe quelques souches résistantes portant *niDf250*, comme MY10. Or dans ce type de souche, le virus est censé pouvoir se répliquer normalement en absence de réponse de l'hôte via les petits ARN. On peut penser que la difficulté qu'a le virus d'Orsay à infecter cette souche provient d'un problème de transmission du virus entre individus, d'entrée dans la cellule hôte ou de réplication. La production de lignées recombinantes (RIL) entre deux souches dont les sensibilités au virus diffèrent permet d'identifier les loci responsables de cette sensibilité. Chez *C. elegans*, la production de RIL est très intéressante car elle permet d'obtenir bien plus de recombinants que ce qui n'est observé dans la nature. Les recombinaisons étant faibles dans les populations naturelles, surtout au centre des chromosomes, la production de RIL permet d'obtenir en laboratoire de nouveaux haplotypes inédits dans la nature en brisant le déséquilibre de liaison. L'obtention d'individus mosaïque aide entre autres à détecter les loci à effet mineur et à révéler les différentes interactions (épistasie) pouvant avoir lieu entre deux loci ou plus.
Je suis actuellement en train de réaliser des lignées recombinantes entre d'un côté MY10 et une souche sensible possédant *niDf250* (JU1580) et des lignées recombinantes entre deux souches possédant l'allèle fonctionnel de *drh-1*, une sensible (CB4852) et l'autre résistante (JU1200). Le séquençage et le phénotypage de ces lignées recombinantes (RIL) va permettre d'identifier, par une analyse QTL, de nouveaux polymorphismes naturels responsables de la résistance ou de la sensibilité au virus d'Orsay chez *C. elegans*.

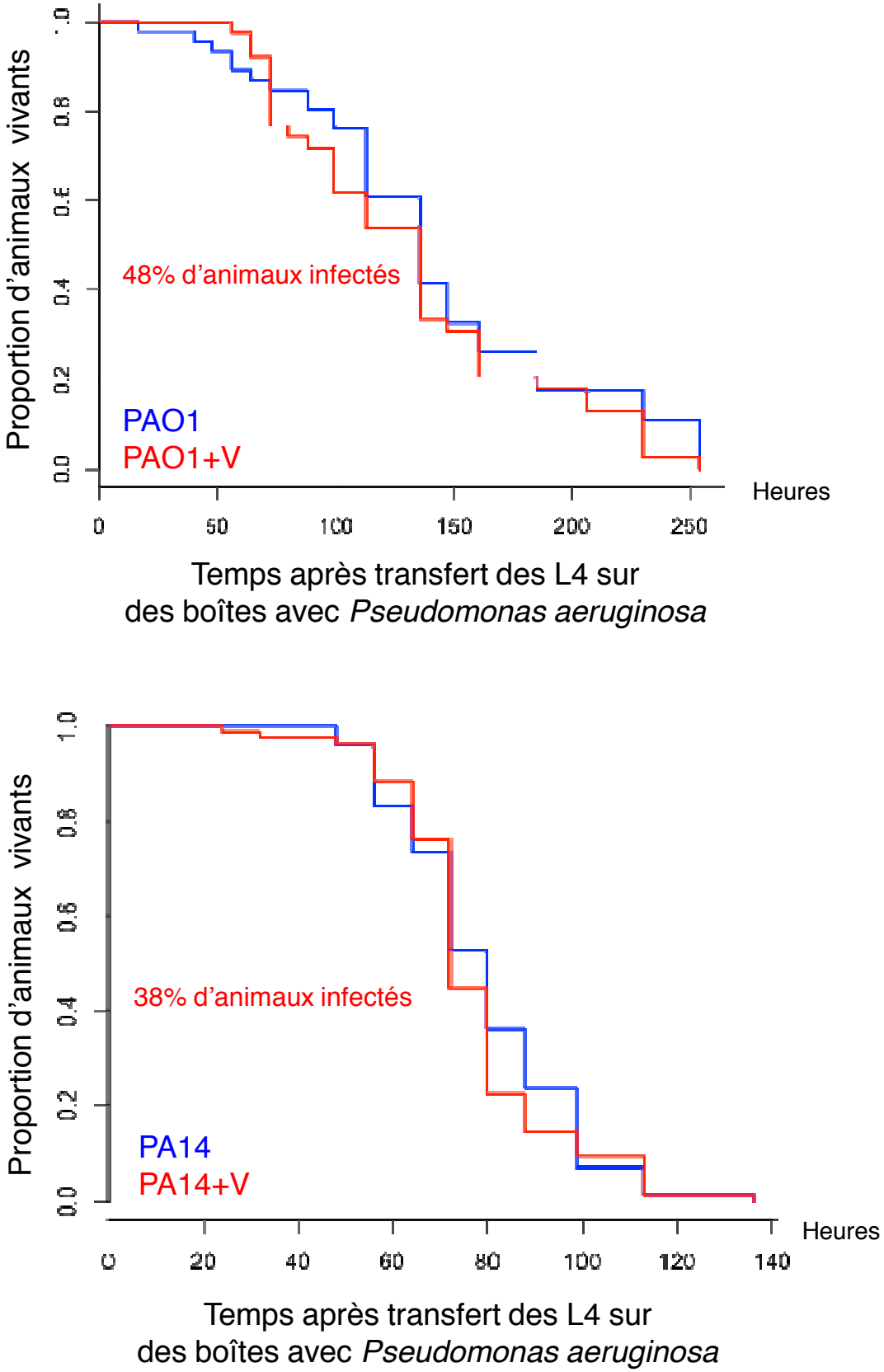


Figure 37. Courbes de survie de *Caenorhabditis elegans* de la souche JU1580 préalablement infectés ou non par le virus d'Orsay (V) nourris exclusivement par *Pseudomonas aeruginosa*. A: La souche de *P. aeruginosa* PA01 n'est pas délétère pour les nématodes. Comme observé précédemment, le virus n'a pas d'effet sur la survie de JU1580. **A:** La souche de *P. aeruginosa* PA14 est pathogène pour *C. elegans*. Aucune différence n'est observée entre les deux traitements. Pour **A** et **B,** le pourcentage d'animaux infectés par le virus dans la population d'origine des L4.

Conclusion

Avec la découverte des premières infections virales des *Caenorhabditis*, nous avons pu mettre en évidence le rôle de *drh-1* dans la reconnaissance des ARN viraux intracellulaires et dans la mise en place d'une réponse antivirale par un mécanisme d'ARN interférence. De manière assez surprenante, l'allèle dérivé de *drh-1*, la délétion *niDf250*, confère une sensibilité au virus d'Orsay, suggérant une évolution vers la sensibilité au virus chez *C. elegans*. Néanmoins, cet allèle appartient à un haplotype plus vaste qui semble avoir été la cible d'un balayage sélectif récent. Nous n'avons pas pu identifier la cause de cette marque de sélection en condition de laboratoire mais il se pourrait que cet haplotype confère un avantage important à son détenteur. L'évolution de l'interaction *C. elegans* - virus d'Orsay dans la nature pourrait donc être conditionnée par l'existence d'une pression de sélection plus importante que celle exercée par le virus d'Orsay. Cette étude illustre parfaitement le fait que les modèles d'évolution d'interaction hôtes parasites présentés en introduction (II.A.2.) ne peuvent être généralisés. Pour chaque étude, il me paraît nécessaire de prendre en compte l'environnement génomique du locus mais également d'avoir une idée des fonctions associées à celui-ci. Même si la pression de sélection exercée par les parasites sur son hôte est forte, les gènes impliqués dans la réponse à l'infection peuvent également être liés physiquement (déséquilibre de liaison) à des loci soumis à d'autres pressions de sélection ou encore posséder des fonctions autres que la réponse au parasite (pléiotropie). La dynamique des allèles à ce locus peut donc être influencée par de multiples facteurs indépendants de l'interaction avec le parasite.

Ces travaux soulignent également le grand potentiel du système *Caenorhabditis*-virus pour l'étude de l'évolution des interactions animal-virus. Ainsi, les très nombreux outils développés chez *C. elegans* ont permis d'identifier dans un laps de temps très court à la fois le polymorphisme associé à la sensibilité au virus mais également les mécanismes impliqués dans la réponse antivirale et ainsi d'avoir une approche intégrative du phénotype de sensibilité.

Néanmoins une meilleure connaissance des interactions entre ces nématodes et leur environnement naturel permettrait de comprendre mieux l'histoire évolutive qui nous est contée par leur génome.

BIBLIOGRAPHIE

Agrawal, A. F. and Lively, C. M. (2003). Modelling infection as a two-step process combining gene-for-gene and matching-allele genetics. *Proc. Biol. Sci.* **270**, 323–334.

Aguinaldo, A. M., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A. and Lake, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489–493.

Akira, S. and Takeda, K. (2004). Functions of toll-like receptors: lessons from KO mice. *C. R. Biol.* **327**, 581–589.

Alcazar, R. M., Lin, R. and Fire, A. Z. (2008). Transmission dynamics of heritable silencing induced by double-stranded RNA in *Caenorhabditis elegans*. *Genetics* **180**, 1275–1288.

Alper, S., McBride, S. J., Lackford, B., Freedman, J. H. and Schwartz, D. A. (2007). Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Mol. Cell. Biol.* **27**, 5544–5553.

Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T. and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in *C. elegans*. *Curr. Biol.* **13**, 807–818.

Andersen, E. C., Gerke, J. P., Shapiro, J. A., Crissman, J. R., Ghosh, R., Bloom, J. S., Félix, M.-A. and Kruglyak, L. (2012). Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat. Genet.* **44**, 285–290.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K. and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *EMBO J.* **26**, 5007–5019.

Aravind, L., Watanabe, H., Lipman, D. J. and Koonin, E. V. (2000). Lineage-specific loss and divergence of functionally linked genes in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11319–11324.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M. P., Cording, A. C., Doebley, A.-L., Goldstein, L. D., Lehrbach, N. J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**, 88–99.

Ausubel, F. M. (2005). Are innate immune signaling pathways in plants and animals conserved? *Nat. Immunol.* **6**, 973–979.

Avery, L. and Thomas, J. H. (1997). Feeding and Defecation. In *C. elegans II* (ed. Riddle, D. L., Blumenthal, T., Meyer, B. J., and Priess, J. R.), Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Backes, S., Langlois, R. A., Schmid, S., Varble, A., Shim, J. V., Sachs, D. and

tenOever, B. R. (2014). The Mammalian Response to Virus Infection Is Independent of Small RNA Silencing. *Cell Reports*.

Bakker, E. G., Toomajian, C., Kreitman, M. and Bergelson, J. (2006). A genome-wide survey of R gene polymorphisms in Arabidopsis. *Plant Cell* **18**, 1803–1818.

Barrière, A. and Félix, M.-A. Natural variation and population genetics of *Caenorhabditis elegans* (December 26, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.43.1, http://www.wormbook.org.

Barrière, A. and Félix, M.-A. (2005b). High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* **15**, 1176–1184.

Barrière, A. and Félix, M.-A. (2007). Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics* **176**, 999–1011.

Barrows, B. D., Haslam, S. M., Bischof, L. J., Morris, H. R., Dell, A. and Aroian, R. V. (2007). Resistance to Bacillus thuringiensis toxin in *Caenorhabditis elegans* from loss of fucose. *J. Biol. Chem.* **282**, 3302–3311.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., Chaves, D. A., Gu, W., Vasale, J. J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* **31**, 67–78.

Bélicard, T. and Félix, M.-A. (2012). [Multigenerational transmission of RNA interference in the nematode *Caenorhabditis elegans*]. *Med Sci (Paris)* **28**, 574–577.

Bélicard, T., Ashe, A., Le Pen, J., Sarkies, P., Frézal, L., Lehrbach, N. J., Félix, M.-A. and Miska, E. A. (2013). A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* **2**, e00994.

Blaxter, M. (1998). *Caenorhabditis elegans* is a nematode. *Science* **282**, 2041–2046.

Blaxter, M. L. (2003). Nematoda: genes, genomes and the evolution of parasitism. *Adv. Parasitol.* **54**, 101–195.

Bolla, R. I., Weinstein, P. P. and Cain, G. D. (1972). Fine structure of the coelomocyte of adult Ascaris suum. *J. Parasitol.* **58**, 1025–1036.

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S. and Sancristobal, M. (2010). Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics* **186**, 241–262.

Borgonie, G., Claeys, M., Dewaele, D. and Coomans, A. (1995). Ultrastructure of the

intestine of the bacteriophagous nematodes *Caenorhabditis*-elegans, *Panagrolaimus superbus* and *Acrobeloides maximus* (Nematoda, Rhabditida). *fundamental and applied nematology* **18**, 123–133.

Britten, R. J. (1995). Active gypsy/Ty3 retrotransposons or retroviruses in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 599–601.

Brockhurst, M. A. and Koskella, B. (2013). Experimental coevolution of species interactions. *Trends Ecol. Evol. (Amst.)* **28**, 367–375.

Brogden, K. A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* **3**, 238–250.

Brown, J. K. M. and Tellier, A. (2011). Plant-Parasite Coevolution: Bridging the Gap between Genetics and Ecology. *Annual Review of Phytopathology* **49**, 345–367.

Buckling, A. and Rainey, P. B. (2002). The role of parasites in sympatric and allopatric host diversification. *Nature* **420**, 496–499.

Bull, J. J. and Turelli, M. (2013). Wolbachia versus dengue: Evolutionary forecasts. *Evol Med Public Health* **2013**, 197–207.

Bull, J. J., Badgett, M. R. and Wichman, H. A. (2000). Big-benefit mutations in a bacteriophage inhibited with heat. *Mol. Biol. Evol.* **17**, 942–950.

Burkhart, K. B., Guang, S., Buckley, B. A., Wong, L., Bochner, A. F. and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. *PLoS Genet.* **7**, e1002249.

Burton, N. O., Burkhart, K. B. and Kennedy, S. (2011a). Nuclear RNAi maintains heritable gene silencing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19683–19688.

Burton, N. O., Burkhart, K. B. and Kennedy, S. (2011b). Nuclear RNAi maintains heritable gene silencing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19683–19688.

Campos, A., Cummings, M. P., Reyes, J. L. and Laclette, J. P. (1998). Phylogenetic relationships of platyhelminthes based on 18S ribosomal gene sequences. *Mol. Phylogenet. Evol.* **10**, 1–10.

Cao, Z., Xiong, J., Takeuchi, M., Kurama, T. and Goeddel, D. V. (1996). TRAF6 is a signal transducer for interleukin-1. *Nature* **383**, 443–446.

Cao, H., Baldini, R. L. and Rahme, L. G. (2001). Common Mechanisms for Pathogens of Plants and Animals. *Annual Review of Phytopathology* **39**, 259–284.

Carius, H. J., Little, T. J. and Ebert, D. (2001). Genetic variation in a host-parasite

association: potential for coevolution and frequency-dependent selection. *Evolution* **55**, 1136–1145.

Casals, F., Sikora, M., Laayouni, H., Montanucci, L., Muntasell, A., Lazarus, R., Calafell, F., Awadalla, P., Netea, M. G. and Bertranpetit, J. (2011). Genetic adaptation of the antibacterial human innate immunity network. *BMC Evolutionary Biology* **11**, 202.

Cassada, R. C. and Russell, R. L. (1975). The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **46**, 326–342.

Casteels, P., Ampe, C., Jacobs, F. and Tempst, P. (1993). Functional and chemical characterization of Hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee (Apis mellifera). *J. Biol. Chem.* **268**, 7044–7054.

Chalfie, M., Horvitz, H. R. and Sulston, J. E. (1981). Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* **24**, 59–69.

Chao, J. A., Lee, J. H., Chapados, B. R., Debler, E. W., Schneemann, A. and Williamson, J. R. (2005). Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nat. Struct. Mol. Biol.* **12**, 952–957.

Chen, S., Cheng, A. and Wang, M. (2013). Innate sensing of viruses by pattern recognition receptors in birds. *Vet. Res.* **44**, 82.

Cheng, T. C. (1986). *General Parasitology*. Academic Press College Division.

Civitello, D. J., Pearsall, S., Duffy, M. A. and Hall, S. R. (2013). Parasite consumption and host interference can inhibit disease spread in dense populations. *Ecol. Lett.* **16**, 626–634.

Combes, C. (2001). *Parasitism: The Ecology and Evolution of Intimate Interactions*. University of Chicago Press.

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D. and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **12**, 3715–3727.

Cutter, A. D. (2006). Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics* **172**, 171–184.

Cutter, A. D. and Payseur, B. A. (2003). Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Mol. Biol. Evol.* **20**, 665–673.

Darnell, J. E., Kerr, I. M. and Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415–1421.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life*. J. Murray.

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., Buhecha, H. R., Gilchrist, M. J., Howe, K. L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **31**, 79–90.

Daszak, P., Cunningham, A. A. and Hyatt, A. D. (2000). Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. *Science* **287**, 443–449.

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van Kerckhoven, L., Ebert, D. and De Meester, L. (2007). Host-parasite "Red Queen" dynamics archived in pond sediment. *Nature* **450**, 870–873.

Deddouche, S., Matt, N., Budd, A., Mueller, S., Kemp, C., Galiana-Arnoux, D., Dostert, C., Antoniewski, C., Hoffmann, J. A. and Imler, J.-L. (2008). The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activity in drosophila. *Nat. Immunol.* **9**, 1425–1432.

Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F. and Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235.

Dennis, S., Sheth, U., Feldman, J. L., English, K. A. and Priess, J. R. (2012). *C. elegans* germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog.* **8**, e1002591.

Dierking, K., Polanowska, J., Omi, S., Engelmann, I., Gut, M., Lembo, F., Ewbank, J. J. and Pujol, N. (2011). Unusual regulation of a STAT protein by an SLC6 family transporter in *C. elegans* epidermal innate immunity. *Cell Host Microbe* **9**, 425–435.

Ding, S.-W. and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. *Cell* **130**, 413–426.

Downing, T., Cormican, P., O'Farrelly, C., Bradley, D. G. and Lloyd, A. T. (2009). Evidence of the adaptive evolution of immune genes in chicken. *BMC Res Notes* **2**, 254.

Duchaine, T. F., Wohlschlegel, J. A., Kennedy, S., Bei, Y., Conte Jr., D., Pang, K., Brownell, D. R., Harding, S., Mitani, S., Ruvkun, G., et al. (2006). Functional Proteomics Reveals the Biochemical Niche of *C. elegans* DCR-1 in Multiple Small-RNA-Mediated Pathways. *Cell* **124**, 343–354.

Edwards, D. D. and Vidrine, M. F. (2006). Host specificity among *Unionicola spp.*

(Acari: Unionicolidae) parasitizing freshwater mussels. *J. Parasitol.* **92**, 977–983.

Eizaguirre, C., Lenz, T. L., Kalbe, M. and Milinski, M. (2012). Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecol. Lett.* **15**, 723–731.

Ellis, J. G., Lawrence, G. J., Luck, J. E. and Dodds, P. N. (1999). Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell* **11**, 495–506.

Engelmann, I., Griffon, A., Tichit, L., Montañana-Sanchis, F., Wang, G., Reinke, V., Waterston, R. H., Hillier, L. W. and Ewbank, J. J. (2011). A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. *PLoS ONE* **6**, e19055.

Estes, K. A., Szumowski, S. C. and Troemel, E. R. (2011). Non-lytic, actin-based exit of intracellular parasites from *C. elegans* intestinal cells. *PLoS Pathog.* **7**, e1002227.

Eulgem, T., Rushton, P. J., Schmelzer, E., Hahlbrock, K. and Somssich, I. E. (1999). Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J.* **18**, 4689–4699.

Ewbank, J. J. (2002). Tackling both sides of the host-pathogen equation with *Caenorhabditis elegans*. *Microbes Infect.* **4**, 247–256.

Excoffier, L., Hofer, T. and Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)* **103**, 285–298.

Fay, J. C. and Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405–1413.

Félix, M.-A. and Braendle, C. (2010). The natural history of *Caenorhabditis elegans*. *Curr. Biol.* **20**, R965–969.

Félix, M.-A. and Duveau, F. (2012). Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biol.* **10**, 59.

Félix, M.-A., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Bélicard, T., Jiang, Y., Zhao, G., Franz, C. J., Goldstein, L. D., et al. (2011). Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol.* **9**, e1000586.

Félix, M.-A., Jovelin, R., Ferrari, C., Han, S., Cho, Y. R., Andersen, E. C., Cutter, A. D. and Braendle, C. (2013). Species richness, distribution and genetic diversity of *Caenorhabditis* nematodes in a remote tropical rainforest. *BMC Evol. Biol.* **13**, 10.

Fenton, A., Antonovics, J. and Brockhurst, M. A. (2009). Inverse-Gene-for-Gene Infection Genetics and Coevolutionary Dynamics. *The American Naturalist* **174**, E230–E242.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998a). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998b). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.

Flajnik, M. F. and Du Pasquier, L. (2004). Evolution of innate and adaptive immunity: can we draw a line? *Trends in Immunology* **25**, 640–644.

Flor, H. (1956). The complementary genetic systems in flax and flax rust. *Advances in Genetics* **8**, 29–54.

Folgarait, P. J. and Gilbert, L. E. (1999). Phorid parasitoids affect foraging activity of *Solenopsis richteri* under different availability of food in Argentina. *Ecol. Entomol.* **24**, 163–173.

Frank, S. A. (1993). Specificity versus detectable polymorphism in host-parasite genetics. *Proc. Biol. Sci.* **254**, 191–197.

Franz, C. J., Zhao, G., Félix, M.-A. and Wang, D. (2012). Complete genome sequence of Le Blanc virus, a third *Caenorhabditis* nematode-infecting virus. *J. Virol.* **86**, 11940.

Franz, C. J., Renshaw, H., Frezal, L., Jiang, Y., Félix, M.-A. and Wang, D. (2014). Orsay, Santeuil and Le Blanc viruses primarily infect intestinal cells in *Caenorhabditis* nematodes. *Virology* **448**, 255–264.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925.

Fu, Y. X. and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709.

Garamszegi, L. Z. (2006). The evolution of virulence and host specialization in malaria parasites of primates. *Ecol. Lett.* **9**, 933–940.

Georgel, P., Naitza, S., Kappler, C., Ferrandon, D., Zachary, D., Swimmer, C., Kopczynski, C., Duyk, G., Reichhart, J.-M. and Hoffmann, J. A. (2001). Drosophila Immune Deficiency (IMD) Is a Death Domain Protein that Activates Antibacterial Defense and Can Promote Apoptosis. *Developmental Cell* **1**, 503–514.

Gilmore, T. D. and Wolenski, F. S. (2012). NF-κB: where did it come from and why? *Immunol. Rev.* **246**, 14–35.

Goh, W.-S. S., Seah, J. W. E., Harrison, E. J., Chen, C., Hammell, C. M. and Hannon, G. J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. *Genes Dev.* **28**, 797–807.

Gokhale, C. S., Papkou, A., Traulsen, A. and Schulenburg, H. (2013). Lotka¿Volterra dynamics kills the Red Queen: population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. *BMC Evol. Biol.* **13**, 254.

Goldmann, D. A. (2000). Transmission of viral respiratory infections in the home. *Pediatr. Infect. Dis. J.* **19**, S97–102.

Graustein, A., Gaspar, J. M., Walters, J. R. and Palopoli, M. F. (2002). Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics* **161**, 99–107.

Gravato-Nobre, M. J., Nicholas, H. R., Nijland, R., O'Rourke, D., Whittington, D. E., Yook, K. J. and Hodgkin, J. (2005). Multiple genes affect sensitivity of *Caenorhabditis elegans* to the bacterial pathogen *Microbacterium nematophilum*. *Genetics* **171**, 1033–1045.

Gray, J. C. and Cutter, A. D. (2014). Mainstreaming *Caenorhabditis elegans* in experimental evolution. *Proc. R. Soc. B* **281**, 20133055.

Greenberg, S. and Grinstein, S. (2002). Phagocytosis and innate immunity. *Current Opinion in Immunology* **14**, 136–145.

Greischar, M. A. and Koskella, B. (2007). A synthesis of experimental work on parasite local adaptation. *Ecol. Lett.* **10**, 418–434.

Grishok, A., Tabara, H. and Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* **287**, 2494–2497.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G. and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34.

Gross, S. R. (1954). Abortive infection of a strain of *Escherichia coli* by coliphage t2. *Journal of Bacteriology* **68**, 36.

Gu, W., Shirayama, M., Conte, D., Jr, Vasale, J., Batista, P. J., Claycomb, J. M., Moresco, J. J., Youngman, E. M., Keys, J., Stoltz, M. J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the *C. elegans* germline. *Mol. Cell* **36**, 231–244.

Guang, S., Bochner, A. F., Pavelec, D. M., Burkhart, K. B., Harding, S., Lachowiec, J. and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. *Science* **321**, 537–541.

Guo, X. and Lu, R. (2013). Characterization of virus-encoded RNA interference suppressors in *Caenorhabditis elegans*. *J. Virol.* **87**, 5414–5423.

Guo, X., Zhang, R., Wang, J., Ding, S.-W. and Lu, R. (2013). Homologous RIG-I-like helicase proteins direct RNAi-mediated antiviral immunity in *C. elegans* by distinct mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16085–16090.

Haber, M., Schüngel, M., Putz, A., Müller, S., Hasert, B. and Schulenburg, H. (2005). Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Mol. Biol. Evol.* **22**, 160–173.

Haft, D. H., Selengut, J., Mongodin, E. F. and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**, e60.

Haskins, K. A., Russell, J. F., Gaddis, N., Dressman, H. K. and Aballay, A. (2008). Unfolded protein response genes regulated by CED-1 are required for *Caenorhabditis elegans* innate immunity. *Dev. Cell* **15**, 87–97.

Hedrick, P. W. and Miller, P. S. (1994). Rare alleles, MHC and captive breeding. *EXS* **68**, 187–204.

Hodgkin, J. and Doniach, T. (1997). Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149–164.

Hodgkin, J., Kuwabara, P. E. and Corneliussen, B. (2000). A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Curr. Biol.* **10**, 1615–1618.

Hodgkin, J., Félix, M.-A., Clark, L. C., Stroud, D. and Gravato-Nobre, M. J. (2013). Two leucobacter strains exert complementary virulence on *Caenorhabditis* including death by worm-star formation. *Curr. Biol.* **23**, 2157–2161.

Hoffman, J. I., Webster, J. P., Ndamba, J. and Woolhouse, M. E. (1998). Extensive genetic variation revealed in adjacent populations of the schistosome intermediate host *Biomphalaria pfeifferi* from a single river system. *Ann Trop Med Parasitol* **92**, 693–698.

Hudson, R. R., Kreitman, M. and Aguadé, M. (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**, 153–159.

Huffman, D. L., Abrami, L., Sasik, R., Corbeil, J., van der Goot, F. G. and Aroian, R. V. (2004). Mitogen-activated protein kinase pathways defend against bacterial

pore-forming toxins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10995–11000.

Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* **18**, 486.

Hurst, L. D. and Smith, N. G. C. (1999). Do essential genes evolve slowly? *Current Biology* **9**, 747–750.

Hutchings, M. R., Kyriazakis, I. and Gordon, I. J. (2001). Herbivore physiological state affects foraging trade-off decisions between nutrient intake and parasite avoidance. *Ecology* **82**, 1138–1150.

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T. and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–838.

Iatsenko, I., Sinha, A., Rödelsperger, C. and Sommer, R. J. (2013). New role for DCR-1/dicer in *Caenorhabditis elegans* innate immunity against the highly virulent bacterium Bacillus thuringiensis DB27. *Infect. Immun.* **81**, 3942–3957.

Irazoqui, J. E. and Ausubel, F. M. (2010). 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: *Caenorhabditis elegans* as a model to study tissues involved in host immunity and microbial pathogenesis. *Clin. Exp. Immunol.* **160**, 48–57.

Irazoqui, J. E., Ng, A., Xavier, R. J. and Ausubel, F. M. (2008). Role for beta-catenin and HOX transcription factors in *Caenorhabditis elegans* and mammalian host epithelial-pathogen interactions. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17469–17474.

Irie, T., Muta, T. and Takeshige, K. (2000). TAK1 mediates an activation signal from toll-like receptor(s) to nuclear factor-κB in lipopolysaccharide-stimulated macrophages. *FEBS Letters* **467**, 160–164.

Iyer, L. M., Koonin, E. V. and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct. Biol.* **3**, 1.

Jansson, H. B. (1994). Adhesion of Conidia of Drechmeria coniospora to *Caenorhabditis elegans* Wild Type and Mutants. *J. Nematol.* **26**, 430–435.

Jiang, H., Franz, C. J., Wu, G., Renshaw, H., Zhao, G., Firth, A. E. and Wang, D. (2014). Orsay virus utilizes ribosomal frameshifting to express a novel protein that is incorporated into virions. *Virology* **450-451**, 213–221.

Jokela, J., Dybdahl, M. F. and Lively, C. M. (2009). The maintenance of sex, clonal dynamics, and host-parasite coevolution in a mixed population of sexual and asexual snails. *Am. Nat.* **174 Suppl 1**, S43–53.

Jovelin, R., Ajie, B. C. and Phillips, P. C. (2003). Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus *Caenorhabditis*. *Mol. Ecol.* **12**, 1325–1337.

Kacsoh, B. Z., Lynch, Z. R., Mortimer, N. T. and Schlenke, T. A. (2013). Fruit flies medicate offspring after seeing parasites. *Science* **339**, 947–950.

Kaltz, O. and Shykoff, J. A. (1998). Local adaptation in host-parasite systems. *Heredity* **81**, 361–370.

Kause, A. and Ødegård, J. (2012). The genetic analysis of tolerance to infections: a review. *Front. Gene* **3**, 262.

Kawli, T. and Tan, M.-W. (2008). Neuroendocrine signals modulate the innate immunity of *Caenorhabditis elegans* through insulin signaling. *Nat. Immunol.* **9**, 1415–1424.

Kerr, J. F., Wyllie, A. H. and Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br. J. Cancer* **26**, 239–257.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. and Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659.

Kim, D. H., Feinbaum, R., Alloing, G., Emerson, F. E., Garsin, D. A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M.-W., et al. (2002). A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297**, 623–626.

Kim, D. H., Liberati, N. T., Mizuno, T., Inoue, H., Hisamoto, N., Matsumoto, K. and Ausubel, F. M. (2004). Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10990–10994.

Kimble, J. and Sharrock, W. J. (1983). Tissue-specific synthesis of yolk proteins in *Caenorhabditis elegans*. *Dev. Biol.* **96**, 189–196.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kiontke, K. C., Félix, M.-A., Ailion, M., Rockman, M. V., Braendle, C., Pénigault, J.-B. and Fitch, D. H. A. (2011). A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new species from rotting fruits. *BMC Evol. Biol.* **11**, 339.

Klassen, G. (1992). Coevolution - a History of the Macroevolutionary Approach to Studying Host-Parasite Associations. *J. Parasitol.* **78**, 573–587.

Knight, S. W. and Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269–2271.

Koch, R., van Luenen, H. G., van der Horst, M., Thijssen, K. L. and Plasterk, R. H. (2000). Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res*. **10**, 1690–1696.

Kominos, S. D., Copeland, C. E., Grosiak, B. and Postic, B. (1972). Introduction of Pseudomonas aeruginosa into a hospital via vegetables. *Appl Microbiol* **24**, 567–570.

Kudlow, B. A., Zhang, L. and Han, M. (2012). Systematic analysis of tissue-restricted miRISCs reveals a broad role for microRNAs in suppressing basal activity of the *C. elegans* pathogen response. *Mol. Cell* **46**, 530–541.

Labouesse, M. Epithelial junctions and attachments (January 13, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.56.1, http://www.wormbook.org.

Laine, A.-L. (2009). Role of coevolution in generating biological diversity: spatially divergent selection trajectories. *J. Exp. Bot*. **60**, 2957–2970.

Lambrechts, L., Quillery, E., Noël, V., Richardson, J. H., Jarman, R. G., Scott, T. W. and Chevillon, C. (2013). Specificity of resistance to dengue virus isolates is associated with genotypes of the mosquito antiviral gene Dicer-2. *Proc. Biol. Sci*. **280**, 20122437.

Lambshead, P. J. D. (1993). Recent developments in marine benthic biodiversity research. *Oceanis* **19**, 5–24.

Langevin, C., Aleksejeva, E., Passoni, G., Palha, N., Levraud, J.-P. and Boudinot, P. (2013). The antiviral innate immune response in fish: evolution and conservation of the IFN system. *J. Mol. Biol*. **425**, 4904–4920.

La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P., Koonin, E., et al. (2008). The virophage as a unique parasite of the giant mimivirus. *Nature* **455**, 100–104.

Lazzaro, B. P., Sceurman, B. K. and Clark, A. G. (2004). Genetic basis of natural variation in *D. melanogaster* antibacterial immunity. *Science* **303**, 1873–1876.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854.

Lee, Y., Jeon, K., Lee, J.-T., Kim, S. and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*. **21**, 4663–4670.

Lee, R. C., Hammell, C. M. and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans*. *RNA* **12**, 589–597.

Lee, H., Choi, M., Lee, D., Kim, H., Hwang, H., Kim, H., Park, S., Paik, Y. and Lee, J. (2012a). Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nat. Neurosci.* **15**, 107–112.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D. and Mello, C. C. (2012b). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **150**, 78–87.

Lemaitre, B. and Hoffmann, J. (2007). The Host Defense of *Drosophila melanogaster*. *Annual Review of Immunology* **25**, 697–743.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983.

Leuckart, R. (1879). *Allgemeine Naturgeschichte der Parasiten mit besonderer Berücksichtigung der bei dem Menschen schmarotzenden Arten*. Leipzig, Heidelberg: C.F. Winter.

Leulier, F., Vidal, S., Saigo, K., Ueda, R. and Lemaitre, B. (2002). Inducible Expression of Double-Stranded RNA Reveals a Role for dFADD in the Regulation of the Antibacterial Response in Drosophila Adults. *Current Biology* **12**, 996–1000.

Levashina, E. A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J. A. and Reichhart, J. M. (1999). Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science* **285**, 1917–1919.

Lewontin, R. C. and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175–195.

Li, Y., Lu, J., Han, Y., Fan, X. and Ding, S.-W. (2013). RNA interference functions as an antiviral immunity mechanism in mammals. *Science* **342**, 231–234.

Liberati, N. T., Fitzgerald, K. A., Kim, D. H., Feinbaum, R., Golenbock, D. T. and Ausubel, F. M. (2004). Requirement for a conserved Toll/interleukin-1 resistance domain protein in the *Caenorhabditis elegans* immune response. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6593–6598.

Lively, C. M. and Dybdahl, M. F. (2000). Parasite adaptation to locally common host genotypes. *Nature* **405**, 679–681.

Lu, R., Maduro, M., Li, F., Li, H. W., Broitman-Maduro, G., Li, W. X. and Ding, S. W. (2005). Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* **436**, 1040–1043.

Lu, R., Yigit, E., Li, W.-X. and Ding, S.-W. (2009). An RIG-I-Like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in *Caenorhabditis elegans*. *PLoS Pathog*. **5**, e1000286.

Luijckx, P., Fienberg, H., Duneau, D. and Ebert, D. (2013). A matching-allele model explains host resistance to parasites. *Curr. Biol*. **23**, 1085–1088.

Lukeš, J., Skalický, T., Týč, J., Votýpka, J. and Yurchenko, V. (2014). Evolution of parasitism in kinetoplastid flagellates. *Mol. Biochem. Parasitol*.

Luna, E. and Ton, J. (2012). The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant Signal Behav* **7**, 615–618.

Luo, M.-F., Pan, H.-J., Liu, Z.-J. and Li, M. (2012). Balancing selection and genetic drift at major histocompatibility complex class II genes in isolated populations of golden snub-nosed monkey (Rhinopithecus roxellana). *BMC Evol. Biol*. **12**, 207.

MacRae, I. J. and Doudna, J. A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr. Opin. Struct. Biol*. **17**, 138–145.

Magwire, M. M., Bayer, F., Webster, C. L., Cao, C. and Jiggins, F. M. (2011). Successive increases in the resistance of Drosophila to viral infection through a transposon insertion followed by a Duplication. *PLoS Genet*. **7**, e1002337.

Magwire, M. M., Fabian, D. K., Schweyen, H., Cao, C., Longdon, B., Bayer, F. and Jiggins, F. M. (2012). Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS Genet*. **8**, e1003057.

Maillard, P. V., Ciaudo, C., Marchais, A., Li, Y., Jay, F., Ding, S. W. and Voinnet, O. (2013). Antiviral RNA interference in mammalian cells. *Science* **342**, 235–238.

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**, 7.

Mallo, G. V., Kurz, C. L., Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y. and Ewbank, J. J. (2002). Inducible antibacterial defense system in *C. elegans*. *Curr. Biol*. **12**, 1209–1214.

May, R. (1988). How Many Species Are There on Earth. *Science* **241**, 1441–1449.

McDonald, J. H. and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**, 652–654.

McGhee, J. D. (2007). The *C. elegans* intestine (March 27, 2007), *WormBook*, ed. The *C.*

elegans Research Community, WormBook, doi/10.1895/wormbook.1.133.1, http://www.wormbook.org.

Means, T. K., Mylonakis, E., Tampakakis, E., Colvin, R. A., Seung, E., Puckett, L., Tai, M. F., Stewart, C. R., Pukkila-Worley, R., Hickman, S. E., et al. (2009). Evolutionarily conserved recognition and innate immunity to fungal pathogens by the scavenger receptors SCARF1 and CD36. *J. Exp. Med.* **206**, 637–653.

Medzhitov, R., Preston-Hurlburt, P., Kopp, E., Stadlen, A., Chen, C., Ghosh, S. and Janeway Jr., C. A. (1998). MyD88 Is an Adaptor Protein in the hToll/IL-1 Receptor Family Signaling Pathways. *Molecular Cell* **2**, 253–258.

Meldal, B. H. M., Debenham, N. J., De Ley, P., De Ley, I. T., Vanfleteren, J. R., Vierstraete, A. R., Bert, W., Borgonie, G., Moens, T., Tyler, P. A., et al. (2007). An improved molecular phylogeny of the Nematoda with special emphasis on marine taxa. *Mol. Phylogenet. Evol.* **42**, 622–636.

Meng, Z. N., Yang, S., Fan, B., Wang, L. and Lin, H. R. (2012). Genetic variation and balancing selection at MHC class II exon 2 in cultured stocks and wild populations of orange-spotted grouper (*Epinephelus coioides*). *Genet. Mol. Res.* **11**, 3869–3881.

Morel, J. B. and Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death Differ.* **4**, 671–683.

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H. and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**, 277–283.

Muzio, M., Ni, J., Feng, P. and Dixit, V. M. (1997). IRAK (Pelle) Family Member IRAK-2 and MyD88 as Proximal Mediators of IL-1 Signaling. *Science* **278**, 1612–1615.

Nappi, A. J. and Christensen, B. M. (2005). Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. *Insect Biochem. Mol. Biol.* **35**, 443–459.

Nappi, A. J. and Ottaviani, E. (2000). Cytotoxicity and cytotoxic molecules in invertebrates. *Bioessays* **22**, 469–480.

Nappi, A. J. and Vass, E. (1993). Melanogenesis and the generation of cytotoxic molecules during insect cellular immune reactions. *Pigment Cell Res.* **6**, 117–126.

Ng, A. C. Y., Eisenberg, J. M., Heath, R. J. W., Huett, A., Robinson, C. M., Nau, G. J. and Xavier, R. J. (2011). Human leucine-rich repeat proteins: a genome-wide bioinformatic categorization and functional analysis in innate immunity. *Proc. Natl. Acad. Sci. U.S.A.* **108 Suppl 1**, 4631–4638.

Nicholas, H. R. and Hodgkin, J. (2004). The ERK MAP kinase cascade mediates tail swelling and a protective response to rectal infection in *C. elegans*. *Curr. Biol.* **14**, 1256–1261.

Ninomiya-Tsuji, J., Kishimoto, K., Hiyama, A., Inoue, J., Cao, Z. and Matsumoto, K. (1999). The kinase TAK1 can activate the NIK-IκB as well as the MAP kinase cascade in the IL-1 signalling pathway. *Nature* **398**, 252–256.

Nonaka, M. and Kimura, A. (2006). Genomic view of the evolution of the complement system. *Immunogenetics* **58**, 701–713.

Nonaka, M. and Miyazawa, S. (2002). Evolution of the initiating enzymes of the complement system. *Genome Biol.* **3**, REVIEWS1001.

Obbard, D. J., Jiggins, F. M., Halligan, D. L. and Little, T. J. (2006). Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr. Biol.* **16**, 580–585.

Obbard, D. J., Callister, D. M., Jiggins, F. M., Soares, D. C., Yan, G. and Little, T. J. (2008). The evolution of TEP1, an exceptionally polymorphic immunity gene in Anopheles gambiae. *BMC Evol. Biol.* **8**, 274.

Obbard, D. J., Welch, J. J., Kim, K.-W. and Jiggins, F. M. (2009). Quantifying adaptive evolution in the Drosophila immune system. *PLoS Genet.* **5**, e1000698.

Oncogenic DNA viruses (2001). , *Published online: 27 November 2001; | doi:10.1038/sj.onc.1204958* **20**,.

O'Rourke, D., Baban, D., Demidova, M., Mott, R. and Hodgkin, J. (2006). Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res.* **16**, 1005–1016.

Orr, H. A. (2002). The population genetics of adaptation: the adaptation of DNA sequences. *Evolution* **56**, 1317–1330.

Osborne, S. E., Leong, Y. S., O'Neill, S. L. and Johnson, K. N. (2009). Variation in antiviral protection mediated by different Wolbachia strains in *Drosophila simulans*. *PLoS Pathog.* **5**, e1000656.

Pak, J. and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315**, 241–244.

Parker, G. S., Eckert, D. M. and Bass, B. L. (2006). RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA-Publ. RNA Soc.* **12**, 807–818.

Parsons, L. M. and Cipollo, J. (2014). Oral ingestion of *Microbacterium nematophilum*

leads to anal-region infection in *Caenorhabditis elegans*. *Microbes and Infection* **16**, 356–361.

Perlman, S. J., Jaenike, J. and Noor, M. (2003). Infection success in novel hosts: an experimental and phylogenetic study of *Drosophila*-parasitic nematodes. *Evolution* **57**, 544–557.

Petersen, C., Dirksen, P., Prahl, S., Strathmann, E. A. and Schulenburg, H. (2014). The prevalence of *Caenorhabditis elegans* across 1.5 years in selected North German locations: the importance of substrate type, abiotic parameters, and *Caenorhabditis* competitors. *BMC Ecol*. **14**, 4.

Pham, L. N., Dionne, M. S., Shirasu-Hiza, M. and Schneider, D. S. (2007). A specific primed immune response in *Drosophila* is dependent on phagocytes. *PLoS Pathog*. **3**, e26.

Philippe, H., Lartillot, N. and Brinkmann, H. (2005). Multigene analyses of bilaterian animals corroborate the monophyly of *Ecdysozoa*, *Lophotrochozoa*, and *Protostomia*. *Mol*. *Biol*. *Evol*. **22**, 1246–1253.

Poulin, R. (1998). Large-scale patterns of host use by parasites of freshwater fishes. *Ecol*. *Lett*. **1**, 118–128.

Poulin, R. and Fitzgerald, G. (1989). Risk of Parasitism and Microhabitat Selection in Juvenile Sticklebacks. *Can*. *J*. *Zool*.-*Rev*. *Can*. *Zool*. **67**, 14–18.

Powell, J. R., Kim, D. H. and Ausubel, F. M. (2009). The G protein-coupled receptor FSHR-1 is required for the *Caenorhabditis elegans* innate immune response. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **106**, 2782–2787.

Pradel, E., Zhang, Y., Pujol, N., Matsuyama, T., Bargmann, C. I. and Ewbank, J. J. (2007). Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proc*. *Natl*. *Acad*. *Sci*. *U*.*S*.*A*. **104**, 2295–2300.

Pujol, N., Link, E. M., Liu, L. X., Kurz, C. L., Alloing, G., Tan, M. W., Ray, K. P., Solari, R., Johnson, C. D. and Ewbank, J. J. (2001). A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Curr*. *Biol*. **11**, 809–821.

Pujol, N., Cypowyj, S., Ziegler, K., Millet, A., Astrain, A., Goncharov, A., Jin, Y., Chisholm, A. D. and Ewbank, J. J. (2008a). Distinct innate immune responses to infection and wounding in the *C*. *elegans* epidermis. *Curr*. *Biol*. **18**, 481–489.

Pujol, N., Zugasti, O., Wong, D., Couillault, C., Kurz, C. L., Schulenburg, H. and Ewbank, J. J. (2008b). Anti-fungal innate immunity in *C*. *elegans* is enhanced by evolutionary diversification of antimicrobial peptides. *PLoS Pathog*. **4**, e1000105.

Pujol, N., Davis, P. A. and Ewbank, J. J. (2012). The Origin and Function of Anti-Fungal Peptides in *C. elegans*: Open Questions. *Front Immunol* **3**, 237.

Putnam, F. W. (1969). Immunoglobulin structure: variability and homology. *Science* **163**, 633–644.

Råberg, L., Sim, D. and Read, A. F. (2007). Disentangling Genetic Variation for Resistance and Tolerance to Infectious Diseases in Animals. *Science* **318**, 812–814.

Rajewsky, K. (1996). Clonal selection and learning in the antibody system. *Nature* **381**, 751–758.

Rechavi, O., Minevich, G. and Hobert, O. (2011). Transgenerational inheritance of an acquired small RNA-based antiviral response in *C. elegans*. *Cell* **147**, 1248–1256.

Reddy, K. C., Andersen, E. C., Kruglyak, L. and Kim, D. H. (2009). A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* **323**, 382–384.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R. and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906.

Ren, M., Feng, H., Fu, Y., Land, M. and Rubin, C. S. (2009). Protein kinase D is an essential regulator of *C. elegans* innate immunity. *Immunity* **30**, 521–532.

Rimer, J., Cohen, I. R. and Friedman, N. (2014). Do all creatures possess an acquired immune system of some sort? *BioEssays* **36**, 273–281.

Rockman, M. V. and Kruglyak, L. (2009). Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet*. **5**, e1000419.

Roeder, T., Stanisak, M., Gelhaus, C., Bruchhaus, I., Grötzinger, J. and Leippe, M. (2010). Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. *Dev. Comp. Immunol*. **34**, 203–209.

ROSS, A. F. (1961). Systemic acquired resistance induced by localized virus infections in plants. *Virology* **14**, 340–358.

Roth, O., Sadd, B. M., Schmid-Hempel, P. and Kurtz, J. (2009). Strain-specific priming of resistance in the red flour beetle, *Tribolium castaneum*. *Proc. Biol. Sci*. **276**, 145–151.

Rothenbuhler, W. C. (1964). Behavior genetics of nest cleaning in honey bees. iv. responses of f1 and backcross generations to disease-killed blood. *Am. Zool*. **4**, 111–123.

Rottschaefer, S. M., Riehle, M. M., Coulibaly, B., Sacko, M., Niaré, O., Morlais, I., Traoré, S. F., Vernick, K. D. and Lazzaro, B. P. (2011). Exceptional Diversity, Maintenance of Polymorphism, and Recent Directional Selection on the APL1 Malaria Resistance Genes of *Anopheles gambiae*. *PLoS Biol* **9**, e1000600.

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**, 1193–1207.

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z. P., Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918.

Sadd, B. M. and Schmid-Hempel, P. (2006). Insect immunity shows specificity in protection upon secondary pathogen exposure. *Curr. Biol.* **16**, 1206–1210.

Sadd, B. M. and Schmid-Hempel, P. (2007). Facultative but persistent trans-generational immunity via the mother's eggs in bumblebees. *Curr. Biol.* **17**, R1046–1047.

Sahl, H.-G., Pag, U., Bonness, S., Wagner, S., Antcheva, N. and Tossi, A. (2005). Mammalian defensins: structures and mechanism of antibiotic activity. *J. Leukoc. Biol.* **77**, 466–475.

Sarkies, P., Ashe, A., Le Pen, J., McKie, M. A. and Miska, E. A. (2013). Competition between virus-derived and endogenous small RNAs regulates gene expression in *Caenorhabditis elegans*. *Genome Res.* **23**, 1258–1270.

Schlee, M. (2013). Master sensors of pathogenic RNA – RIG-I like receptors. *Immunobiology* **218**, 1322–1335.

Schott, D. H., Cureton, D. K., Whelan, S. P. and Hunter, C. P. (2005). An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18420–18424.

Schulenburg, H. and Ewbank, J. J. (2004). Diversity and specificity in the interaction between *Caenorhabditis elegans* and the pathogen *Serratia marcescens*. *BMC Evol. Biol.* **4**, 49.

Schulenburg, H., Kurz, C. L. and Ewbank, J. J. (2004). Evolution of the innate immune system: the worm perspective. *Immunol. Rev.* **198**, 36–58.

Schulenburg, H., Hoeppner, M. P., Weiner, J., 3rd and Bornberg-Bauer, E. (2008). Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* **213**, 237–250.

Schulte, R. D., Hasert, B., Makus, C., Michiels, N. K. and Schulenburg, H. (2012). Increased responsiveness in feeding behaviour of *Caenorhabditis elegans* after experimental coevolution with its microparasite *Bacillus thuringiensis*. *Biol. Lett.* **8**, 234–236.

Sharma, S., Sachdeva, P. and Virdi, J. S. (2003). Emerging water-borne pathogens. *Appl. Microbiol. Biotechnol.* **61**, 424–428.

Shen, J., Araki, H., Chen, L., Chen, J.-Q. and Tian, D. (2006). Unique evolutionary mechanism in R-genes under the presence/absence polymorphism in *Arabidopsis thaliana*. *Genetics* **172**, 1243–1250.

Shen, Y., Wollam, J., Magner, D., Karalay, O. and Antebi, A. (2012). A steroid receptor-microRNA switch regulates life span in response to signals from the gonad. *Science* **338**, 1472–1476.

Sifri, C. D., Begun, J. and Ausubel, F. M. (2005). The worm has turned--microbial virulence modeled in *Caenorhabditis elegans*. *Trends Microbiol.* **13**, 119–127.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H. and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**, 465–476.

Simon, D. J., Madison, J. M., Conery, A. L., Thompson-Peer, K. L., Soskis, M., Ruvkun, G. B., Kaplan, J. M. and Kim, J. K. (2008). The microRNA miR-1 regulates a MEF-2-dependent retrograde signal at neuromuscular junctions. *Cell* **133**, 903–915.

Sivasundar, A. and Hey, J. (2003). Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics* **163**, 147–157.

Sivasundar, A. and Hey, J. (2005). Sampling from natural populations with RNAI reveals high outcrossing and population structure in *Caenorhabditis elegans*. *Curr. Biol.* **15**, 1598–1602.

Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N. and Maine, E. M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Curr. Biol.* **10**, 169–178.

Sterken, M. G., Snoek, L. B., Bosman, K. J., Daamen, J., Riksen, J. A. G., Bakker, J., Pijlman, G. P. and Kammenga, J. E. (2014). A heritable antiviral RNAi response limits Orsay virus infection in *Caenorhabditis elegans* N2. *PLoS ONE* **9**,

e89760.

Stuart, L. M. and Ezekowitz, R. A. (2008). Phagocytosis and comparative innate immunity: learning on the fly. *Nat Rev Immunol* **8**, 131–141.

Styer, K. L., Singh, V., Macosko, E., Steele, S. E., Bargmann, C. I. and Aballay, A. (2008). Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* **322**, 460–464.

Swan, K. A., Curtis, D. E., McKusick, K. B., Voinov, A. V., Mapa, F. A. and Cancilla, M. R. (2002). High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res*. **12**, 1100–1105.

Tabara, H., Yigit, E., Siomi, H. and Mello, C. C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* **109**, 861–871.

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **123**, 585.

Thomas, W. K. and Wilson, A. C. (1991). Mode and tempo of molecular evolution in the nematode *Caenorhabditis*: cytochrome oxidase II and calmodulin sequences. *Genetics* **128**, 269–279.

Troemel, E. R., Félix, M.-A., Whiteman, N. K., Barrière, A. and Ausubel, F. M. (2008). Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. *PLoS Biol*. **6**, 2736–2752.

Van Valen, L. (1973). A new evolutionary law. *Evol. Theory* **1**, 1–30.

Vasale, J. J., Gu, W., Thivierge, C., Batista, P. J., Claycomb, J. M., Youngman, E. M., Duchaine, T. F., Mello, C. C. and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3582–3587.

Vastenhouw, N. L., Brunschwig, K., Okihara, K. L., Müller, F., Tijsterman, M. and Plasterk, R. H. A. (2006). Gene expression: long-term gene silencing by RNAi. *Nature* **442**, 882.

Voight, B. F., Kudaravalli, S., Wen, X. and Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biol*. **4**, e72.

Walossek, D. and Müller, K. J. (1994). Pentastomid parasites from the Lower Palaeozoic of Sweden. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh* **85**, 1–37.

Wambua, S., Mwangi, T. W., Kortok, M., Uyoga, S. M., Macharia, A. W., Mwacharo, J. K., Weatherall, D. J., Snow, R. W., Marsh, K. and Williams, T.

N. (2006). The effect of alpha+-thalassaemia on the incidence of malaria and other diseases in children living on the coast of Kenya. *PLoS Med*. **3**, e158.

Wang, G. and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Curr. Biol*. **18**, 861–867.

Wang, D., Fang, L., Li, T., Luo, R., Xie, L., Jiang, Y., Chen, H. and Xiao, S. (2008). Molecular cloning and functional characterization of porcine IFN-beta promoter stimulator 1 (IPS-1). *Vet. Immunol. Immunopathol*. **125**, 344–353.

Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., Beeson, K. Y., Bibbs, L., Bolanos, R., Keller, M., et al. (2003). The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. *Proc. Natl. Acad. Sci. U.S.A*. **100**, 12984–12988.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* **7**, 256–276.

Wayne, M. L., Pienaar, J., Telonis-Scott, M., Sylvestre, L.-S., Nuzhdin, S. V. and McIntyre, L. M. (2011). Expression of defense genes in Drosophila evolves under a different selective regime from expression of other genes. *Evolution* **65**, 1068–1078.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R. A., Moss, S. M. M., Cording, A. C., Ahringer, J., Martinez-Perez, E. and Miska, E. A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in *C. elegans*. *Genes Dev*. **28**, 783–796.

Weiss, M. R. (2006). Defecation behavior and ecology of insects. *Annu. Rev. Entomol*. **51**, 635–661.

Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. and Bull, J. J. (1999). Different trajectories of parallel evolution during viral adaptation. *Science* **285**, 422–424.

Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. and Plasterk, R. H. (2001). Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet*. **28**, 160–164.

Wilfert, L. and Jiggins, F. M. (2010). Host-parasite coevolution: genetic variation in a virus population and the interaction with a host gene. *J. Evol. Biol*. **23**, 1447–1455.

Wilfert, L. and Schmid-Hempel, P. (2008). The genetic architecture of susceptibility to parasites. *BMC Evol. Biol*. **8**, 187.

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. A., Chow, M. and Machaca, K. (2005). RNA interference is an antiviral defence mechanism in *Caenorhabditis*

elegans. *Nature* **436**, 1044–1047.

Wilson, K., Knell, R., Boots, M. and Koch-Osborne, J. (2003). Group living and investment in immune defence: an interspecific analysis. *Journal of Animal Ecology* **72**, 133–143.

Wong, D., Bazopoulou, D., Pujol, N., Tavernarakis, N. and Ewbank, J. J. (2007). Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. *Genome Biol*. **8**, R194.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics* **16**, 97–159.

Xu, P., Chen, F., Mannas, J. P., Feldman, T., Sumner, L. W. and Roossinck, M. J. (2008). Virus infection improves drought tolerance. *New Phytologist* **180**, 911–921.

Yasukochi, Y. and Satta, Y. (2013). Current perspectives on the intensity of natural selection of MHC loci. *Immunogenetics* **65**, 479–483.

Yeaman, M. R. and Yount, N. Y. (2007). Unifying themes in host defence effector polypeptides. *Nat. Rev. Microbiol*. **5**, 727–740.

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C.-C. G., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J. and Mello, C. C. (2006). Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127**, 747–757.

Zasloff, M. (2002). Antimicrobial Peptides in Health and Disease. *New England Journal of Medicine* **347**, 1199–1200.

Zhang, H., Yoshida, S., Aizawa, T., Murakami, R., Suzuki, M., Koganezawa, N., Matsuura, A., Miyazawa, M., Kawano, K., Nitta, K., et al. (2000). In vitro antimicrobial properties of recombinant ASABF, an antimicrobial peptide isolated from the nematode *Ascaris suum*. *Antimicrob. Agents Chemother*. **44**, 2701–2705.

Zhang, L., Ding, L., Cheung, T. H., Dong, M.-Q., Chen, J., Sewell, A. K., Liu, X., Yates, J. R. and Han, M. (2007). Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol. Cell* **28**, 598–613.

Ziegler, K., Kurz, C. L., Cypowyj, S., Couillault, C., Pophillat, M., Pujol, N. and Ewbank, J. J. (2009). Antifungal innate immunity in *C. elegans*: PKCdelta links G protein signaling and a conserved p38 MAPK cascade. *Cell Host Microbe* **5**, 341–352.

Zou, J., Chang, M., Nie, P. and Secombes, C. J. (2009). Origin and evolution of the RIG-I like RNA helicase gene family. *BMC Evol. Biol*. **9**, 85.

Zugasti, O. and Ewbank, J. J. (2009). Neuroimmune regulation of antimicrobial peptide expression by a noncanonical TGF-beta signaling pathway in *Caenorhabditis elegans* epidermis. *Nat. Immunol.* **10**, 249–256.

Zugasti, O., Bose, N., Squiban, B., Belougne, J., Kurz, C. L., Schroeder, F. C., Pujol, N. and Ewbank, J. J. (2014). Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. *Nat. Immunol.* **15**, 833–838.

ANNEXE

démontré, chez la souris, qu'il favorise l'expression de l'allèle paternel en réduisant la transcription de l'antisens par un mécanisme distinct de la méthylation du centre de contrôle et qui reste à découvrir. L'identification du mécanisme d'action du Topo *via* cette activation de l'expression de l'allèle paternel du gène *UBE3A* pourrait permettre le développement de nouvelles approches pharmacologiques pour traiter les patients atteints du syndrome d'Angelman.

De fait, de nombreuses interrogations persistent avant que l'on puisse affirmer que ces résultats obtenus chez la souris peuvent s'appliquer à des humains. Par exemple, les effets de cette molécule sur l'expression du gène *Ube3a* sont-ils stables à long terme chez la souris ? Quels seraient les effets du Topo sur l'expression clinique du syndrome d'Angelman chez l'animal ? Quelles seraient les doses nécessaires ? Y aurait-il des effets secondaires ? Le traitement serait-il sûr et efficace chez des humains ? À quel type de patient l'administrer ? Ce n'est que lorsqu'une démarche scientifique aura permis de clarifier ces aspects que l'on pourra envisager dans quelle mesure des molécules comme le Topo seraient à même de soulager les symptômes des individus atteints du syndrome d'Angelman.

Ce sont de nouvelles questions qui soulèvent l'enthousiasme des équipes qui y travaillent tout en n'oubliant pas que l'éthique de la science médicale impose de suivre des chemins balisés et progressifs qui ne mènent aux essais cliniques qu'après obtention des réponses à toutes les questions préliminaires. ◊

A new Topo to targeted management of Angelman syndrome?

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Dan B. *Angelman syndrome.* London : Mac Keith Press, Wiley-Blackwell, 2008 : 256 p.
2. Kishino T, Lalande M, Wagstaff J. *UBE3A/E6-AP* mutations cause Angelman syndrome. *Nat Genet* 1997 ; 15 : 70-3.
3. Matsuura T, Sutcliffe JS, Fang P, *et al. De novo* truncating mutations in E6-AP ubiquitin-protein ligase gene (*UBE3A*) in Angelman syndrome. *Nat Genet* 1997 ; 15 : 74-7.
4. Greer PL, Hanayama R, Bloodgood BL, *et al.* The Angelman syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell* 2010 ; 140 : 704-16.
5. Weeber EJ, Jiang YH, Elgersma Y, *et al.* Derangements of hippocampal calcium/calmodulin-dependent protein kinase II in a mouse model for Angelman mental retardation syndrome. *J Neurosci* 2003 ; 23 : 2634-44.
6. Cheron G, Servais L, Wagstaff J, Dan B. Fast cerebellar oscillation associated with ataxia in a mouse model of Angelman syndrome. *Neuroscience* 2005 ; 130 : 631-7.
7. Rougeulle C, Cardoso C, Fontes M, *et al.* An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat Genet* 1998 ; 19 : 15-6.
8. Bird LM, Tan WH, Bacino CA, *et al.* A therapeutic trial of pro-methylation dietary supplements in Angelman syndrome. *Am J Med Genet* 2011 ; 155A : 2956-63.
9. Huang HS, Allen JA, Mabb AM, *et al.* Topoisomerase inhibitors unsilence the dormant allele of *Ube3a* in neurons. *Nature* 2011 ; 481 : 185-9.

NOUVELLE

Transmission multigénérationnelle de l'interférence à l'ARN chez le nématode *Caenorhabditis elegans*

Tony Bélicard, Marie-Anne Félix

Institut de biologie de l'École normale supérieure (ENS), CNRS-Inserm-ENS, 46, rue d'Ulm, 75230 Paris Cedex 05, France.

felix@biologie.ens.fr

› Chez de nombreux organismes, l'introduction artificielle d'un fragment d'ARN double brin (ARNdb) correspondant à la séquence d'un gène inactive spécifiquement l'expression de ce gène. Ce phénomène d'interférence à l'ARN (ARNi) a été mis en évidence chez le nématode *C. elegans* [1], un petit ver d'un millimètre qui est devenu un organisme modèle de laboratoire. L'ARNi est devenu un outil pratique d'inactivation des gènes, et cette découverte a valu le prix Nobel de médecine à Andy Fire et Craig Mello en 2006. Des résultats récents montrent que la réponse ARNi peut être transmise de génération en génération en affectant l'état des histones et la régulation transcriptionnelle du gène. Le rôle physiologique de cette transmission héréditaire pourrait-il être la défense antivirale ?

Inactivation post-transcriptionnelle… mais aussi transcriptionnelle

Le mécanisme de l'ARNi a été amplement étudié depuis 15 ans [2]. Chez *C. elegans*, l'injection ou l'ingestion d'un ARNdb peut déclencher la réponse d'interférence. La ribonucléase Dicer/DCR-1 clive dans le cytoplasme l'ARNdb en fragments de 21 nucléotides, nommés siARN (si pour *small interfering*) *(Figure 1)*. Ces siARN-1 recrutent et guident la protéine Argonaute RDE-1 vers une autre copie de l'ARN messager complémentaire (ARNm) qui sert de matrice à la synthèse de siARN 2° antisens par une ARN polymérase dépendante de l'ARN. Ces siARN 2° amplifiés lient alors d'autres protéines Argonaute, et ces complexes induisent la

DOI : 10.1051/medsci/2012286006

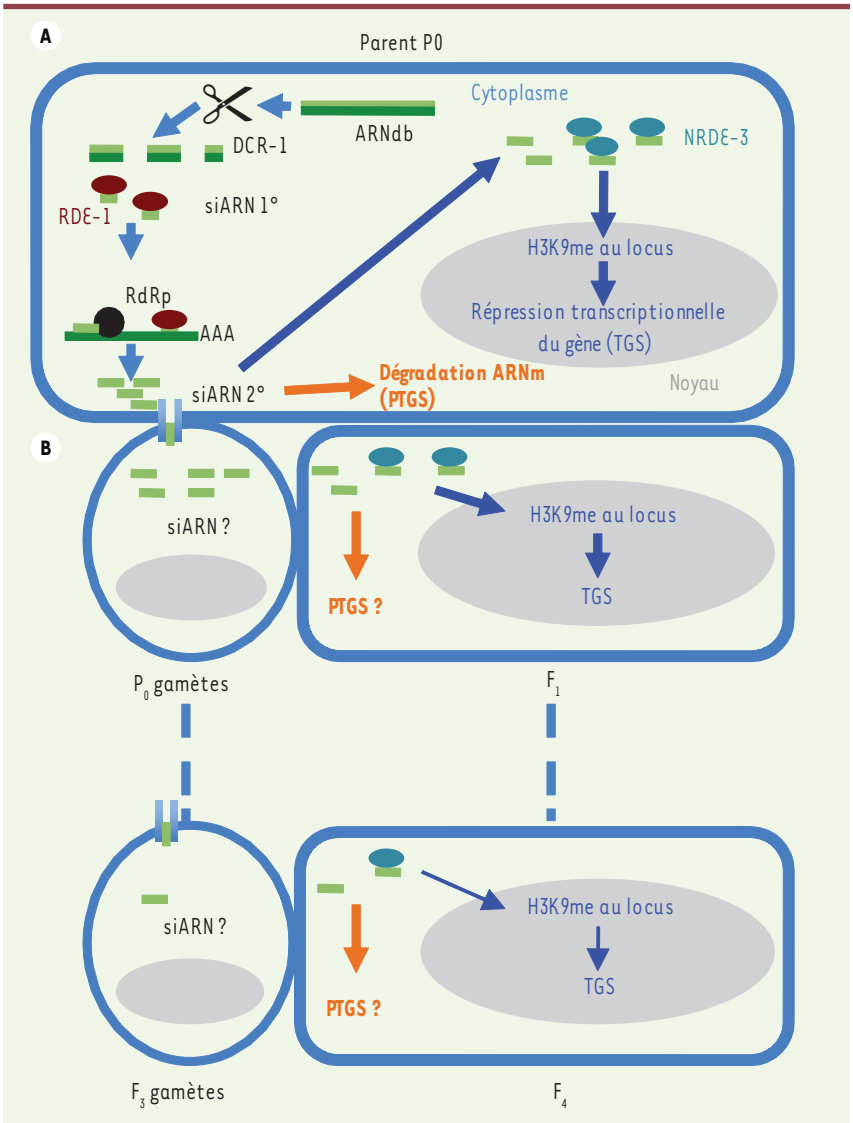


Figure 1. ***Mécanismes transcriptionnels et post-transcriptionnels de l'extinction de gènes par l'ARN double-brin et transmission multigénérationnelle. A.*** Dans le parent P0 exposé à des ARNdb, deux mécanismes de répression de l'expression du gène correspondant peuvent être déclenchés. L'ARNdb (en vert) est fragmenté par l'enzyme DCR-1. Les petits siARN 1° produits (extrémité 5' monophosphate) sont pris en charge par l'Argonaute RDE-1. Ce complexe s'hybride à son ARNm complémentaire et produit des siARN antisens dits 2° (extrémité 5' triphosphate). Ces derniers vont, soit participer directement au clivage des ARNm correspondants (voie PTGS ou *post-transcriptional gene silencing*, en orange), soit être pris en charge par l'Argonaute NRDE-3. Le complexe siARN 2°/NRDE-3 entre dans le noyau et réprime la transcription du gène correspondant (voie TGS ou *transcriptional gene silencing*, en bleu). ***B.*** Transmission de l'extinction du gène dans les générations suivantes (F_1, etc.). Des siARN (vraisemblablement 1°) semblent être transmis dans les gamètes et, même en absence d'ARNdb exogène dans la descendance, permettent l'extinction des gènes ciblés *via* la méthylation de l'ADN. La quantité de siARN transmis décroît au fil des générations.

dégradation des ARNm correspondants. Ceci conduit à l'extinction post-transcriptionnelle (PTGS pour *post-transcriptional gene silencing*) du gène.

Un second mécanisme d'action des siARN 2° a été mis en évidence chez *C. elegans* par le laboratoire de Scott Kennedy. Les siARN 2° se lient à la protéine Argonaute NRDE-3, et le complexe entre dans le noyau. L'extinction du gène correspondant se fait alors au niveau de la transcription (TGS pour *transcriptional gene silencing*) *(Figure 1)*. Les gènes impliqués dans ce phénomène ont été identifiés par un crible génétique et appelés nrde (*nuclear-RNAi deficient*).

Extinction transgénérationnelle de certains gènes à partir d'un ARNdb spécifique

Chez *C. elegans*, l'ARNi peut, dans certains cas, persister pendant plusieurs générations, mimant les effets transgénérationnels de l'extinction de gènes connus chez certains champignons et plantes. Ainsi, chez *C. elegans*, l'extinction d'un transgène exprimant la protéine fluorescente GFP (*green fluorescent protein*) peut être transmise jusqu'à 20 générations après l'ingestion d'ARNdb spécifique de la GFP [3]. Cette transmission de l'extinction fonctionne pour certains gènes seulement (13/171 gènes testés sont ainsi éteints sur plusieurs générations) [3-5], sans que l'on comprenne pour l'instant ce qui les caractérise.

Une diminution de la quantité d'ARNdb injectée réduit la durée de l'extinction du gène, suggérant que la transmission de l'effet ARNi est diluée au fil des générations [5]. La même étude du laboratoire d'Andrew Fire révèle que les deux types de gamètes, mâle et femelle, sont capables de transmettre ces facteurs, mais que les spermatozoïdes le font plus efficacement.

Deux mécanismes de transmission transgénérationnelle

La transmission de l'extinction ARNi ne requiert pas que la descendance puisse elle même produire des siRNA 1°. En effet, cette transmission se fait normalement chez un descendant incapable de produire des siARN 1° (mutant *rde-1* ou *rde-4*) [4]. En revanche, les gènes nécessaires à l'amplification des siARN 2°, comme *mut-7* ou *rde-2*,

doivent, eux, être fonctionnels dans la descendance [4], comme doivent l'être aussi les gènes *nrde-2*, *nrde-3* et *nrde-4* nécessaires à la voie de régulation transcriptionnelle en aval des siARN 2° [6]. La transmission transgénérationnelle requiert donc, à chaque génération, la formation de siARN 2° et, en aval, une voie de répression transcriptionnelle intacte *(Figure 1)*.

La répression transcriptionnelle par les siARN 2° semble passer par des modifications d'histones. Les histones sont des protéines de la chromatine qui régulent le degré de condensation et l'accessibilité de cette dernière pour que s'enclenche la transcription. En effet, la voie nucléaire des siARN active la méthylation de la lysine 9 de l'histone 3 (H3K9me) [6], une modification qui inhibe localement la transcription. De plus, la transmission de l'extinction d'expression génique est abolie chez des mutants affectant la structure chromatinienne [3].

Gu *et al.* [7], du laboratoire d'Andrew Fire, ont récemment réussi à suivre la quantité de siARN 2° spécifiques et le taux de méthylation de H3K9 au locus correspondant chez un ver adulte ayant absorbé des ARNdb spécifiques et dans sa descendance pendant quatre générations. Dans la génération parentale (P0), la synthèse des siARN spécifiques est d'abord détectée, puis une marque H3K9me au niveau du gène ciblé est observée. Les siARN 2° et la méthylation sont ensuite détectés chez les descendants et ce jusqu'à trois générations dans cette étude, mais avec un marquage d'intensité décroissante au fil des générations [7].

Quel peut être le facteur transmis par les gamètes qui permet l'héritabilité de l'extinction transcriptionnelle ? Les modifications d'histones sont transmissibles lors des divisions cellulaires, mais elles sont généralement « réinitialisées » dans la lignée germinale chez *C. elegans*. De plus, l'extinction transcriptionnelle peut aussi s'appliquer à une copie du gène nouvellement introduite par croisement, et l'action des siARN 2° sur la méthylation des histones semble devoir être répétée à chaque génération. La transmission de siARN *via* les gamètes pourrait permettre une persistance et une réamplification de l'extinction transcriptionnelle de certains gènes au fil des générations. Pour l'instant, il n'est pas clair s'il s'agit de siARN 2° ou 1° (voir plus loin).

L'enjeu évolutif de la transmission de l'ARNi serait-il l'immunisation de la descendance contre des pathogènes naturels ?

En principe, la transmission transgénérationnelle du statut de l'expression des gènes permet à un parent de transmettre à sa descendance les conséquences qu'ont imprimées à son organisme les influences de l'environnement (effet « lamarckien »). Ceci peut concerner par exemple la défense antivirale. En effet, chez *C. elegans* comme chez les plantes et les insectes, la voie d'ARNi joue un rôle dans la défense antivirale - de nombreux virus ayant un génome de type ARNdb.

Le virus d'Orsay - découvert récemment - est le premier virus naturel de *C. elegans* [8]. Son génome s'apparente à celui des nodavirus et est composé de deux segments d'ARN simple brin positif (ARN1 et ARN2). L'ARN1 code pour une ARN polymérase dépendante de l'ARN (RdRP) qui synthétise le brin complémentaire des ARN viraux et, par conséquent, crée des ARNdb. Ceux-ci déclenchent alors une réponse siARN contre le génome viral. Les mutants *C. elegans* des gènes impliqués dans la réponse siARN sont plus sensibles au virus que le génotype sauvage de référence [8]. La réponse siARN est donc requise pour la défense antivirale chez *C. elegans*.

La réponse à un virus peut-elle se transmettre d'une génération à l'autre, et ainsi « immuniser » plusieurs générations ? Chez *C. elegans*, nous n'avons pas encore la réponse avec un virus infectieux comme le virus d'Orsay, mais un système artificiel a testé cette hypothèse.

Un système transgénique artificiel reconstituant la réplication du *Flock house virus* (FHV), un nodavirus d'arthropode qui n'infecte pas naturellement *C. elegans*, a été mis au point il y a quelques années [9]. Le génome du virus FHV est composé de deux ARN dont l'un code pour une RdRP et pour la protéine B2 qui inhibe la réponse ARNi de l'hôte. En remplaçant cette protéine B2 par la GFP et en construisant un ver transgénique dans lequel les séquences ADN correspondant à RdRP et B2 sont sous le contrôle d'un promoteur inductible par un choc thermique, la réplication de l'ARN1 viral dans le ver (qui est assurée ensuite de façon autonome par la RdRP) et la réponse d'extinction peuvent être reconstituées et suivies en mesurant l'expression ou l'extinction de la GFP [9]. Rechavi *et al.* [10] ont récemment étudié la transmissibilité intergénérationnelle de petits ARN antiviraux dérivés de l'expression d'un tel transgène. Comme c'est le cas pour un gène endogène ou la GFP (voir plus haut), l'extinction des ARN viraux peut se transmettre en absence de la machinerie de formation des siARN 1° chez les descendants (mutant *rde-1*), et peut se transmettre à une nouvelle copie du gène testé introduite par croisement. Cependant, cette réponse requiert à chaque génération l'amplification de siARN 2° - elle est abolie dans le mutant *rrf-1* affectant une RdRP du ver. Des siARN 1° correspondant aux séquences virales peuvent être détectés sur plusieurs générations, quoique en très petit nombre dès la troisième génération. Ceci suggère donc que la transmission peut s'effectuer *via* les siARN 1°, qui pourraient enclencher l'amplification *de novo* de siARN 2° à chaque génération.

Pour ce qui est de la portée de cette étude sur la transmission d'une défense antivirale, on peut regretter que les séquences virales y soient portées par un transgène artificiel. En effet,

il est possible que l'interférence agisse ici exclusivement au niveau transcriptionnel. Ce ne peut pas être le cas avec un virus relativement proche, tel que celui d'Orsay, dont le génome n'existe pas sous une forme ADN.

En conclusion

L'extinction d'un gène par l'ARNdb correspondant peut se transmettre au fil des générations chez des animaux, comme chez les plantes et les champignons. Chez *C. elegans*, elle semble utiliser l'amplification de petits ARN spécifiques et la répression transcriptionnelle à chaque génération. L'impact de ce processus de transmission sur l'adaptation phénotypique des organismes en milieu naturel reste cependant à démontrer. Il est possible que la cible principale de cette extinction multigénérationnelle soit les transposons insérés dans le génome. ◊

Multigenerational transmission of RNA interference in the nematode *Caenorhabditis elegans*

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

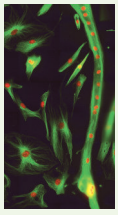
REMERCIEMENTS

L'équipe est financée par le CNRS, l'ENS, l'Inserm, et pour la recherche sur les virus, la subvention de l'ANR11 BSV3 01301. Nous remercions Valérie Robert (ENS Lyon) pour ses commentaires.

RÉFÉRENCES

1. Fire A, Xu S, Montgomery MK, *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998 ; 391 : 806-11.
2. Fischer SEJ. Small RNA-mediated gene silencing pathways in *C. elegans*. *Int J Biochem Cell Biol* 2010 ; 42 : 1306-15.
3. Vastenhouw NL, Brunschwig K, Okihara KL, *et al.* Gene expression: long-term gene silencing by RNAi. *Nature* 2006 ; 442 : 882.
4. Grishok A, Tabara H, Mello CC. Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* 2000 ; 287 : 2494-7.
5. Alcazar RM, Lin R, Fire AZ. Transmission dynamics of heritable silencing induced by double-stranded RNA in *Caenorhabditis elegans*. *Genetics* 2008 ; 180 : 1275-88.
6. Burton NO, Burkhart KB, Kennedy S. Nuclear RNAi maintains heritable gene silencing in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 2011 ; 108 : 19683-8.
7. Gu SG, Pak J, Guang S, *et al.* Amplification of siRNA in *Caenorhabditis elegans* generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. *Nat Genet* 2012 ; 44 : 157-64.
8. Félix M-A, Ashe A, Piffaretti J, *et al.* Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol* 2011 ; 9 : e1000586.
9. Lu R, Maduro M, Li F, *et al.* Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* 2005 ; 436 : 1040-3.
10. Rechavi O, Minevich G, Hobert O. Transgenerational inheritance of an acquired small RNA-based antiviral response in *C. elegans*. *Cell* 2011 ; 147 : 1248-56.

NOUVELLE



Positionnement du noyau dans les muscles

Nouvel acteur de la fonction musculaire

Vincent Gache[1], Bruno Cadot[1], Edgar R. Gomes[1,2]

[1]UMR S 787 Inserm, université Paris 6 Pierre et Marie Curie, 75634 Paris, France ; [2]Groupe hospitalier Pitié-Salpêtrière, institut de myologie, 105, boulevard de l'Hôpital, 75634 Paris Cedex 13, France. edgar.gomes@upmc.fr.

V. Gache et B. Cadot ont contribué équitablement à ce travail.

Maturation des fibres du muscle squelettique et positionnement des noyaux

Les fibres des muscles squelettiques sont les principales cellules des vertébrés spécialisées dans la contraction, et sont responsables de tous les mouvements volontaires de l'organisme. Chaque fibre composant le muscle squelettique est un syncytium issu de la fusion de nombreuses cellules spécialisées : les myoblastes. Ces cellules, dont le potentiel prolifératif est élevé, ont la capacité de sortir du cycle cellulaire au début de la différenciation en exprimant de manière coordonnée les protéines spécifiques du muscle. Les myoblastes deviennent alors des myocytes, cellules spécialisées ayant acquis la particularité de pouvoir fusionner entre elles et de conduire à la formation de myotubes contenant plusieurs noyaux. Ces myotubes sont caractérisés par un positionnement central des noyaux qu'ils contiennent. Au cours de la maturation des myotubes, des unités contractiles appelées sarcomères vont se former par l'association précise de réseaux d'actines et de moteurs moléculaires, les myosines, permettant la contraction des fibres musculaires. Ces myotubes matures deviennent des myofibres et localisent les noyaux à la périphérie, près de la membrane. Ces myofibres sont les briques élémentaires du muscle [1, 2].

In vivo, tandis que dans les fibres matures les noyaux se localisent à la périphérie, ils sont alignés au centre de la fibre régénérée *(Figure 1A, B)* [3]. Afin d'étudier ce phénomène, nous avons reproduit en laboratoire les premières étapes de la formation des fibres musculaires, et nous avons suivi les déplacements des noyaux au cours du temps. Nous avons pu observer qu'à la suite de la fusion d'un myocyte avec un myotube, le noyau du myocyte migre rapidement vers le centre du myotube. Par la suite, ces noyaux s'alignent le long du syncytium *(Figure 1C)*. Ces observations nous ont

Photo : myotube formé *in vitro* après 5 jours de différenciation (microtubules en vert, noyaux en rouge) (© V. Gache, UMR S 787 Inserm).