



**HAL**  
open science

# Identification, polymorphisme et évolution moléculaire de gènes du pouvoir pathogène chez le nématode à kyste de la pomme de terre *Globodera pallida*

Alexandra Blanchard

## ► To cite this version:

Alexandra Blanchard. Identification, polymorphisme et évolution moléculaire de gènes du pouvoir pathogène chez le nématode à kyste de la pomme de terre *Globodera pallida*. Biochimie [q-bio.BM]. Université Rennes 1, 2006. Français. NNT: . tel-00132028

**HAL Id: tel-00132028**

**<https://theses.hal.science/tel-00132028>**

Submitted on 20 Feb 2007

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

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

N°ORDRE : 3433

# THESE

présentée

DEVANT L'UNIVERSITE DE RENNES 1

pour obtenir

le grade de : *DOCTEUR DE L'UNIVERSITE DE RENNES 1*

Mention : Biologie

Par

**Alexandra BLANCHARD**

Unité mixte de recherche INRA-Agrocampus Rennes

Biologie des organismes et des populations appliquée à la protection des plantes

Equipe : Biologie et génétique des nématodes

Ecole doctorale : Université de Rennes 1 / Agrocampus Rennes Vie-Agro-Santé

**IDENTIFICATION, POLYMORPHISME ET EVOLUTION MOLECULAIRE DE  
GENES DU POUVOIR PATHOGENE CHEZ LE NEMATODE A KYSTE DE LA  
POMME DE TERRE *GLOBODERA PALLIDA***

Soutenue le 19 décembre 2006 devant la commission d'examen

Composition du jury :

Vivian BLOK	Chargée de recherche	Rapporteur
Pierre ABAD	Directeur de recherche	Rapporteur
Serge MORAND	Directeur de recherche	Examineur
Daniel BOUJARD	Professeur	Examineur
Véronique LEFEBVRE	Directeur de recherche	Examineur
Eric GRENIER	Chargé de recherche	Directeur de thèse

*"Les individus sont des artifices  
inventés par les gènes pour se  
reproduire, ce qui n'est pas facile à  
accepter pour nous qui sommes des  
individus"*

*P-H Gouyon*

## Remerciements

A l'issue de ce travail de thèse, je tiens à témoigner toute ma reconnaissance et ma sympathie à Eric GRENIER pour m'avoir confié ce projet, m'avoir fait confiance et m'avoir orientée au cours des 4 dernières années. Merci Eric pour ton implication, ta disponibilité, ton soutien, ton aide, ta sympathie. Ces 4 années ont débutées pour moi sur les chapeaux de roue, je me suis complètement plue dans ce projet que j'ai pu voir et faire évoluer...L'expression "formation par la recherche" a petit à petit pris tout son sens...Cette liberté que tu m'as laissée prendre au fil du temps m'a permis de m'épanouir scientifiquement et personnellement dans mon projet de thèse qui arrive à son terme avec, pour moi, un grand succès. Cette période a été pour moi très riche sur de très nombreux plans !! Je suis maintenant prête à voguer vers d'autres horizons et à prendre d'autres responsabilités. Alors pour tout ça, tu as toute ma reconnaissance et mon admiration.

Je voudrais remercier chaleureusement Olivier PLANTARD : pour tes conseils, ton aide notamment lors de l'analyse des données, ton soutien et ton implication dans mon comité de thèse, ta disponibilité à toute épreuve, ta gentillesse et ton partage du savoir. Nos échanges ont toujours été très riches d'enseignements. Je t'en suis très reconnaissante et suis extrêmement admirative devant tes énormes capacités d'adaptation à divers sujets. Un grand merci également pour m'avoir fait découvrir d'autres horizons plus évolutionnistes chaque jour, et dans un autre registre, un autre Antony !

Je souhaite remercier tout particulièrement Magali ESQUIBET, dite "mini chef", qui m'a guidée dans mes premiers pas au laboratoire, m'a soutenue sur le plan technique dès mon arrivée. Merci également pour le travail d'équipe que nous avons réalisé avec Didier et Eric. Un grand merci pour ta sympathie, ta bonne humeur de tous les jours, ta gentillesse, ton amitié et pour m'avoir soutenue sur le plan personnel, même à l'extrême (saut à l'élastique, saut en parachute...mariage !!! Quelle aventure que cette thèse !). Merci de m'avoir suivie dans les séances de gym effrénées même si les excuses pour y échapper ne manquaient pas !

Merci à Didier FOUVILLE pour son implication technique dans ce travail de thèse, sa disponibilité, sa gentillesse de tous les instants. Cette thèse est le fruit d'un travail



d'équipe auquel tu as largement contribué et je t'en remercie vivement. J'espère que nous arriverons maintenant à valoriser tout ce travail au mieux.

Merci à tous les autres membres de l'équipe de Nématologie : Sylvie – bravo pour ce joli bout de chou...Gabriel –, Catherine – merci pour ta disponibilité à chaque fois que je t'ai sollicitée ! –, Laurent – et viva Espagne ! –, Lionel – et Renaud ... –, M. Mugniéry – merci pour votre accueil, votre aide et pour mon initiation à la nématologie hors BM ! –, Claudia – merci pour ta gentillesse, ton soutien et les cotis' café ! Le binôme a plutôt bien fonctionné ! –, Elsa – merci pour ta contribution à mon initiation à l'identification visuelle des nématodes ! Et merci pour les étudiants de l'ENSAR –, Phan – bon courage pour la suite –, Marie-Christine – alias McD, merci pour ta gentillesse, ton amitié et pour ces tablettes de chocolats partagées ! chuuut ! –, Sylvain – merci de m'avoir acceptée au "trou" pendant ces 4 années au cours desquelles les échanges (amicaux, cela va sans dire !) ont fusés et également pour ton grand sens du compliment ! –, M. Rivoal – merci pour vos encouragements notamment lors de notre escapade bulgare –, Yves – merci pour m'avoir fait découvrir ces fabuleuses interactions nématodes/taupins ! Mais aussi pour m'avoir souvent faite tourner la tête...sur des danses bretonnes diverses et variées !!! –, et enfin Katell – Bonne chance pour la suite ! Merci à tous pour m'avoir accueillie à bras ouverts et m'avoir volontiers consacré un peu de votre temps lorsque que j'ai eu besoin, de votre aide, de vos conseils ou de votre expertise. Je remercie tout particulièrement Géraldine pour m'avoir soutenue dès mon arrivée dans l'équipe, pour tes conseils, ta disponibilité, ta sympathie et ton amitié...mais aussi pour nos longues discussions, nos séances de natation effrénées. A moi maintenant de te donner le goût pour d'autres rythmes plus...terrestres ! Merci Damien de t'être dévoué corps et âme dans cette thèse ! Mais aussi pour ta sympathie, ton dynamisme, ta bonne humeur à toutes épreuves. Merci Pauline pour ta contribution graphique à ce manuscrit mais aussi à ta description graphique des nématodes Letorta et Blancharda, c'est ça !?! Tu devrais publier ! Ce fût un très beau cadeau. Merci Abou pour ta gentillesse (même quand tu fais ton Abou !) et ton amitié. Merci à Laure et Virgil (un peu plus nématologiste chaque jour depuis le congrès de Blagoevgrad !) pour votre soutien, votre sympathie, votre amitié. Encore un peu de courage à tous les deux...je suis certaine que les séances de gym ne seront qu'un plus !!

Merci à tous pour m'avoir encouragée, soutenue dans tout ce que j'ai pu entreprendre au laboratoire.

Merci à Patrick LHOMME pour avoir contribué à ce travail par le biais de son stage de master II.

Merci à l'équipe BM, Morgane, Nathalie, Lucie, Agnès, Lionel et Manu pour leur aide, leurs précieux conseils, leur disponibilité au quotidien. Votre aide à tous a largement contribué au bon déroulement de ce projet de thèse. Merci à vous tous pour votre sympathie au labo mais aussi en dehors, pour toutes les activités plus ou moins extrêmes que nous avons partagées et pour votre amitié ! J'espère que nous aurons d'autres occasions de mettre ensemble notre courage à rude épreuve ! Je fais à cette occasion un clin d'oeil aux conjoints : Matthieu, Isa, Guillaume, Anne-Sophie, Patrick, Frank et aux enfants (Tom, Yoann, Hugo, Maé, Coline & Romane) avec qui j'ai également partagé de très bons moments plus ou moins extrêmes !! Et que tous ces beaux et adorables enfants aient une vie pleine de joies et d'épanouissements, et apportent encore beaucoup de bonheur à leurs parents !!

Merci à Claude RISPE et à Jean-Pierre GAUTHIER pour leur aide bioinformatique lors de l'analyse de mes données.

Merci aux membres de mon comité de thèse, Abdelhak EL-AMRANI, Pierre DELEPORTE, Marie-Noëlle ROSSO, pour les conseils qu'ils ont pu me donner notamment en début de thèse pour orienter au mieux le projet et au soutien qu'ils m'ont apporté.

Merci à John JONES, Vivan BLOK et Mark PHILLIPS du SCRI de Dundee pour leur accueil lors de mes deux séjours très intéressants et bénéfiques en Ecosse.

Merci à Denise, Dominique, Bruno, Michèle, Géraldine, Anne-Sophie... toujours opérationnels d'un point de vue logistique !

Merci à Magali ERMEL pour m'avoir fait découvrir la Bretagne sous l'angle de la danse. Et merci à tout le groupe de danse bretonne. Kenavo !

Merci le CFU pour les bons moments de festivités bisannuels qui apportent tant de bonne humeur et de bonheur pour les papilles !!

Merci aux autres thésards, anciens et actuels à qui je souhaite bon courage pour la suite et bonne route !

Merci à tous les amateurs de café et de thé qui nous permettent de partager quelques douceurs de temps à autre...

Merci également aux adeptes de gym qui ont été de plus en plus nombreux au fil des années...je compte sur vous pour vous motiver (Magali en particulier !) quand je ne serai plus là pour regrouper les troupes le mardi soir !!!

Merci à toutes les personnes que j'ai côtoyées de près ou de loin et qui m'ont permis de réaliser au mieux se travail de thèse et de rendre ces 4 années très agréables et prolifiques je l'espère ! Merci à tous les membres de l'UMR BiO3P qui contribuent au quotidien au bon déroulement de tous nos travaux. Merci pour votre sympathie qui rend notre dur labeur plus agréable tous les jours ! Je vous souhaite à toutes et tous une bonne continuation dans vos activités diverses et variées, qu'elles soient riches d'enseignements et fructueuses !

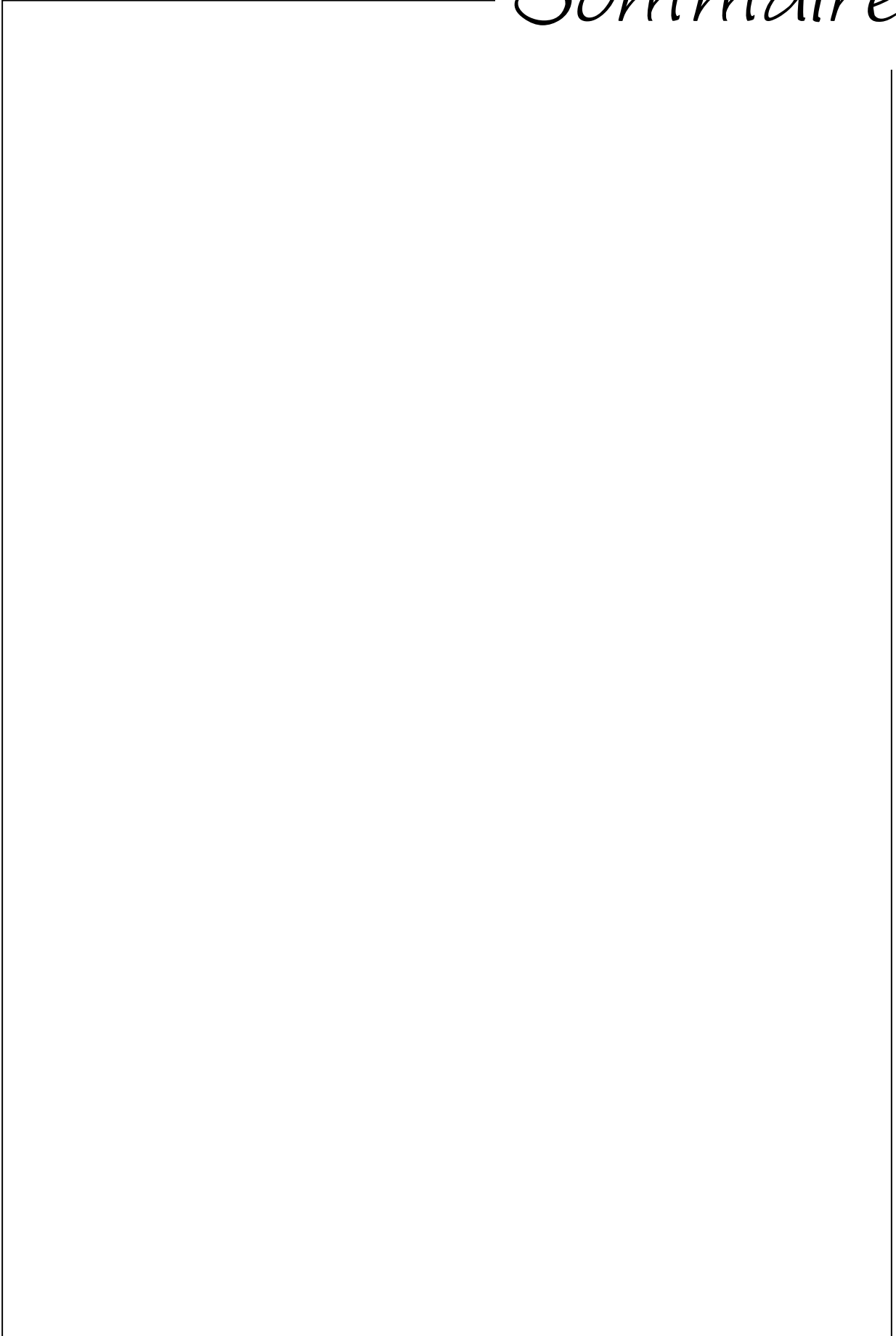
Pour terminer, un grand merci chaleureux à tous ceux qui ont partagé le plus grand jour de notre vie...ce sont des souvenirs intarissables.

Un grand merci également à mes parents et mes frères qui m'ont toujours encouragée à poursuivre mes études et m'ont toujours soutenue dans mes décisions.

Merci à tous mes amis extérieurs au labo en particulier le groupe handballistique, et aussi à Nico, Marion et Nohan pour m'avoir donné le doux surnom de "Mme Patate"...j'aurais au moins réussi à faire passer le message de mon travail relatif aux "vers parasites de la pomme de terre", même si je n'ai pas trouvé la "patate magique" (pour Nico) !!!

Enfin, je ne pourrais clore ce chapitre sans remercier la personne qui a vécu cette thèse aussi intensément que moi, mais dans l'ombre...Un grand merci à Anthony à qui je dédie cette thèse du fond du cœur. Merci de m'avoir soutenue sur tous les plans depuis si longtemps...

# *Sommaire*



# SOMMAIRE

## PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES

## GLOSSAIRE

## ABREVIATIONS

INTRODUCTION GENERALE : CONTEXTE ET OBJECTIFS DE LA THESE.....	1
<b>PREMIERE PARTIE : INTRODUCTION BIBLIOGRAPHIQUE.....</b>	<b>5</b>
1. Le phylum Nemata.....	5
1.1. Diversité des nématodes.....	5
1.2. Eléments de systématique.....	6
1.3. Organisation structurale des nématodes.....	9
2. Les nématodes phytoparasites : Généralités.....	9
2.1. Les stratégies parasitaires des nématodes phytophages.....	9
2.2. Dégâts liés aux nématodes phytophages.....	11
3. Les nématodes a kyste : <i>Globodera</i> & <i>Heterodera</i> .....	12
3.1. Fonctionnement des populations de nématodes à kyste.....	12
3.2. Relations phylogénétiques et classification.....	13
A. <i>Les nématodes du genre Heterodera</i> .....	13
B. <i>Les nématodes du genre Globodera</i> .....	15
4. Méthodes de lutte contre les nematodes à kyste .....	15
4.1. Lutte génétique : utilisation de variétés résistantes.....	17
4.2. Lutte génétique : résistances artificielles.....	18
A. <i>Cibler les protéines sécrétées par le nématode</i> .....	19
B. <i>Cibler directement les gènes de nématodes</i> .....	20
C. <i>Limite des résistances artificielles</i> .....	20
5. Biologie des nématodes a kyste : <i>Globodera</i> & <i>Heterodera</i> .....	21
5.1. Cycle de développement .....	21
5.2. Organisation de la partie antérieure des nématodes .....	23
6. Implication des sécrétions salivaires dans les étapes clé du parasitisme.....	24
6.1. Les phases précoces du parasitisme : pénétration et migration dans la racine hôte.....	25
A. <i>Conséquences physiologiques pour la plante hôte</i> .....	25
B. <i>Sécrétions salivaires impliquées dans les stades précoces du parasitisme</i> .....	26
6.2. Les phases tardives du parasitisme : élaboration et maintien du site nourricier....	27
A. <i>Modifications physiologiques de la plante hôte</i> .....	27
B. <i>Sécrétions salivaires impliquées dans les phases tardives</i> .....	29

## RESULTATS

### DEUXIEME PARTIE : CARACTERISATION DE NOUVEAUX GENES DU POUVOIR PATHOGENE CHEZ *GLOBODERA PALLIDA* ET *G. "MEXICANA"*

PRESENTATION DE LA PARTIE II.....	31
INTRODUCTION.....	33
CHAPITRE 1	
1. Eléments de contexte : Le genre <i>Globodera</i> .....	35
2. Article 1 : Sequence polymorphism of two pioneer genes in phytoparasitic nematodes showing different host ranges.....	39
CHAPITRE 2	
1. Eléments de contexte : Les RanBPM.....	57
2. Article 2 : RanBPM homologues genes characterised in the cyst nematodes <i>Globodera pallida</i> and <i>Globodera "mexicana"</i> .....	61
CHAPITRE 3	
1. Eléments de contexte : La famille multigénique des RanBPM.....	69
2. Article 3 : Characterisation of the RanBPM-like gene family of the plant parasitic nematode <i>Globodera pallida</i> .....	75
CONCLUSION.....	85

### TROISIEME PARTIE : VARIABILITE DE GENES DU POUVOIR PATHOGENE DES NEMATODES A KYSTE ET EVOLUTION MOLECULAIRE DE CES GENES

PRESENTATION DE LA PARTIE III.....	89
INTRODUCTION.....	91
CHAPITRE 1	
1. Eléments de contexte : Choix des gènes et des espèces.....	93
1.1. Choix des populations et espèces au sein de la famille des Heteroderidae.....	93
1.2. Choix des gènes du pouvoir pathogène.....	95
A. Les pectate lyase.....	96
B. Les cathepsine L.....	100
C. Le facteur d'élongation 1 $\alpha$ .....	103
2. Article 4 : Analysis of parasitism genes in cyst nematodes to recount the evolutionary history of the Heteroderidae.....	107
3. Résultats complémentaires : caractérisation des gènes du parasitisme dans les 40 populations de <i>Globodera</i> et d' <i>Heterodera</i> .....	129
3.1. Amplification et caractérisation du gène cathepsine L.....	129
3.2. Amplification et caractérisation du gène pectate lyase.....	132

## CHAPITRE 2

1. Eléments de contexte : contraintes sélectives de la relation hôte/pathogène.....	135
1.1. Evolution des gènes du pouvoir pathogène.....	136
1.2. Pressions de sélection exercées par la plante hôte.....	137
2. Article 5 : Analyse des pressions de sélections.....	138
Introduction.....	138
Résultats.....	139
1. Répartition des mutations pour chaque gène, le long de la séquence amplifiée en fonction de la position dans le codon.....	139
1.1. Le gène cathepsine L.....	139
1.2. Le gène pectate lyase.....	140
1.3. Le gène facteur d'élongation 1 $\alpha$ .....	142
2. Identification de la sélection agissant sur les gènes du pouvoir pathogène et sur le gène de ménage.....	143
2.1. Analyse des ratios des taux de substitution non synonymes vs non synonyme par fenêtre glissante ou "sliding window".....	144
2.2. Analyses des taux de substitutions différentiels entre les sites et les espèces pour les gènes du pouvoir pathogène : recherche de sélection positive.....	146
Discussion.....	152
1. Quel rôle des sites sous sélection positive.....	152
2. Comment expliquer une accélération de l'évolution des gènes du parasitisme dans certaines populations et/ou espèces ?.....	152
3. En quoi les informations sur les zones à fort potentiel évolutif sont-elles utiles ?....	154

## QUATRIEME PARTIE : CONCLUSION ET PERSPECTIVES..... 155

1. Identification et caractérisation des gènes du pouvoir pathogène des nématodes à kyste et mise en évidence de leur rôle dans l'interaction plante/nématode.....	157
1.1. Etude de lignées hybrides <i>G. pallida</i> x <i>G. "mexicana"</i> : polymorphisme des gènes <i>la7</i> et <i>IVg9</i> .....	157
1.2. IC5 devenu <i>rbp-1</i> : la grande famille des gènes RanBPM-like.....	157
1.3. Un transfert horizontal des gènes du parasitisme.....	159
2. Polymorphisme nucléotidique et évolution moléculaire des gènes du pouvoir pathogène.....	160
2.1. Comment s'assurer de la comparaison de gènes orthologues ?.....	160
2.2. Les gènes du pouvoir pathogène étudiés ont-ils évolué différemment des gènes de ménage ?.....	161
2.3. Tous les gènes du pouvoir pathogène évoluent-ils sur le même mode ?.....	161

<b>3. Perspectives.....</b>	<b>162</b>
3.1. Apport des résultats pour l'étude des gènes du pouvoir pathogène.....	162
3.2. Apport des résultats pour la gestion des gènes de résistance.....	163
3.3. Apport des résultats pour l'étude du fonctionnement des populations.....	164

<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>166</b>
---	------------

## **ANNEXES**

### **ANNEXE 1 : AMORCES**

Liste des amorces utilisées.....	177
Localisation des amorces <i>rbp-1</i> sur le transcrit pleine longueur chez <i>G. pallida</i> .....	179
Localisation des amorces pectate lyase sur le transcrit pleine longueur chez <i>G. rostochiensis</i> .....	180
Localisation des amorces cathepsine L sur le transcrit pleine longueur chez <i>G. pallida</i> .....	181

### **ANNEXE 2 : MATERIEL ET METHODES**

Création et criblage de la banque SSH.....	182
5'RACE.....	184
Extraction d'ADN au CTAB.....	186
Extraction d'ARNs totaux à partir d'un juvénile de nématode à kyste.....	188
Extraction d'ADN génomique à partir d'un juvénile de nématode à kyste.....	189
Southern Blot.....	190
Hybridation <i>in situ</i> .....	195
Tampons pour l'hybridation <i>in situ</i> .....	198
Amplification des gènes pectate lyase, cathepsine L, facteur d'élongation 1 $\alpha$ et <i>rbp-1</i> .....	199
Analyse des pressions de sélection.....	203

### **ANNEXE 3 : ALIGNEMENTS DE SEQUENCES**

Alignement des séquences génomiques pectate lyase (introns et exons).....	206
Alignement des séquences génomiques cathepsine L (introns et exons).....	214
Alignement des séquences génomiques du facteur d'élongation (exons).....	222

<b>ANNEXE 4 : COMMUNICATIONS AFFICHÉES.....</b>	<b>228</b>
---	------------



## PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES

### PUBLICATIONS SCIENTIFIQUES :

- ♦ **Blanchard A.**, Esquibet M., Fouville D. and E. Grenier (2005). *Ranbpm* homologue genes charaterised in the cyst nematode *Globodera pallida* and *Globodera "mexicana"*. ***Physiological and Molecular Plant Pathology* 67:15-22.**
- ♦ **Blanchard A.**, Fouville D., Esquibet M., Mugniéry D. and E. Grenier. Sequence polymorphism of two pioneer genes in phytoparasitic nematodes showing different host ranges. **Soumise à *Journal of Heredity*.**

### COMMUNICATIONS SCIENTIFIQUES :

- ♦ **Blanchard A.**, Esquibet M., Fouville D. and E. Grenier. Four new *Globodera* genes putatively involved in plant-nematode interaction. **(Poster) XXVII ESN Symposium, June 2004, Rome, Italy.**
- ♦ **E. Grenier, Blanchard A.**, Esquibet M. and D. Fouville. Identification and polymorphism of new potato cyst nematode genes putatively involved in plant-nematode interaction. **(Poster) Conférences Jacques Monot "Ecology and evolution of host-parasite relationships", September 2004, Roscoff.**
- ♦ **Blanchard A.**, Pylypenko L., Thirugnanasambandam A., Block V., Chapman S., Grenier E., Lilley C., Neatham J., Phillips M. and **J.T. Jones**. Characterisation on a ranbpm multigene gene family from ESTs of *Globodera pallida*. **(Communication orale) 58<sup>th</sup> International symposium on crop protection, May 23, 2006, Gent, Belgium.**
- ♦ **Blanchard A.**, Fouville D., Esquibet M. and E. Grenier. Evolution of parasitism genes among the cyst nematodes. **(Communication orale) XXVIII ESN Symposium, June 2006, Blagoevgrad, Bulgaria.**
- ♦ **Blanchard A.**, Castagnone P., Rosso M.N., Fouville D. and E. Grenier. Evolution of parasitism genes in the Heteroderidae plant parasitic nematodes family. **(Poster & courte communication orale) XX Evolutionary Biology Meeting, September 2006, Marseille, France.**

# GLOSSAIRE

**Avirulence** : incapacité d'un pathogène à se reproduire sur un hôte possédant un gène de résistance.

**Clade** : Groupe d'organismes vivants qui ont une origine génétique commune et qui satisfont le critère de monophylie.

**Coévolution** : évolution conjointe, apparition d'adaptations réciproques chez deux ou plusieurs espèces qui ont des interactions durables (parasitisme, prédation...) et constituent des pressions de sélection réciproques.

**Convergence évolutive** : correspond à une analogie. Evènement qui fait que deux structures sont semblables d'un point de vue structural, mais qui ne possèdent pas d'éléments fondamentaux communs.

**Diversité interspécifique** : différences morphologiques, anatomiques, génétiques entre espèces.

**Diversité intra spécifique** : différences, variabilité entre les individus au sein d'une même espèce.

**Espèce** : les espèces sont des groupes de population réellement ou potentiellement capables de se croiser et qui sont isolées des autres groupes ayant les mêmes propriétés (Mayr, 1942).

**Exon** : région codante d'un gène.

**Gène d'avirulence** : gène du pathogène qui code une protéine qui est reconnue ou qui produit un composé reconnu par le gène de résistance correspondant pour induire une réaction de défense.

**Gène de pathogénicité** : gène du pathogène dont le produit a un rôle direct ou indirect dans le succès parasitaire.

**Gène de résistance** : gène de plante codant un déterminant spécifique permettant la reconnaissance et l'activation d'une réponse de défense contre un pathogène.

**Gènes orthologues** : gènes homologues qui ont divergés entre eux après un évènement de spéciation.

**Gènes paralogues** : gènes homologues qui ont divergés entre eux après duplication génique.

**Groupe monophylétique** : ensemble composé d'un ancêtre commun et de toutes les espèces qui en descendent.

**Homologie** : deux caractères sont homologues s'ils présentent des éléments structuraux communs, n'ayant pas nécessairement la même fonction.

**Intron** : région non codante d'un gène.

**Pathotype** : pour certaines espèces de nématodes à kyste, une population est définie par son succès qualitatif ou parfois quantitatif ou alors par son échec à produire une descendance sur une gamme de génotypes hôtes spécifiques, chacun présentant des résistances différentes.

**Phylogénie** : méthode systématique de construction d'arbres évolutifs montrant les relations entre un ancêtre commun et les espèces qui en dérivent.

**Phylum** : lignée génétique complexe d'être vivants.

**Polymorphisme** : variation observée entre deux séquences, que ce soit en terme de taille de la séquence ou en terme de composition nucléotidique.

**Population** : En ce qui concerne les nématodes à kyste, une population est relative à un bassin de production (Picard *et al.*, 2004).

**Potentiel évolutif** : estimation de la capacité à évoluer à partir des évolutions passées observées.

**Réaction d'hypersensibilité** : mort cellulaire programmée qui se produit lorsque le pathogène portant un gène d'avirulence est reconnu par la plante qui porte le gène de résistance correspondant.

**Résistance artificielle** : résistance générée par introduction d'un transgène dans une plante.

**RNAi** : Inactivation de gènes par interférence d'ARN.

**Sélection positive ou sélection/pression diversificatrice** : une mutation qui favorise un individu aura plus de chance de se fixer dans la population, c'est la sélection positive ou sélection Darwinienne, ou encore la sélection adaptative.

**Sélection/pression purificatrice** : si une mutation a des effets délétères pour un individu, la probabilité qu'elle se fixe dans une population est réduite, c'est la sélection négative ou purificatrice.

**Spéciation** : Formation d'une ou de plusieurs nouvelles espèces à partir d'une autre (ancêtre commun).

**Substitution non synonyme** : modification nucléotidique qui engendre une modification de l'acide aminé correspondant.

**Substitution synonyme** : modification nucléotidique qui n'engendre pas de modification de l'acide aminé correspondant.

**Synapomorphie** : caractère dérivé, partagé.

**Tempo d'évolution** (vitesse d'évolution) : nombre de mutations fixées par unité de temps et de longueur de molécules (par gène, par site ponctuel).

**Théorie neutraliste de l'évolution** : théorie définie par Motoo Kimura dans les années 60 qui considère que la plupart des régions géniques sont soumises à sélection purificatrice.

**Variabilité horizontale (d'un gène)** : polymorphisme des copies paralogues d'un gène.

**Variabilité verticale (d'un gène)** : polymorphisme des copies orthologues d'un gène.

**Virulence** : capacité d'un pathogène à se développer sur un hôte exprimant un gène de résistance.

# ABREVIATIONS

- aa** : Acide aminé
- ADN** : Acide désoxyribonucléique
- ADNc** : Acide désoxyribonucléique complémentaire
- AFLP** (Amplified Fragment-Length Polymorphism) : Polymorphisme de longueur de fragments amplifiés
- ARN** : Acide ribonucléique
- ARNm** : ARN messagé
- BCIP** : 5-bromo-4-chloro-3-indolyl phosphate
- CathL** : Cathepsine L
- CTAB** : Cetyltriméthylammonium bromide
- Dig** : Digoxygénine
- dNTP** : Désoxyribonucléotide triphosphate
- DTT** : Dithiothreitol
- EDTA** : Ethylène diamine tétra acétate de sodium
- EF** (Elongation factor) : Facteur d'élongation
- EST** (Expressed Sequence Tag) : Etiquettes des séquences exprimées
- GAP** : GTPase Activating Protein
- GP** : *Globodera pallida*
- GM** : *Globodera "mexicana"*
- GR** : *Globodera rostochiensis*
- GT** : *Globodera tabacum*
- GTP** : Guanosine 5'-triphosphate
- ITS** : Internal Transcribed Spacer
- J2** : Juvénile de second stade de développement
- kDa** : Kilo Dalton
- min** : Minute (s)
- NBT** : Nitro Blue Tetrazolium chloride
- ORF** (Open Reading Frame) : Cadre ouvert de lecture
- pb** : Paire de bases
- PCR** (Polymerisation Chain Reaction) : Réaction de Polymérisation en Chaîne
- PL** : Pectate lyase
- RACE** (Rapid Amplification of cDNA Ends) : Amplification rapide des extrémités d'ADNc
- RNAi** : ARN interférence
- RT** : Reverse transcription
- SDS** : Dodécyl sulfate de sodium
- SSH** : Hybridation suppressive et soustractive
- SSC** : Citrate de sodium-chlorure de sodium
- TDF** (Transcribed Derived Fragments) : Fragment de transcrit

**TdT** : Terminal deoxynucleotidyl transferase

**Tris** : Trihydroxyméthyl aminométhane

**U** : Unité

**UP** : Ultrapure

**UV** : Ultra-violets

*Introduction*

*Générale*

# CONTEXTE ET QUESTIONS DE RECHERCHE

## CONTEXTE

Les interactions hôte/parasite sont extrêmement diverses et semblent en général perdurer dans le temps de telle sorte qu'elles ont été qualifiées d'interactions durables (Combes, 1992). Comment expliquer qu'une interaction dans laquelle un seul des protagonistes est bénéficiaire puisse s'installer dans la durée ? Les organismes libres devenus parasites ont probablement trouvé un habitat plus favorable auprès de leur hôte. Quant aux organismes devenus hôtes, ils ont soit évolué de manière à tolérer leur(s) parasite(s) du fait du coût trop important pour les éliminer, ou bien ils y ont également trouvé des avantages. Quels sont alors les avantages de la vie parasitaire ? Dans le cas des nématodes phytoparasites, la plupart d'entre eux sont parasites du système racinaire des plantes provoquant des difficultés d'absorption de l'eau et des nutriments du sol. Il semble peu probable que la plante tire un quelconque profit dans ce cas. Côté nématode, l'environnement hostile du sol (bactéries parasites, prédateurs, propriétés physico-chimique défavorables...) a certainement conduit les nématodes à développer des stratégies pour passer le moins de temps possible dans le sol et se loger dans la plante, c'est l'endoparasitisme.

Dans notre étude, nous nous sommes intéressés aux relations hôte/parasite des nématodes à kystes, endoparasites, qui mettent en place une relation très étroite avec leur hôte, nécessaire pour effectuer leur cycle de développement et atteindre les stades adultes auxquels ils se reproduisent. Les femelles nématodes sont sédentaires et dépendantes de la racine qu'elles parasitent pour la nourriture jusqu'à ce qu'elles aient généré leur descendance. Les mâles sont eux libres dans le sol et se déplacent à la recherche de femelle(s) à féconder. Les nématodes à kystes, sont oligophages et à ce titre sont très dépendants des caractéristiques de leur plante hôte qui peut ainsi influencer l'évolution du parasite qui doit s'adapter en permanence. Ces évolutions conjointes passent par l'évolution des gènes qui sont mis en jeu dans ces interactions et non pas par l'évolution des individus en tant que tels. Combes (2001) définit différents types de gènes régissant les associations d'organismes : des gènes pour rencontrer, pour éviter, pour tuer ou encore pour survivre (Fig. 1). Chaque partenaire dans une interaction possède donc des gènes qui évoluent et lui permettent de rester le mieux adapté à son environnement et donc à son hôte.

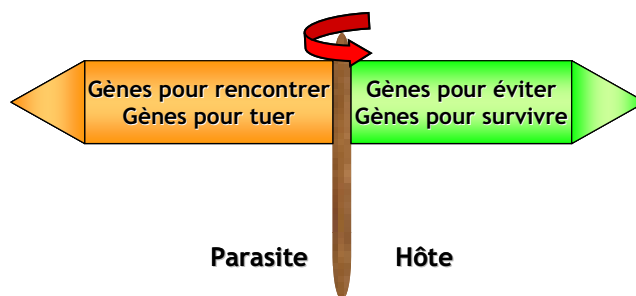


Figure 1 : Différents gènes régissant les interactions hôte/parasites.

Ainsi, comme dans toutes les interactions plante pathogène, les nématodes possèdent une batterie de protéines qui leur permettent de parasiter la plante qui elle, possède un large spectre de gènes de résistance pour lutter contre les attaques des nématodes mais aussi d'autres pathogènes (bactéries, champignons, virus, plantes supérieures). Ces gènes coévoluent au gré des interactions.

Le phénomène de coévolution est le processus par lequel deux adversaires acquièrent sans cesse de nouvelles adaptations pour ne pas être distancés par "l'autre". Ces phénomènes d'adaptation sont possibles grâce à l'évolution des gènes impliqués dans la relation parasitaire et sont contraints par les pressions qu'exerce chaque partenaire sur ces gènes. La compréhension des mécanismes moléculaires sous jacents est rendue difficile par la complexité écologique du milieu dans lequel le couple se trouve (Gouyon, 2001). Les disciplines visant à étudier les mécanismes d'interaction ainsi que l'évolution de ces mécanismes, doivent donc nécessairement être couplées à des études au sein des populations afin d'avoir une vision plus globale de l'évolution de la relation hôte/parasite et de comprendre ce type de relation non pas à l'échelle de deux individus, mais plutôt à l'échelle de deux populations se développant dans un milieu donné. En effet, la sélection n'agit pas sur les individus, mais sur l'information génétique qu'ils portent et qu'ils transmettent au fil des cycles de reproduction (Gouyon, 2001).

Dans le modèle nématode/plante que nous avons choisi, peu de données sont connues et disponibles sur les déterminismes du pouvoir pathogène de ces organismes et sur les résistances mise en place par la plante hôte. De ce fait, l'évolution des relations hôte/parasite de ce modèle est encore peu connue. Pourquoi faut-il s'y intéresser ?

Les nématodes à kyste du genre *Globodera* sont des parasites telluriques responsables de dégâts considérables à l'échelle de la planète sur une gamme d'hôte quasiment restreinte aux Solanacées. La lutte chimique qui est actuellement la plus efficace contre ces organismes est progressivement abandonnée en raison des problèmes environnementaux que posent les produits nématicides (ces produits ne sont pas sélectifs et les doses utilisées sont nettement supérieures à celles utilisées contre les insectes) ainsi qu'en raison des problèmes économiques ou réglementaires (interdiction progressive). Parmi les autres systèmes de lutte, la lutte variétale est privilégiée pour limiter le développement des populations de nématodes à kyste. En effet, les plantes possèdent un large spectre de mécanismes de défense qui peuvent être utilisés et introgressés dans les plantes d'intérêt agronomique parasitées par les nématodes. Les résistances identifiées à ce jour sont en général durables cependant, des évènements de contournement ont été mis en évidence par l'existence de populations de nématodes non contrôlées par ces résistances chez *Heterodera avenae*, *Heterodera schachtii* et *Globodera rostochiensis* par exemple. Ces constats ont favorisé les réflexions sur les possibilités de mettre en place de nouvelles méthodes de lutte, telles que les résistances artificielles. La principale difficulté réside dans le choix de la ou les cibles (gènes du pouvoir pathogène) contre lesquelles il pourrait être intéressant de générer des résistances artificielles avec comme principal objectif de créer des résistances durables dans le temps et efficaces contre un large spectre de nématodes. Les méthodes de lutte les plus prometteuses



reposent donc actuellement sur l'utilisation combinée des sources de résistance naturelle existantes et sur la mise au point de nouvelles variétés plus résistantes par le biais des biotechnologies.

L'identification des bases moléculaires de l'interaction plante/nématodes contribue à améliorer la compréhension des mécanismes mis en jeu au cours du parasitisme. Actuellement peu de données sont disponibles pour permettre de lever le voile sur tous ces mécanismes et mettre en évidence les clés de la relation plante/nématode. Si certains gènes sont mieux connus et ont une activité définie dans l'interaction, l'existence de plusieurs copies d'un même gène pouvant avoir des fonctions différentes est souvent évoquée mais pas encore élucidée pour les gènes de nématodes. Il est donc nécessaire de proposer de nouvelles approches d'étude des gènes du pouvoir pathogène. Les études fonctionnelles par transgénèse ne sont pas au point chez les nématodes phytoparasites et les études par "silencing" sont en cours de développement. Ces deux types d'approches se situent à l'échelle individuelle. Des approches populationnelles seraient complémentaires pour mieux évaluer la variabilité des gènes du pouvoir pathogène. Cette variabilité génétique repose sur l'accumulation de mutations nécessaires aux diverses adaptations des populations et espèces aux conditions environnementales. La caractérisation et l'analyse du polymorphisme des gènes du pouvoir pathogène devraient permettre de définir les pressions de sélection subies par ces gènes au cours de l'histoire évolutive des nématodes. Ces pressions de sélection conditionnent en partie les capacités adaptatives de certaines populations, espèces ou encore genres.

## **QUESTIONS DE RECHERCHE ET OBJECTIFS DE LA THESE**

Dans le contexte que nous venons de définir, mes travaux de thèse se sont orientés autour des aspects de caractérisation de nouveaux gènes du pouvoir pathogène et sur une approche innovante d'étude de la variabilité et des pressions de sélection subies par ces gènes à l'échelle d'une famille de nématodes, les Heteroderidae. A ce titre, les questions sous-jacentes étaient :

- 1- Quels sont les gènes du pouvoir pathogène des nématodes à kyste qui jouent un rôle clé dans le parasitisme ?
- 2- Comment ces gènes ont-ils évolué par rapport à divers critères tels que les différences de pathogénicité, de spécificité ou de gamme d'hôtes, d'origine géographique ?
- 3- Comment pourraient-ils évoluer s'ils étaient soumis à des pressions de sélection liées à une nouvelle source de résistance (exemple des résistances artificielles) ?

Afin d'apporter de nouvelles données permettant de répondre à ces questions, nous avons choisi de développer trois grands axes de recherche :

- 1- l'identification de nouveaux gènes du pouvoir pathogène chez le nématode à kyste *Globodera pallida*, afin d'apporter de nouvelles données pour mieux comprendre les bases moléculaires de l'interaction plante/nématode.
- 2- la caractérisation de la variabilité de gènes du pouvoir pathogène à l'échelle de la famille des nématodes à kyste en se focalisant sur le genre *Globodera*, afin de proposer une approche novatrice et complémentaire des études fonctionnelles des gènes du parasitisme.
- 3- l'identification des pressions de sélection subies par ces gènes au cours de l'évolution afin de mieux comprendre la diversité rencontrée et d'avoir une vision plus globale des modes évolutifs de ces gènes.

Ces axes devraient nous permettre d'apporter quelques réponses sur l'intérêt de l'utilisation de ce type d'approche pour mieux comprendre le fonctionnement des interactions. En effet, les connaissances en matière d'interaction plante/nématode doivent nous permettre de mieux appréhender les bases moléculaires de la diversité de pathogénicité chez les nématodes à kyste. Ces données permettront aussi d'évaluer l'efficacité et la durabilité de nouvelles méthodes de lutte dérivées du pathogène en terme de spectre de populations et d'espèces potentiellement visées et en terme de durabilité de ces résistances par évaluation du potentiel évolutif de ces gènes. Ces données pourraient également nous permettre de cibler les sources de résistances naturelles à rechercher chez les plantes hôtes en fonction des gènes du pouvoir pathogène qui seront identifiés comme les plus intéressants pour le développement de ces nouvelles méthodes de lutte.

Ce manuscrit est organisé en quatre parties rapportant les résultats de mes travaux.

La première partie constitue l'introduction bibliographique retraçant l'état des connaissances actuelles sur la biologie des nématodes à kystes, les relations phylogénétiques entre les différents genres et espèces ainsi que sur les stratégies parasitaires et les moyens mis en place par ces nématodes au cours de l'infestation de leur plante hôte.

Les deux parties suivantes rapportent les résultats obtenus pour les trois axes de recherche mentionnés précédemment. La deuxième partie aborde la caractérisation de gènes du pouvoir pathogène des nématodes à kyste et la troisième partie rapporte les résultats de l'étude de la variabilité et des pressions de sélections subies par ces gènes au cours de l'évolution.

La quatrième et dernière partie de ce manuscrit est constituée par la conclusion générale et les perspectives qui se dégagent de mes travaux.

*Introduction*  
*Bibliographique*

*Partie I*

# PARTIE I :

## INTRODUCTION BIBLIOGRAPHIQUE

Les nématodes sont très peu connus au regard d'autres organismes de la biosphère. Ce sont des animaux peu visibles dans l'environnement du fait de leur petite taille, et assez peu abordés, voire pas du tout, au cours des cursus scolaires et universitaires. Le nématode le mieux connu est *Caenorhabditis elegans*, organisme modèle dont le génome a été entièrement séquencé en 1998 et dont la structure simple permet l'étude de nombreux mécanismes cellulaires et moléculaires. Les nématodes suscitent pourtant beaucoup de curiosité de par leur caractère ubiquiste sur la planète et leur rôle important dans le fonctionnement de divers écosystèmes. Nous allons dans cette introduction présenter brièvement le phylum Nemata afin d'apporter quelques éclaircissements sur sa diversité, ses caractéristiques puis ensuite nous focaliser sur les nématodes parasites de plantes qui mobilisent de nombreuses recherches du fait de leur incidence en agronomie. Nous détaillerons l'état actuel des connaissances en matière de mécanismes moléculaires régissant les relations nématodes/plantes. Enfin, nous montrerons en quoi une meilleure connaissance de ces relations, notamment d'un point de vue évolutif, pourrait permettre de mieux lutter contre ces parasites.

### 1. LE PHYLUM NEMATA

#### 1.1. Diversité des nématodes

Les nématodes sont des organismes vermiformes cylindriques non segmentés occupant des niches écologiques très diverses sur la planète. S'ils comprennent différentes formes, libres ou parasites d'animaux ou de végétaux, les nématodes sont tous des animaux aquatiques. Ils exploitent différents milieux tels que les océans et les mers, les eaux douces mais aussi les fluides corporels, les films d'eau dans le sol ou sur les végétaux. Ce sont les organismes les plus abondants de tous les métazoaires en terme de nombre d'individus dans de nombreux écosystèmes, notamment ceux du sol ( $1.20.10^{11}$  individus/m<sup>2</sup> contre  $10^5$ /m<sup>2</sup> pour les acariens qui constituent le deuxième groupe le plus abondant, Kevan, 1965). Excepté leur morphologie très homogène, les nématodes présentent une très grande diversité avec un nombre total d'espèces dans le phylum Nemata estimé entre 40000 et 10 millions (Blaxter *et al.*, 1998; Dorris *et al.*, 1999 ; Blumenthal *et al.*, 2004). Ce grand nombre d'espèces (26000 décrites, Hugot *et al.*, 2001) les place au deuxième rang dans le règne animal après les insectes. Les nématodes ont également des régimes alimentaires très diversifiés. Certaines espèces sont bactériophages, comme le nématode "modèle" *C. elegans*, d'autres sont entomopathogènes (ex : *Steinernema* spp ou *Heterorhabditis* spp), parasites d'animaux (ex : *Ascaris* spp, *Brugia* spp ou *Trichinella* spp) ou encore prédatrices (ex : *Mononchus* spp). Enfin, parmi toutes

les espèces de nématodes décrites, seulement 15% sont des parasites de plantes (ex : *Globodera* spp, *Meloidogyne* spp, *Pratylenchus* spp, *Ditylenchus* spp...) (Fig. 1).

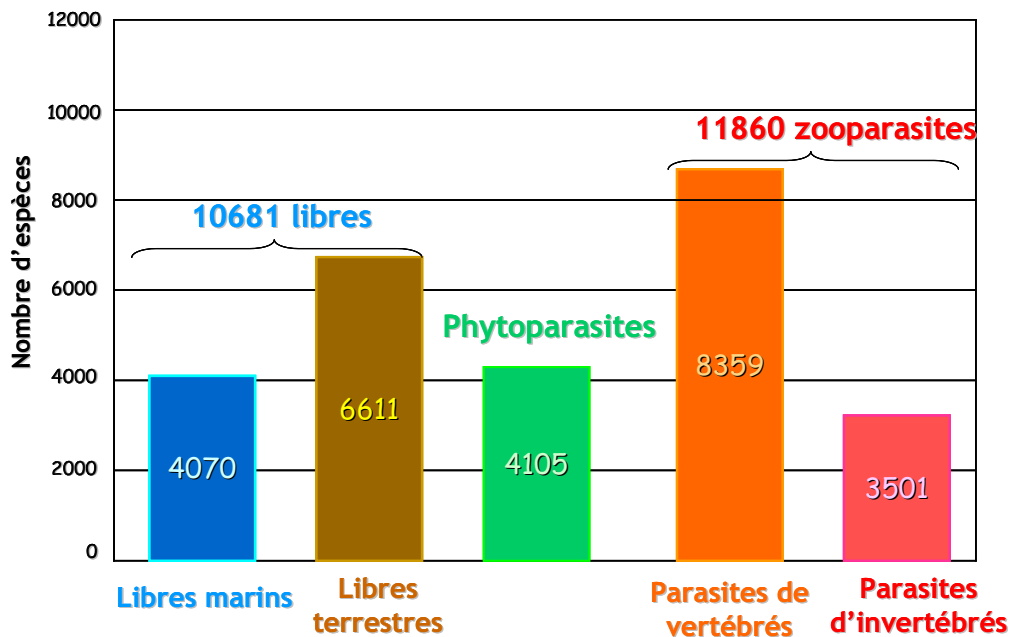


Figure 1 : Mode de vie des espèces de nématodes décrites, d'après Hugot *et al.*, 2001.

## 1.2. Éléments de systématique

Les nématodes sont des organismes triploblastiques (trois feuillet embryonnaires) et possèdent une cavité interne (formée par l'endoderme) non complètement recouverte de mésoderme. Au sein des métazoaires, les nématodes avaient été placés en fonction de ce critère morphologique dans les pseudo-coelomates (Fig. 2A). Avec l'apparition des techniques de biologie moléculaire, des phylogénies ont pu être réalisées. Basée sur la séquence 18S de l'ARNr, la phylogénie des métazoaires a alors été reconsidérée. Même si cette conception suscite encore de nombreuses interrogations, un consensus se dégage pour considérer que les nématodes peuvent être regroupés avec les arthropodes (Fig. 2B) pour former les Ecdysozoa (animaux capables de renouveler leur cuticule par des mues) (Aguinaldo *et al.*, 1997; Adoutte *et al.*, 1999). Ce groupe rassemble les nématodes, les insectes, les arachnides et les crustacés. Ce nouveau cadre phylogénétique est important, notamment lors des études moléculaires des organismes pour lesquelles les recherches de séquences homologues dans les bases de données (logiciels de type Blasts) sont la principale source d'identification de nouveaux gènes. Un cadre phylogénétique bien établi permet de mieux interpréter la proximité de certaines séquences en fonction de l'évolution des organismes (ex : occurrence de transferts horizontaux de gènes entre des bactéries et des eucaryotes tels que les nématodes phytoparasites). En effet, ces comparaisons ne sont possibles et pertinentes que si les organismes comparés possèdent un ancêtre commun relativement proche.

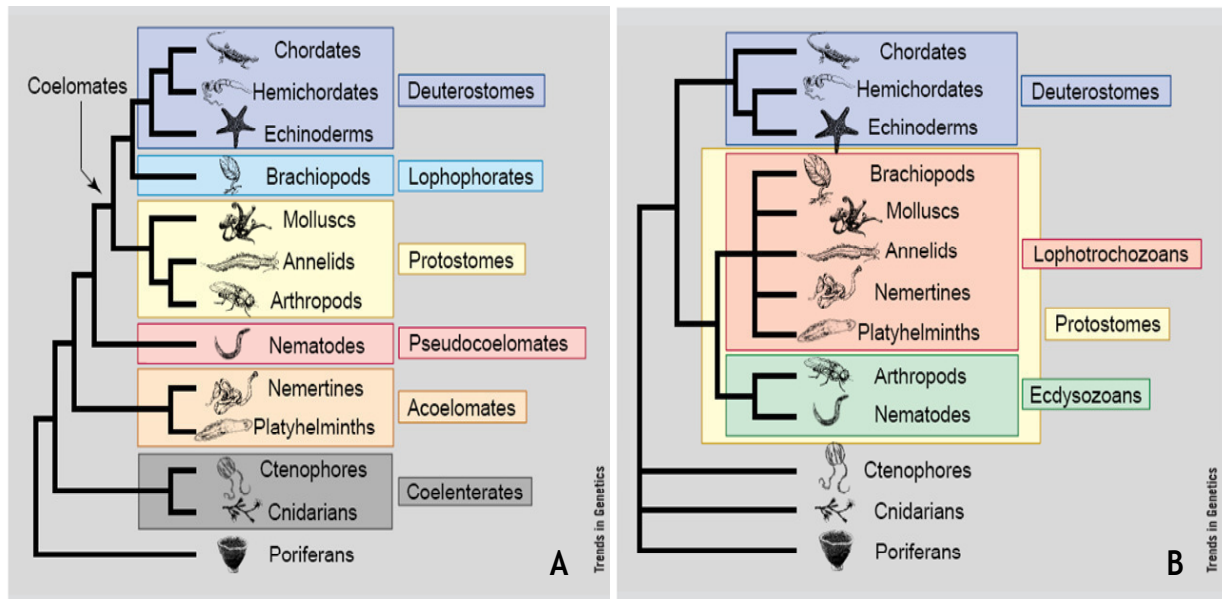


Figure 2 : Phylogénie des métazoaires (Adoutte *et al.* 1999). A : Phylogénie "traditionnelle" des métazoaires ; B : Phylogénie basée sur les séquences 18S de l'ARN ribosomal.

Les nématodes ont été répartis par Blaxter *et al.* (1998) en cinq clades majeurs dans la phylogénie du phylum Nemata basée sur les séquences de la petite sous unité de l'ARN ribosomal (Fig. 3). Chacun de ces cinq clades contient des espèces ayant des modes de vie divers et compte tous des espèces parasites d'animaux ou de végétaux. Le parasitisme des végétaux serait apparu indépendamment au moins trois fois au cours de l'évolution dans différents ordres : les Dorylaimida, les Triplonchida, les Aphelenchida et les Tylenchida (Blaxter *et al.*, 1998, Fig. 3). Les Dorylaimida et les Triplonchida sont surtout problématiques parce qu'ils sont vecteurs de virus, plus que pour les dégâts directs qu'ils provoquent sur les plantes. Les Tylenchida constituent l'ordre le plus important des nématodes phytoparasites à la fois en terme de nombre d'espèces, mais aussi en terme de dégâts causés aux plantes qu'ils parasitent. Cet ordre regroupe neuf familles de nématodes dont les Heteroderidae (Ferraz & Brown, 2002). Cette famille elle-même englobe une vingtaine de genres dont les nématodes à kyste des genres *Globodera* et *Heterodera*, sur lesquels nous nous sommes focalisés, et les nématodes à galle du genre *Meloidogyne* (même si de récents travaux remettent en cause la proximité de ces genres dans la même famille; Holterman *et al.*, 2006). Les Heteroderidae sont responsables des dégâts les plus importants sur les plantes. De ce fait, ils font actuellement l'objet du plus grand nombre d'études visant à élucider les mécanismes fins qui permettent le dialogue moléculaire étroit entre le nématode et sa plante hôte. En effet, ils établissent avec leur hôte une relation très complexe *via* l'établissement d'un site nourricier, dont les mécanismes d'induction sont encore mal connus. Nous reviendrons sur ce point ultérieurement.

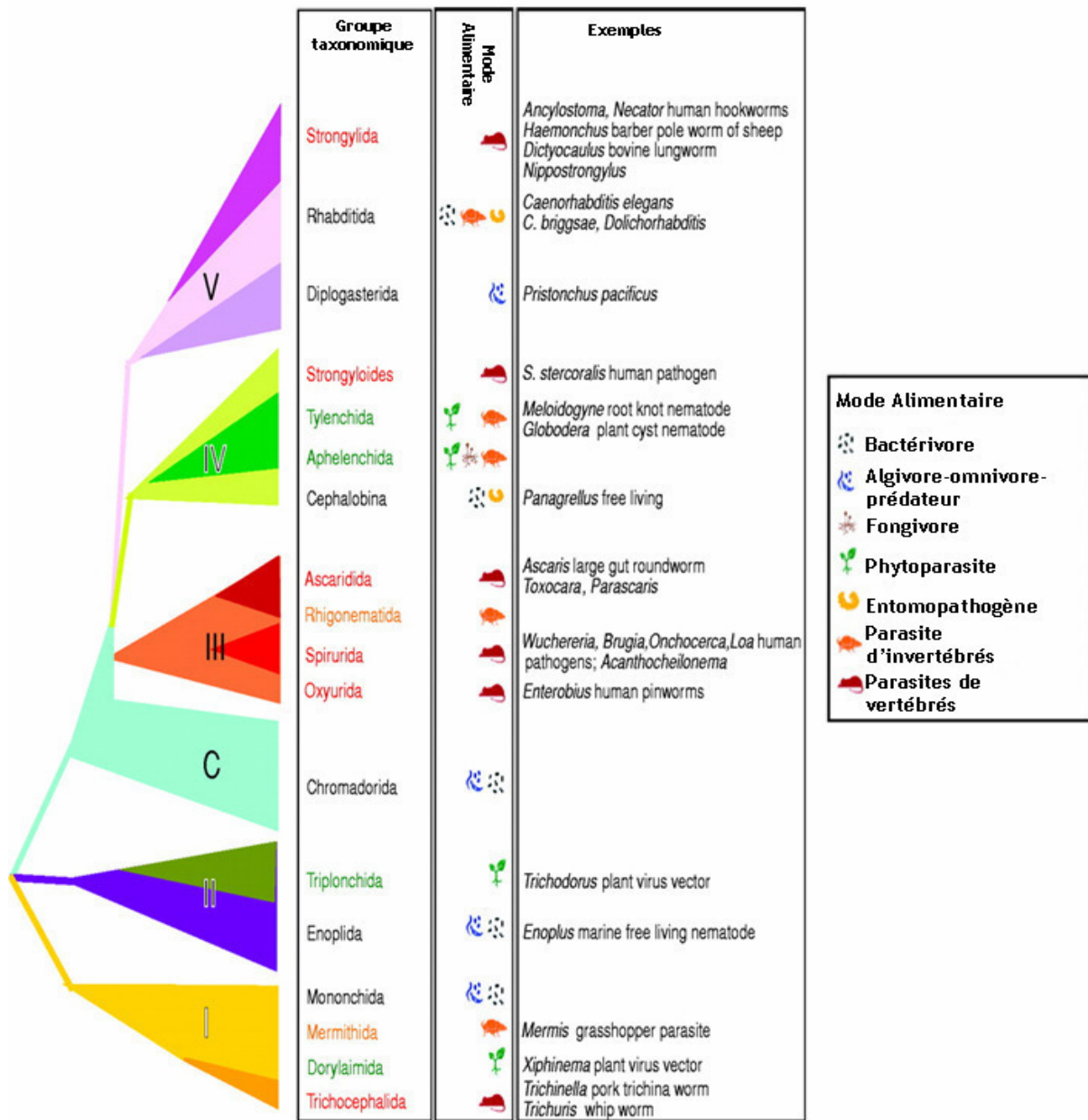


Figure 3 : Phylogénie du phylum Nematoda basée sur la séquence complète de la petite sous-unité de l'ADN ribosomale (Blaxter *et al.*, 1998).

### 1.3. Organisation structurale des nématodes

Les nématodes sont des organismes vermiformes à symétrie bilatérale recouverts d'une cuticule continue et souple mais très résistante. Ils sont ainsi contraints à croître de façon discontinue en passant par quatre mues larvaires avant d'atteindre la forme adulte. Même si leur taille est très variable, de 100 µm à 6 m (Blumenthal *et al.*, 2004), l'immense majorité des espèces ne dépasse pas 1 à 2 mm. Ils possèdent une musculature longitudinale qui entoure le tube digestif rectiligne - se terminant par la bouche et l'anus aux extrémités - et les glandes génitales (Fig. 4). Les cellules longitudinales musculaires sont connectées aux cordes nerveuses par des expansions (cellules neuromusculaires). Les nématodes possèdent des organes sensoriels, les amphides situées à l'extrémité antérieure et les phasmides à l'extrémité postérieure. Les nématodes n'ont ni système circulatoire, ni système respiratoire. Enfin ces organismes possèdent un hypoderme produisant deux cordes longitudinales hébergeant les cordes nerveuses.

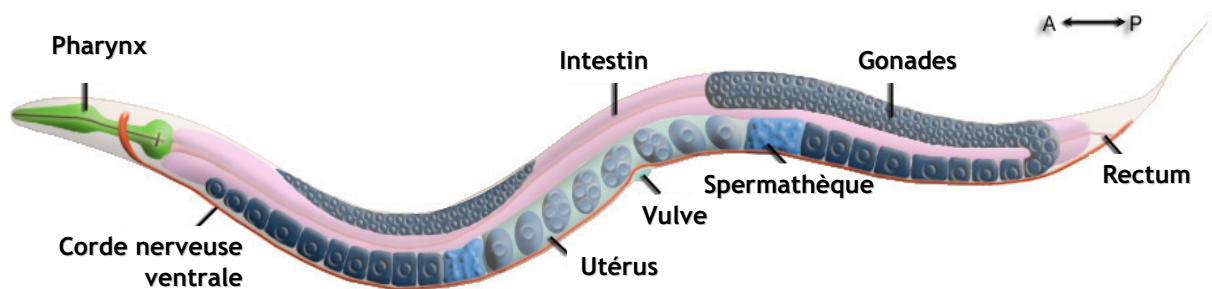


Figure 4 : Structure d'un nématode (d'après by Z. F. Altun & D. H. Hall, dans Atlas of *C. elegans* anatomy, <http://www.wormatlas.org/handbook/contents.htm>).

## 2. LES NEMATODES PHYTOPARASITES : GENERALITES

### 2.1. Les stratégies parasitaires des nématodes phytophages

Les nématodes sont le plus souvent invisibles à l'œil nu (0,2 à 3 mm de long) et parasitent les parties aériennes et souterraines des plantes. Ils ne provoquent pour la plupart pas de symptômes visibles caractéristiques sur les parties aériennes des plantes, ce qui rend le diagnostic délicat. Comme nous l'avons vu précédemment, les nématodes phytoparasites regroupent quatre ordres : les Aphelenchida (ex : *Aphelenchoides* spp) qui sont des parasites des parties aériennes, les Dorylaimida (ex : *Longidorus* spp) et les Triplonchida (ex : *Trichodorus* spp) qui sont des parasites des racines, et enfin les Tylenchida qui parasitent les racines, les tubercules ou les rhizomes, ainsi que les parties aériennes pour certains genres. Bien qu'ayant divergé les uns des autres très précocement au cours de l'évolution du phylum Nemata (Fig. 3), ces quatre ordres possèdent un caractère commun dérivé, le stylet, aiguille creuse (pour la plupart) située dans la partie



antérieure. Cette structure leur permet de se nourrir des composés cellulaires de la plante hôte et est adaptée aux différentes stratégies parasitaires rencontrées dans les quatre ordres de nématodes phytoparasites (Fig. 5) de racine.

On distingue les ectoparasites, les ecto-endoparasites et les endoparasites migrants et sédentaires. Les ectoparasites ont des stylets plutôt longs et fins leur permettant de piquer les cellules végétales profondes depuis l'extérieur de la racine. Les nématodes à kyste, qui sont des endoparasites ont au contraire un stylet plutôt court et robuste pour leur permettre de détruire mécaniquement les cellules végétales et pénétrer la racine (Fig. 6). Les nématodes à kyste sur lesquels nous nous focaliserons par la suite sont des endoparasites sédentaires. Ils établissent une relation très élaborée avec leur hôte en mettant en place une structure très complexe qui leur permet de se nourrir lors de leur phase sédentaire sur laquelle nous reviendrons ultérieurement.

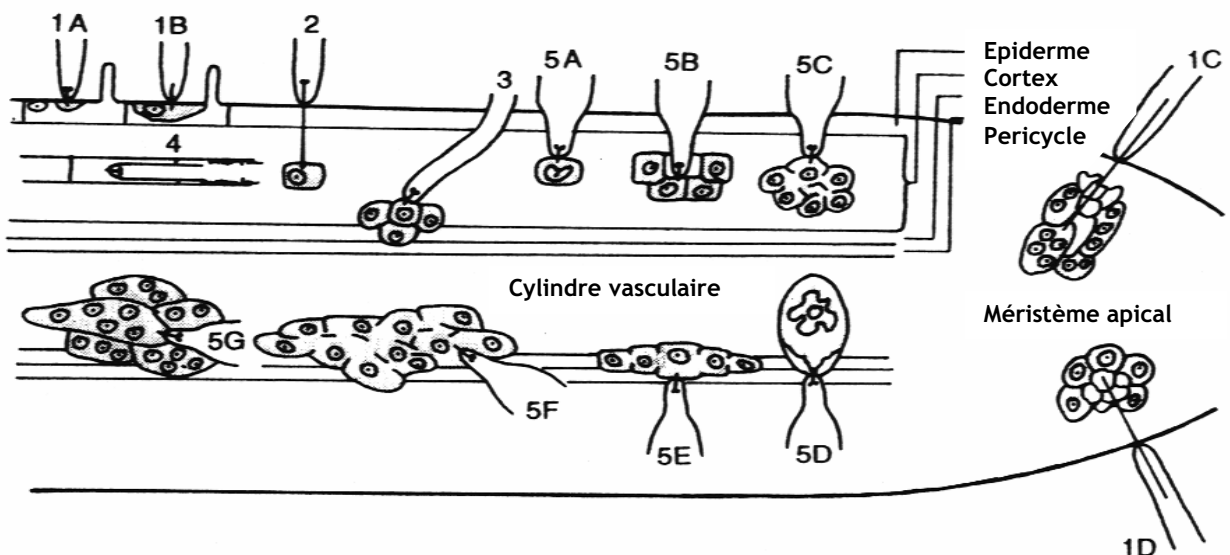


Figure 5 : Représentation des modes de parasitisme des nématodes parasites de racines, d'après Sijmons *et al*, 1994. **Ectoparasites migrants** : 1A : *Tylenchorhynchus dubius*, 1B : *Trichodorus* spp, 1C : *Xiphinema index*, 1D : *Longidorus elongatus*; **Ectoparasites sédentaires** : 2 : *Criconemella xenoplax*; **Ecto-endoparasites migrants** : 3 : *Helicotylenchus* spp; **Endoparasites migrants** : 4 : *Pratylenchus* spp; **Endoparasites sédentaires** : 5A : *Trophotylenchulus obscurus*, 5B : *Tylenchulus semipenetrans*, 5C : *Verutus volvingentis*, 5D : *Cryphodera utahensis*, 5E : *Rotylenchulus reniformis*, 5F : *Heterodera* spp, 5G : *Meloidogyne* spp.

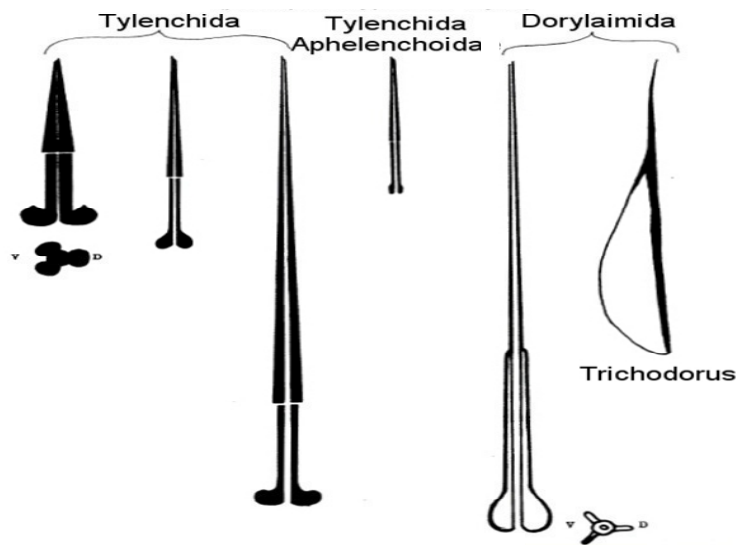


Figure 6 : Différentes structures des stylets et odontostyles des nématodes phytoparasites.

## 2.2. Dégâts liés aux nématodes phytophages

Les nématodes phytophages sont des parasites obligatoires occasionnant des dégâts considérables sur les grandes cultures à travers le monde, représentant un coût d'environ 100 milliards d'euros (Sasser *et al.*, 1987; Haq *et al.* 2004). En effet, pratiquement aucune culture n'échappe à l'attaque d'au moins une espèce de nématodes, même s'il existe des différences quantitatives importantes suivant les espèces. Les coûts engendrés par les attaques de nématodes sont imputables aux :

- baisses de rendement,
- problèmes de qualité des plantes (aspect) qui les rendent impropres à la commercialisation,
- augmentations d'irrigation pour pallier les perturbations subies par le système racinaire des plantes parasitées,
- interdictions d'exportation du fait du statut de quarantaine de certaines espèces,
- traitements nématicides très coûteux.

S'ajoutent à cela les problèmes environnementaux liés à la toxicité des produits nématicides actuellement sur le marché. Ainsi, les conséquences économiques et environnementales liées aux problèmes que posent les attaques de nématodes en matière de protection des plantes, sont à l'origine d'efforts importants pour mettre au point des méthodes de lutte durables et plus respectueuses de l'environnement. L'accent est particulièrement mis sur les nématodes qui causent le plus de dégâts, notamment les nématodes à kyste et à galle.

### 3. LES NÉMATODES A KYSTE : *GLOBODERA* & *HETERODERA*

Les nématodes phytoparasites sédentaires comprennent principalement les nématodes à kyste (*Globodera* spp. et *Heterodera* spp.) et les nématodes à galle (*Meloidogyne* spp.) qui causent les dégâts les plus conséquents sur les cultures. Ces genres ont des caractéristiques biologiques très proches et ont longtemps été regroupés dans la classification basée sur des critères morphologiques. En effet, si ces nématodes ont des gammes d'hôtes très différentes et génèrent des symptômes également très variés sur les racines hôtes, ils ont un cycle de développement très proche et mettent en place une relation avec leur hôte relativement similaire : un site nourricier. Pourtant les modifications cellulaires de la plante sont assez différentes. Récemment, suite à des études sur la phylogénie moléculaire des Tylenchida (Subbotin *et al.*, 2006 ; Holterman *et al.*, 2006), il s'est avéré que ces deux groupes de nématodes ne sont pas des groupes frères puisqu'ils ne partagent pas d'ancêtre commun proche. Leur mode de vie sédentaire serait donc apparu indépendamment à partir de groupes de nématodes migrants. Les similitudes de leur mode de vie et de leur morphologie (femelles renflées) ne seraient que des convergences évolutives. Étant donné cet éloignement phylogénétique apparent, les comparaisons entre ces deux groupes peuvent être difficiles, notamment d'un point de vue moléculaire. C'est pourquoi il semble plus judicieux de se focaliser dans un premier temps sur l'un ou l'autre des groupes lors des études moléculaires des interactions plante/nématode.

#### 3.1. Fonctionnement des populations de nématodes à kyste

Le cycle biologique des nématodes permet deux modes de dispersion :

1- actif : les juvéniles infestants se déplacent dans le sol même si leur petite taille (~ 500 µm) ne leur permet pas de migrer au delà d'un mètre. Ils ne peuvent donc pas se disperser facilement de champ en champ.

2- passif : les kystes, forme de survie des nématodes à kyste contenant les œufs peuvent facilement être transportés par les activités humaines (le matériel agricole, les chaussures, ou encore l'eau d'irrigation). Le vent peut également propager les kystes, très légers, à plus longue distance. Enfin, des transports occasionnels mais sur de très longues distances peuvent se faire au cours des importations/exportations de plantes ou de tubercules contaminés.

Connaissant les différents modes de dispersion possibles, Picard (2005) ont étudié la génétique des populations de *G. pallida* au Pérou et ont défini les contours d'une population d'un point de vue spatial (Picard et Plantard 2006) : dans un bassin de production de pomme de terre, les nématodes situés dans un rayon de 50 km sont génétiquement semblables et sont donc considérés comme appartenant à la même population. Une population est donc relative à un bassin de production. Picard *et al.* (2004) ont également montré que le maximum de variabilité génétique entre les individus est retrouvé à l'échelle d'une parcelle et non d'un champ ou d'une région. Ces données ont appuyé notre choix de travailler au niveau populationnel, et c'est en ce sens que nous utiliserons le terme population dans le reste du manuscrit.

### 3.2. Relations phylogénétiques et classification

Les études des relations phylogénétiques entre les nématodes ne sont pas seulement nécessaires pour établir la taxonomie des nématodes mais permettent aussi de proposer un cadre pour l'interprétation de l'évolution de ce groupe et de mieux comprendre la biologie de ces nématodes en tant que pathogènes de plantes. Différentes phylogénies des nématodes à kyste ont été proposées, basées sur des critères morphologiques. Cependant, elles se sont avérées relativement limitées par le petit nombre de caractères discriminants au sein de ces groupes morphologiquement très homogènes. Les études moléculaires ont permis d'apporter une autre vision sur l'évolution des caractères liés aux espèces, comme décrit par Subbotin *et al.* en 2001.

Les nématodes à kyste ont été classés en six clades définis sur la base des séquences ITS de 40 espèces et sous-espèces de la famille des Heteroderidae (Subbotin *et al.*, 2001; Fig. 7). Ces six clades représentent deux sous-familles :

- les Punctoderinae regroupant les espèces des genres *Punctodera*, *Cactodera*, *Dolichodera* et *Globodera* (Clade I)
- les Heteroderinae regroupant les *Heterodera* et les *Afenestrata* répartis en 5 clades (Clades II à VI).

Nous nous sommes intéressés dans cette étude aux genres *Globodera* et *Heterodera* qui sont deux genres à fort intérêt agronomique.

#### A. Les nématodes du genre *Heterodera*

Les nématodes du genre *Heterodera* sont plus polyphages que les *Globodera*. Ils sont divisés en six groupes d'espèces (Subbotin *et al.*, 2001 ; Fig. 7) : "Schachtii", "Goettingiana", "Avenae", "Humuli", "Cyperi" et "Sacchari". La phylogénie moléculaire des Heteroderidae a permis de mettre en évidence une coévolution de ces nématodes avec leur plante hôte (Subbotin *et al.*, 2001). Deux des groupes, Schachtii et Humuli, ont évolué avec les dicotylédones alors que les groupes Sacchari et Avenae semblent avoir évolué avec les monocotylédones. Parmi ces groupes, certains tels que Humuli ou Goettingiana ont développé la capacité à se développer sur une large gamme de familles botaniques (Fig. 7). Au contraire, d'autres sont très spécialisés d'une famille, comme les groupes Cyperi, Avenae et Sacchari qui se développent exclusivement sur les plantes de la famille des Poaceae, ou encore le groupe Schachtii dont les espèces se développent principalement sur les Fabaceae, excepté *H. schachtii* qui semble avoir acquis la capacité à se développer sur d'autres familles botaniques telles que les Chenopodiaceae ou les Brassicaceae.

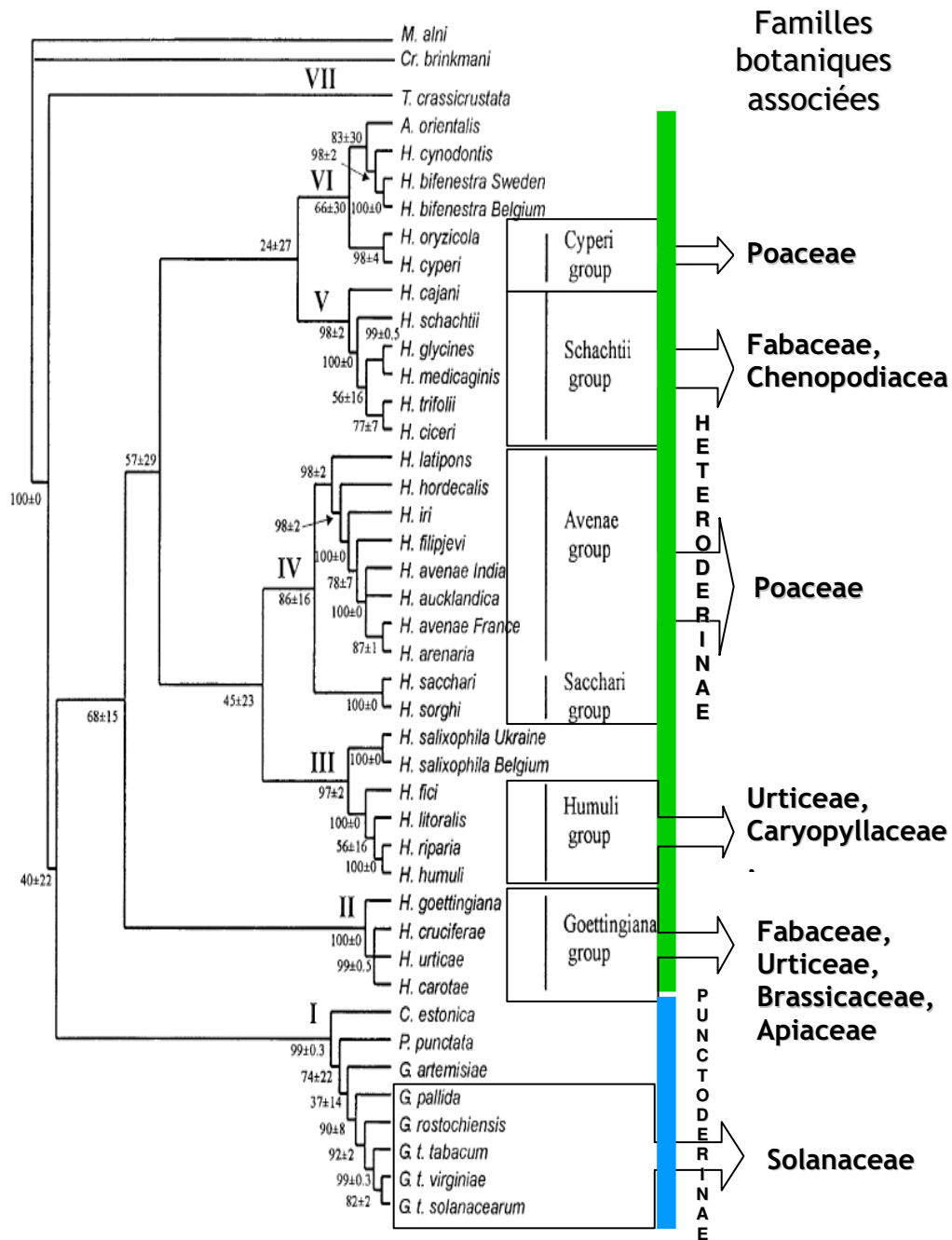


Figure 7 : Phylogénie des nématodes à kyste basée sur les séquences des régions ITS de 40 espèces et sous-espèces de la famille des Heteroderidae (d'après Subbotin *et al.*, 2001).

## B. Les nématodes du genre *Globodera*

Les nématodes à kyste du genre *Globodera* sont des nématodes présentant une très grande spécificité d'hôtes et sont principalement inféodés aux Solanacées (Fig. 7), excepté par exemple *G. artemisiae* qui se développe aussi sur les Asteraceae et les Rosaceae. Trois espèces principales composent ce genre : *G. pallida* (Stone, 1972) originaire de la moitié nord des Andes (Pérou), *G. tabacum* (Lownsbery & Lownsbery, 1954) originaire d'Amérique du nord (Virginie, Caroline) et *G. rostochiensis* (Wollenweber, 1923) originaire de la moitié sud des Andes (Bolivie). Actuellement, les nématodes à kyste de la pomme de terre, *G. pallida* et *G. rostochiensis*, sont définis comme des organismes de quarantaine. (Voir site internet OEPP - Organisation Européenne et méditerranéenne pour la Protection des Plantes - [www.eppo.org/QUARANTINE/nematodes/Globodera\\_rostochiensis/F-hetdro.pdf](http://www.eppo.org/QUARANTINE/nematodes/Globodera_rostochiensis/F-hetdro.pdf)). A ces trois espèces s'ajoute *G. "mexicana"*\*, espèce très proche de *G. pallida* originaire du Mexique décrite dans la thèse non publiée de Campos-Vela (1967).

Les différences morphologiques ou de gamme d'hôtes entre les espèces sont subtiles, excepté pour *G. rostochiensis*. Mugniéry *et al.* (1992) ont montré que ces quatre espèces sont hybridables, mais que les croisements donnent rarement des descendants viables et féconds.

Les analyses phylogénétiques dont nous disposons à ce jour se sont focalisées sur les relations inter-spécifiques. A l'échelle spécifique, peu de travaux consacrés à l'analyse de la variabilité des populations indigènes ou importées ont été menés. Picard (2005) a récemment montré une grande diversité génétique des populations de *G. pallida* péruviennes. Nous reviendrons sur ce point dans le chapitre 3. Sur la base des analyses ITS de Subbotin *et al.* (2001), les espèces *G. rostochiensis* et *G. tabacum* sont regroupées et *G. pallida* constitue un groupe extérieur (Fig. 7). Marché *et al.* (2001) ont montré que l'espèce *G. tabacum* présente au moins quatre sous-espèces présentant des pathogénies différentes sur une gamme de variétés de tabac.

## 4. METHODES DE LUTTE CONTRE LES NEMATODES À KYSTE

L'utilisation de substances chimiques est actuellement la méthode de lutte la plus utilisée. Néanmoins, face à l'interdiction progressive et planifiée (plus d'utilisation des fumigants d'ici à 2009) des nématicides du fait de leur toxicité pour l'environnement ainsi que pour les utilisateurs, il est urgent de proposer des méthodes de lutte alternatives.

Certaines relevant de pratiques culturales ou d'interventions techniques existent mais ne semblent pas être suffisamment efficaces : prophylaxie, lutte culturale, lutte physique (voir encadré 1). La lutte biologique a également été envisagée mais ne semble pas être applicable en conditions de culture pour le moment (voir encadré 1). Actuellement, les méthodes de lutte les plus prometteuses reposent sur l'utilisation de variétés résistantes suffisamment efficaces et qui peuvent assurer un contrôle durable des populations de nématodes.

\* les guillemets systématiquement ajoutés au nom d'espèce "mexicana" signifient que cette espèce n'a été décrite que dans la thèse non publiée de Campos-Vela en 1967 et n'a pas été validée selon le code international de nomenclature zoologique.

## ENCADRE N° 1 : METHODES DE LUTTE CONTRE LES NEMATODES À KYSTE

### 1- Prophylaxie : éviter la dissémination des pathogènes

- contrôle des végétaux aux frontières pour éviter l'introduction de nouvelles populations sur un territoire
- nettoyage des machines agricoles pour éviter les contaminations inter parcelles
- rotations de culture pour éviter la multiplication du pathogène

Cependant les capacités de survie des nématodes dans le sol sont supérieures à dix années et rendent cette méthode difficilement applicable. Il est admis qu'un minimum de sept ans est nécessaire entre deux cultures de pomme de terre (Mugniéry & Phillips, sous presse).

### 2- Lutte culturale

Il existe trois types de lutte culturale :

- utiliser des variétés qui résistent le mieux aux attaques de nématodes
- abaisser le niveau de population au dessous du seuil de nuisibilité par utilisation de plantes nématicides ou de plantes pièges (Scholte *et al.*, 2000)
- modifier les pratiques culturales pour éviter la multiplication du nématode : récolte précoce des pommes de terre avant maturité des nématodes, par exemple.

Ces méthodes ne sont pas forcément les plus adaptées ou les plus facile à mettre en place.

### 3- Lutte physique

Il existe deux moyens de lutte physique :

- la solarisation (augmentation de la température du sol, en surface, par bâchage)
- l'inondation (les nématodes meurent par asphyxie).

Ce sont deux moyens très peu utilisés pour des raisons pratiques (manque d'ensoleillement, utilisation des parcelles difficile après inondation, coût).

### 4- Lutte chimique

Il existe trois types de traitements chimiques :

- 1- les fumigants qui ont des propriétés nématicides, mais aussi bactéricides, fongicides et herbicides
- 2- les organo-phosphorés et 3- les carbamates qui sont aussi insecticides.

Ils sont très efficaces, induisant 80 à 90% de mortalité. Cependant, en Europe, leur utilisation est limitée ou interdite du fait de leur toxicité pour l'environnement et pour l'utilisateur.

### 5- Lutte biologique

Certains micro-organismes sont connus pour contrôler des nématodes phytoparasites des genres *Globodera*, *Meloidogyne* et *Pratylenchus* :

- champignons du sol : prédateurs (*Arthrobotrys irregularis*) ou parasites des nématodes (*Paecylomyces lilacinus*)
- certaines bactéries du genre *Bacillus* ou *Pasteuria*

Si ces méthodes semblent parfois efficaces au laboratoire et en milieu clos (sous serre), elles paraissent difficilement applicables au champ du fait de la présence d'antagonistes potentiels dans l'environnement.

### 6- Lutte génétique : utilisation de résistances naturelle et/ou artificielles

Les détails de ces méthodes sont développés dans la partie 3.

#### 4.1. Lutte génétique : utilisation de variétés résistantes

Les plantes sont constamment confrontées aux attaques de pathogènes de toutes natures (virus, bactéries, champignons, insectes, nématodes). Pour se défendre, elles ont développé au cours de leur évolution un système de défense constitué par un large spectre de gènes de résistance. La difficulté est d'identifier ces ressources génétiques afin de les utiliser pour lutter contre les pathogènes. En nématologie, une plante est considérée comme résistante aux nématodes à kyste lorsqu'elle permet de réduire très fortement ou totalement le nombre de femelles formées par rapport à un témoin non résistant. Chez ces nématodes, deux modes de résistance ont été observés en fonction de leur délai d'action après attaque par le parasite (Mugniéry *et al.*, 2001). Les réactions les plus précoces sont souvent des réactions d'hypersensibilité pendant l'induction du site nourricier. Cette réaction semble être identique à celles observées contre d'autres pathogènes (Williamson *et al.* 2006). Les juvéniles de second stade n'évoluent alors pas jusqu'au stade adulte. Un deuxième type de réaction rencontré chez les plantes provoque la masculinisation des nématodes adultes. C'est notamment le cas du gène *Hero A* qui initie tardivement la réponse à une infestation par le nématode, induisant une atrophie ou un développement anormal du site nourricier provoquant la formation de mâles (la détermination du sexe chez les nématodes à kyste est épigénétique).

Même si l'utilisation de résistances naturelles semble être la méthode la plus attractive notamment parce qu'elle est respectueuse de l'environnement et moins coûteuse que l'application de nématicides, elle présente cependant certaines limites dans l'état actuel des connaissances. Les sources de résistances utilisées présentent souvent une grande spécificité et ne sont efficaces que contre une espèce, voire une population. C'est le cas notamment du gène de résistance *H1* de la pomme de terre qui ne permet de contrôler que les populations de *G. rostochiensis* de pathotype Ro1/4 ou du gène *Gpa2* qui ne contrôle que deux populations hollandaises (Tableau 1). Contre les autres pathotypes de *G. rostochiensis* ou *G. pallida*, seules des résistances partielles ont été identifiées tel que de gène *HeroA* (Tableau 1). L'utilisation des sources de résistance nécessite est de s'assurer de leur durabilité dans le temps. Des cas de contournement de gènes de résistance utilisés contre *H. schachtii* (Hs1<sup>pro1</sup>, Müller *et al.*, 1992) ou contre *G. rostochiensis* (H1) ont été recensés. Le risque principal lié à ces phénomènes est l'appauvrissement rapide des ressources génétiques disponibles. En effet, chaque événement de ce type ne permettrait plus d'utiliser les résistances dont nous disposons pour lutter contre les nématodes. De plus les contournements des résistances favorisent la sélection de populations de nématodes de plus en plus virulentes. Une utilisation raisonnée de ces résistances est donc nécessaire afin d'en assurer la durabilité. Une autre limite de l'utilisation de variétés de plantes résistantes est le risque de modification de l'équilibre des populations dans le sol ce qui ne ferait que déplacer le problème. C'est le cas notamment de l'utilisation intensive du gène *H1* aux Pays-Bas, efficace contre une population de *G. rostochiensis*, qui a été utilisé durant dix années consécutives et qui a permis de réduire le niveau d'infestation par ce nématode, mais en contrepartie, les populations de *G. pallida* initialement présentes en faible proportion dans le sol, se sont développées.



Les recherches actuelles portent principalement sur la compréhension des interactions entre un hôte et son ou ses pathogènes. Les nouvelles données devraient permettre de réfléchir aux moyens les plus efficaces de préserver ces résistances naturelles et également de proposer de nouvelles méthodes de lutte. En effet, à défaut de gènes de résistance efficaces vis-à-vis de plusieurs espèces et durables, ceux-ci pourraient être complétés par des résistances dites "artificielles", basées sur les connaissances moléculaires en matière d'interaction entre le nématode et sa plante hôte.

Tableau 1 : Gènes de résistance aux nématodes endoparasites sédentaires clonés et cartographiés (d'après Williamson *et al.*, 2006).

Gène	Plante	Nématode
<b>Clonés</b>		
<i>Hs1<sup>pro-1</sup></i>	Betterave	Nématodes à kyste de la betterave : <i>Heterodera schachtii</i>
<i>Mi-1</i>	Tomate	Nématodes à galle : <i>Meloidogyne incognita</i> , <i>M. javanica</i> , <i>M. arenaria</i> . (également efficace contre le puceron de la pomme de terre <i>Macrosiphum euphorbiae</i> et la mouche blanche <i>Bemisia tabaci</i> )
<i>Hero A</i>	Tomate	Nématode à kyste de la pomme de terre : <i>Globodera rostochiensis</i> , pathotypes <i>Ro1</i> , <i>Ro3</i> et <i>Ro5</i> ; <i>Globodera pallida</i> , pathotypes <i>Pa2et Pa3</i> , et <i>Pa2/3</i>
<i>Gpa2</i>	Pomme de terre	Nématode à kyste de la pomme de terre : <i>G. pallida</i> (populations <i>D383</i> et <i>D372</i> )
<i>Gro1-4</i>	Pomme de terre	Nématodes à kyste de la pomme de terre : <i>G. rostochiensis</i> , pathotype <i>Ro1</i>
<i>Rhg1</i> et <i>Rhg4</i>	Soja	Nématode à kyste du soja : <i>Heterodera glycines</i> type 0
<b>Cartographiés<sup>a</sup></b>		
<i>H1</i>	Pomme de terre	Nématodes à kyste de la pomme de terre : <i>G. rostochiensis</i> , pathotypes <i>Ro1/4</i>
<i>Mi-3</i>	Tomate	Nématodes à galle : <i>Meloidogyne incognita</i> , <i>M. javanica</i> , <i>M. arenaria</i>
<i>Mi-9</i>	Tomate	Nématodes à galle : <i>M. incognita</i>
<i>Cre1</i>	Blé	Nématodes à kyste des céréales : <i>Heterodera avenae</i> , pathotype australiens et européens
<i>Cre3</i>	Blé	Nématodes à kyste des céréales : <i>H. avenae</i> , pathotype Australiens
<i>Ma</i>	Prunier	Nématodes à galle : toutes les espèces testées
<i>Hsa-1<sup>0g</sup></i>	Riz	Nématodes à kyste : <i>Heterodera sacchari</i>
<i>Me3</i>	Piment	Nématodes à galle : <i>Meloidogyne incognita</i> , <i>M. arenaria</i> , <i>M. javanica</i> et quelques populations de <i>M. hapla</i>
<i>Rmc1</i>	Pomme de terre	Nématode à galle : <i>Meloidogyne chitwoodi</i> , <i>M. fallax</i> et quelques populations de <i>M. hapla</i>

<sup>a</sup> Cette liste de gènes cartographiés n'est pas exhaustive. En bleu sont indiqués les gènes de résistance aux nématodes du genre *Globodera*.

## 4.2. Lutte génétique : résistances artificielles

Les données moléculaires acquises et en cours d'acquisition sur les mécanismes moléculaires qui régissent le succès du processus parasitaire du nématode sur son hôte ont une importance d'un point de vue cognitif mais également appliqué. En effet, de meilleures connaissances du

fonctionnement du pathogène permettraient d'avoir une meilleure vision des stratégies de lutte à adopter. Si les molécules clé du succès parasitaire étaient identifiées, tels que les gènes du pouvoir pathogène, elles pourraient constituer des cibles d'intérêt pour la création de résistances artificielles, autrement dit de plantes génétiquement modifiées capables d'inhiber la mise en place du parasitisme par les nématodes.

Les deux stratégies de création de résistances artificielles sont :

1- d'empêcher l'alimentation du nématode par la destruction ou l'atténuation spécifique du fonctionnement du site nourricier (ex : composés phytotoxiques),

2- de cibler directement le nématode (locomotion, perception chémosensorielle, neurotransmission, mue, digestion, reproduction, développement...).

Même si cette deuxième possibilité est attrayante du fait du large spectre de nématodes qui pourrait être ciblé, la première stratégie semble la plus envisageable au regard des connaissances actuelles sur le système sécrétoire des nématodes. Ainsi, des études se sont focalisées sur l'utilisation de techniques pouvant cibler les sécrétions salivaires et d'autres plus prospectives sur l'extinction de gènes du parasite. Nous allons illustrer ces études par des exemples qui ont été développés.

#### **A. Cibler les protéines sécrétées par le nématode**

Les sécrétions salivaires semblent être une cible d'intérêt pour bloquer le nématode dans sa phase de migration ou lors de l'initiation de son site nourricier. Une possibilité serait de faire produire à la plante des anticorps, appelés planticorps, spécifiquement dirigés contre une protéine impliquée dans le pouvoir pathogène des nématodes à kyste. Des tests ont d'ores et déjà été faits sur des plantes (Baum *et al.*, 1996) ou des protoplastes (Rosso *et al.*, 1996) de tabac, en faisant exprimer des anticorps monoclonaux ou simplement les chaînes légères des anticorps spécifiquement dirigés contre les sécrétions de nématodes. Si la spécificité des anticorps exprimés semblait être satisfaisante avec les protoplastes, aucune incidence sur le parasitisme n'a été obtenue avec les plants de tabac transgéniques infectés par les nématodes.

Afin d'empêcher l'alimentation du nématode au cours du parasitisme, les protéases semblent être des cibles de choix puisqu'elles sont spécifiques du mode d'alimentation des nématodes sédentaires et qu'elles sont vitales pour ces nématodes. Urwin *et al.* (2001) ont mené la seule étude en plein champ sur l'impact de l'utilisation de variétés de pomme de terre modifiées, exprimant une cystatine de façon constitutive. La cystatine est un inhibiteur de protéases à cystéine notamment de celles qui sont sécrétées par les nématodes sédentaires. Ainsi des variétés de *Solanum tuberosum tuberosum* transformées ont été cultivées pour lutter contre les populations de *G. pallida* du sol. S'ils ont observé qu'il n'y a pas d'effet sur le poids des tubercules, ni sur la croissance de la plante, ils n'ont pas constaté de meilleure résistance qu'une variété semi résistante déjà utilisée au champ. De plus cette cystatine étant exprimée de façon systémique dans la plante, le problème de la spécificité de la cible se pose en utilisant ce type de construction.

### **B. Cibler directement les gènes de nématodes**

Si le "silencing" ou extinction des gènes de nématode est actuellement une méthode utilisée pour démontrer l'implication d'un gène dans l'interaction avec la plante, cette méthode pourrait être envisagée en tant que méthode de lutte (Bakhetia *et al.*, 2005). Il s'agit d'inhiber l'expression d'un gène par fixation d'ARNs de petite taille empêchant la transcription. De nombreuses recherches sont actuellement en cours sur les nématodes sédentaires dont le pompage d'une solution peut être stimulé à l'aide d'agents chimiques tels que l'octopamine. Les ARNs double brin utilisés comme interférents sont alors ingérés par le nématode induisant le phénomène de silencing. Il existe deux obstacles principaux à l'utilisation de cette méthode pour la lutte : 1- il faut que la plante exprime spécifiquement des ARNs double brin au niveau des cellules nourricières afin que le nématode les ingère lorsqu'il s'alimente sur la plante et 2- il faut que l'efficacité du silencing soit durable. Pour le moment les études menées notamment sur les nématodes à galle (Rosso *et al.*, 2005) montrent que l'efficacité du silencing induit de cette manière est relativement transitoire (retour à 80% d'expression du gène après 68h) compte tenu de la durée du cycle de développement du nématode (30 jours).

### **C. Limite des résistances artificielles**

Quelle que soit la méthode utilisée, le problème de spécificité est la principale préoccupation lorsqu'il s'agit de créer des plantes transgéniques. En effet, il faudrait combiner une spécificité cellulaire pour conserver une plante en état de se développer, une spécificité tissulaire pour que la plante soit propre à la consommation sans risque pour le consommateur - l'idée étant de ne cibler que les parties radiculaires, non consommées -, et une spécificité d'action sur les nématodes et pas sur d'autres micro-organismes du sol pouvant s'alimenter sur la plante. Ces stratégies nécessitent, par exemple, d'identifier un promoteur spécifique des cellules nourricières, de manière à ne faire exprimer le transgène qu'en présence du nématode, et uniquement dirigé contre ce pathogène. Un autre impératif qui vient se greffer à cette première barrière, est le spectre de nématodes visé. En effet, si des études sont engagées dans ce domaine, l'objectif n'est pas de cibler une population ou une espèce de nématode mais plutôt un spectre le plus large possible. Dans un souci de durabilité des résistances mises en place, le choix de la cible peut s'avérer relativement difficile. En effet, il faut choisir un gène clé, qui a évolué suffisamment pour être spécifique des nématodes ciblés mais pas trop variable pour qu'il ne soit pas spécifique d'une seule espèce ou d'une seule population de nématodes. Des études moléculaires évolutives de ces gènes semblent alors nécessaires de manière à identifier les pressions de sélection subies par les gènes entiers ou par des domaines fonctionnels. En effet, afin de créer des résistances artificielles spécifiquement dirigées contre un gène ou une portion de gène, les données sur l'évolution moléculaire de celui-ci seront très utiles pour évaluer la spécificité et également de la durabilité de la résistance qui pourrait y être associée.

Ce chapitre montre clairement les difficultés rencontrées en matière de lutte contre les nématodes à kyste. Un moyen pour pallier ces déficiences et raisonner les méthodes de lutte actuelles, est 1- de connaître les gènes du pouvoir pathogène, 2- d'appréhender l'évolution de ces gènes au sein des populations, des espèces, voire des genres pour estimer l'efficacité dans le temps et dans l'espace des moyens de lutte basés sur ces données. Pour cela, il semble nécessaire

d'acquérir un maximum de connaissances sur les interactions plante/nématode d'un point de vue moléculaire afin de mieux comprendre le fonctionnement des résistances actuellement identifiées et utilisées, et de proposer de nouvelles méthodes de lutte. En effet, une meilleure connaissance des molécules régissant le succès parasitaire de la plante devrait permettre 1- de disposer d'un plus large éventail de possibilités de méthodes de lutte, 2- d'être en mesure de faire des choix sur les cibles les plus appropriées afin de gérer durablement les populations de nématodes phytoparasites.

Nous allons développer dans les parties suivantes l'état des connaissances actuelles, les informations qu'elles nous apportent dans la compréhension des mécanismes moléculaire qui existent entre les nématodes phytoparasites et leur plante hôte, ainsi que l'apport de ces connaissances en matière de lutte contre les nématodes à kyste.

## **5. BIOLOGIE DES NEMATODES A KYSTE : *GLOBODERA* & *HETERODERA***

### **5.1. Cycle de développement**

Afin de lutter contre ces nématodes, il apparaît essentiel de connaître leur biologie afin d'identifier les étapes clés du parasitisme et notamment les mécanismes moléculaires régissant l'induction des différentes étapes. En connaissant mieux les interactions et la biologie de ces organismes, il devrait être plus évident de définir les phases auxquelles il faut agir pour lutter de façon efficace contre ces parasites.

Les seconds stades larvaires (stades infestants, Fig. 8A) contenus dans les kystes perçoivent les exsudats radiculaires émis par la plante hôte et sortent du chorion (Fig. 9). Ils se dispersent alors dans le sol et se dirigent vers les zones d'élongation de la racine hôte. Les juvéniles pénètrent la racine de façon mécanique à l'aide de leur stylet perforateur dans la zone apicale et migrent de façon intra cellulaire jusqu'au cylindre central. Les juvéniles utilisent leur stylet pour percer une cellule du cylindre central et induisent un site nourricier appelé syncytium par digestion des parois des cellules adjacentes. Le cytoplasme de ces cellules devient dense, la vacuole se rétrécit et le noyau devient multilobé (Mugniéry & Phillips sous presse). C'est la phase de sédentarisation du nématode. Le syncytium peut être formé de près de quinze cellules et fait office de cellule nourricière permettant l'alimentation et le développement des nématodes. Ils déploient une grande diversité de protéines au cours de leur stratégie parasitaire. Les principales clés de réussite du parasitisme sont liées à leurs sécrétions salivaires qui sont supposées intervenir tant lors des phases de migration que lors de la mise en place du parasitisme. Bien que de nombreuses protéines aient déjà été identifiées, les mécanismes d'induction du syncytium sont encore mal connus. Une fois le site nourricier induit, les nématodes sédentarisés, perdent leur aspect vermiforme et subissent trois mues successives pour passer au stade adulte. Les mâles ressortent de la racine pour migrer activement dans le sol.

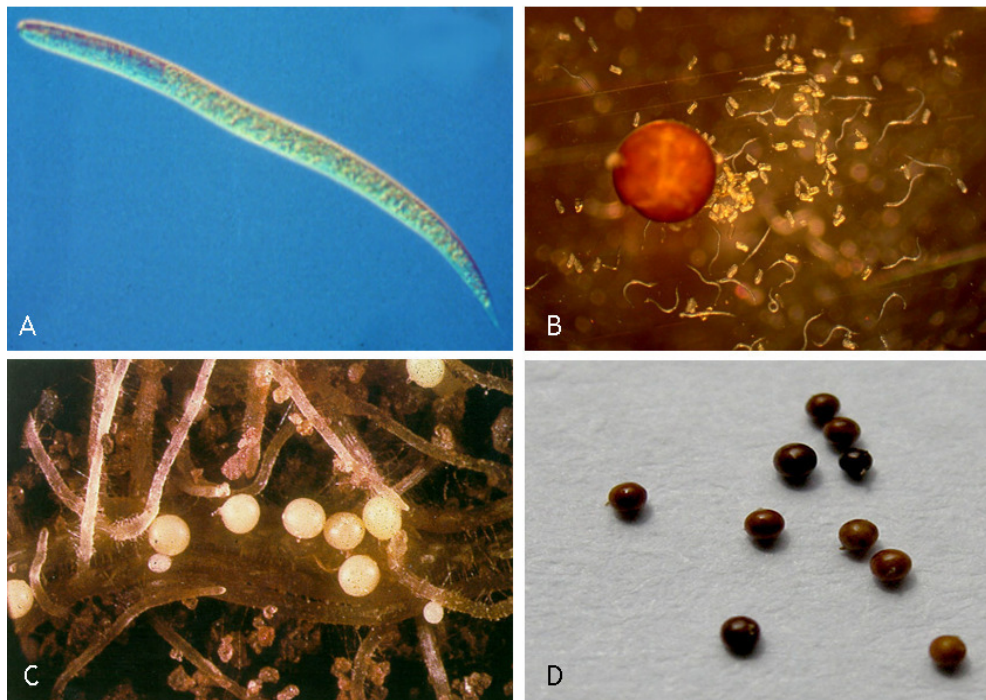


Figure 8 : Nématodes à kyste du genre *Globodera*. A : Juvénile de second stade de *G. pallida* (photo : S. Sardanelli), B : kyste ouvert de *G. pallida* (Photo : INRA), montrant les œufs de nématodes, C : femelles de *G. pallida* faisant saillie à la surface de racines de pomme de terre (Photo : Crown), D : kystes de *G. pallida* (Photo : INRA).

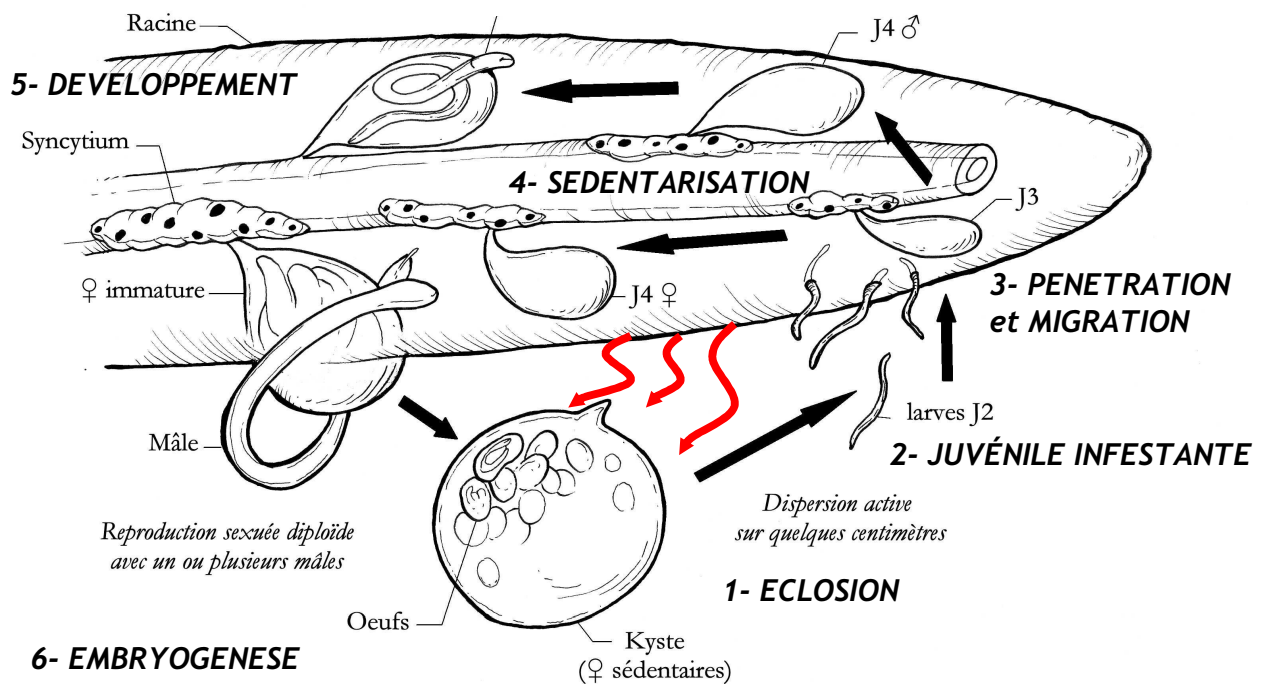


Figure 9 : Cycle biologique des nématodes à kyste du genre *Globodera*. Les flèches rouges symbolisent l'émission d'exsudats radiculaires par la plante hôte qui sont ensuite perçus par les nématodes dans le kyste.

Les femelles, quant à elles, se renflent et deviennent citriforme (présence d'un cône vulvaire) pour les *Heterodera* et piriforme (pas de cône vulvaire) pour les *Globodera* (Figure 8C). Elles font saillie à la surface de la racine hôte. Les nématodes à kyste ont une reproduction amphimictique. Les mâles sont attirés par les phéromones libérées par la cuticule de la femelle. La femelle fécondée produit des œufs puis meurt et se transforme alors en kyste (Figure 8D) qui se désolidarise de la racine. Ce kyste qui contient jusqu'à 1000 œufs (Figure 8B) qui entrent en diapause, peut résister de nombreuses années dans le sol. Le cycle de développement complet est achevé en 30 jours à 20°C. En général, une seule génération par an est observée en conditions naturelles chez les *Globodera*.

Une connaissance approfondie du cycle biologique a permis de définir les étapes dites "clés" du parasitisme, essentielles au succès du parasite à se développer à l'intérieur de l'hôte. Les recherches portent donc sur les mécanismes impliqués dans ces étapes afin de mieux déterminer les stratégies de lutte à mettre en place de façon efficace. Nous allons détailler dans cette partie les modifications physiologiques que subissent les cellules végétales lors des deux grandes étapes du parasitisme par les nématodes à kyste.

## 5.2. Organisation de la partie antérieure des nématodes

Les nématodes à kyste possèdent plusieurs organes de sécrétions : l'intestin, la cuticule, le pore excréteur, les amphides et les glandes salivaires. Les structures de la partie antérieure des nématodes à kyste (glandes salivaires et stylet) jouent un rôle central dans le parasitisme (pénétration et migration, mise en place et maintien du site nourricier). Nous allons donc détailler la structure du nématode et les organes mis en jeu au cours du processus parasitaire.

Chez les Tylenchida, l'œsophage comprend une partie glandulaire et une partie musculaire parfois confondues. La partie musculaire, appelée bulbe médian (Fig. 10), est une pompe aspirante et refoulante permettant notamment l'injection de sécrétions salivaires dans les cellules végétales *via* le stylet et l'absorption du contenu prédigéré des cellules (De Guiran, 1983). La partie glandulaire est composée de trois glandes salivaires : une dorsale et deux ventrales. Les deux glandes subventrales débouchent au niveau du bulbe médian et la glande dorsale débouche à la base du stylet (Fig. 10). Les glandes salivaires ont une structure en forme de bouteille, avec dans leur région basale, un lobe cellulaire et une extension vers l'ampoule. Les glandes subventrales sont connectées au lumen de l'œsophage par un canal possédant une valve élaborée, dans le bulbe médian, juste derrière la chambre de pompage qui joue un rôle de valve unidirectionnelle. La glande dorsale s'ouvre sur l'œsophage juste derrière la base du stylet (Smant *et al.*, 1998). Chaque glande constituée d'une cellule possède un gros noyau et un grand nombre de complexes de Golgi incluant de petites vésicules. Les granules de sécrétions migrent souvent antérieurement, à travers l'extension cellulaire, vers la partie distale de la glande salivaire appelée ampoule. Ces granules de sécrétion sont accumulés dans l'ampoule en attendant un signal de sécrétion (Smant *et al.*, 1998). Même si des études ont montré des différences d'activités des deux types glandes salivaires avec des rôles spécifiques aux différentes phases parasitaires (Davis *et al.*, 2000), leurs activités ne semblent pas si tranchées. En effet, les glandes salivaires des nématodes à kyste du genre *Globodera* semblent toutes trois actives dès l'éclosion de par la présence d'un grand nombre de granules de

sécrétions dès le stade infestant dans chacune des glandes (Mugniéry, com. pers.). Qin *et al.* (2001) ont également identifié par cDNA-AFLP des sécrétions provenant tant des glandes salivaires ventrales que de la glande dorsale dès ce stade larvaire.

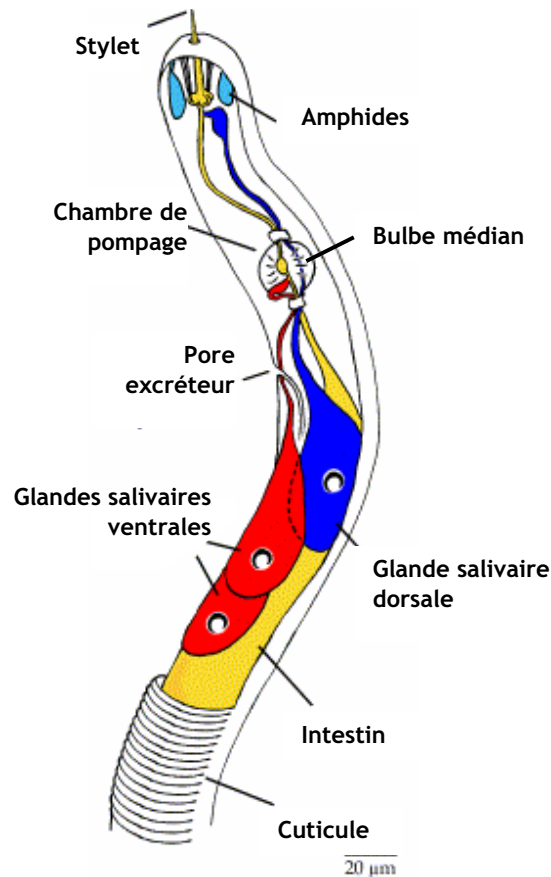


Figure 10 : Structure de la partie antérieure d'un nématode à kyste (D'après Vanholme *et al.*, 2004).

## 6. IMPLICATION DES SECRETIONS SALIVAIRES DANS LES ETAPES CLE DU PARASITISME

Les protéines sécrétées par les nématodes *via* leur stylet se retrouvent directement à l'interface hôte/parasite et de ce fait peuvent jouer un rôle fondamental dans le processus parasitaire ou sont directement injectés dans les cellules végétales (Fig. 11). Actuellement, peu de preuves directes existent quant au rôle des sécrétions salivaires dans le parasitisme. Dans certains cas comme par exemple pour les chorismates mutases et les cellulases, une activité enzymatique a été montrée, confortant l'hypothèse d'une activité au cours du parasitisme (Wang *et al.*, 1999, Lambert *et al.*, 1999; Gao *et al.*, 2004). Pour les autres candidats, nous ne disposons que des homologues identifiés dans les bases de données pour inférer une fonction potentielle aux nouveaux gènes de nématodes identifiés.



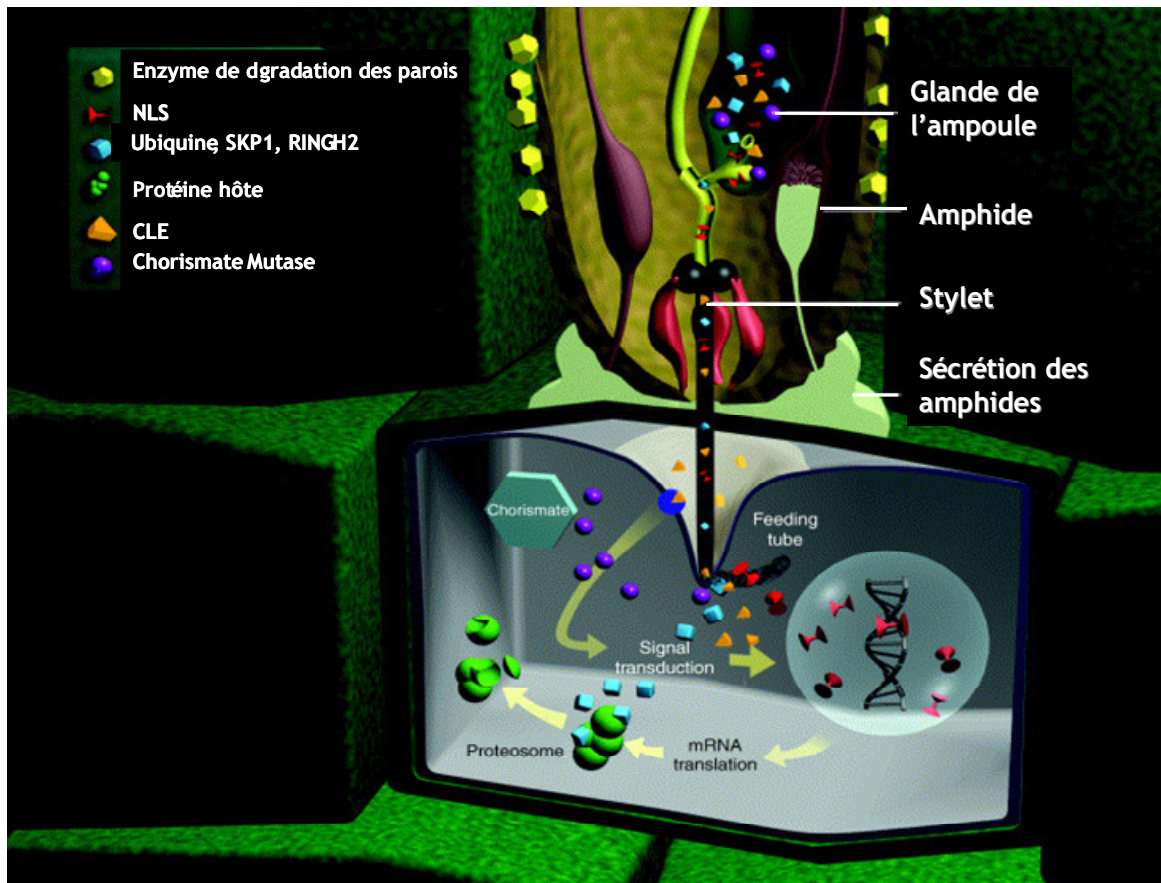


Figure 11 : Modèle d'interaction potentielle entre les produits de sécrétions des nématodes et la cellule végétale (d'après Davis *et al.* 2004).

## 6.1. Les phases précoces du parasitisme : pénétration et migration dans la racine hôte

### A. Conséquences physiologiques pour la plante hôte

Les phases dites précoces de l'invasion de la plante par les nématodes à kyste sont la pénétration dans la racine hôte et la migration intracellulaire jusqu'au cylindre central. Le juvénile de second stade attiré par les messages chimiques délivrés par les racines hôtes à travers les exsudats, se dirige vers les zones apicales des racines pour initier la pénétration. Les zones plus matures semblent être moins propices à la réussite du parasitisme à cette phase.

Si cette étape fait certainement appel à l'action mécanique puissante du stylet, elle nécessite probablement la mise en place d'une batterie d'enzymes, d'une part pour digérer les parois des cellules végétales et se frayer un accès au cylindre central de la racine, d'autre part pour se protéger contre les réactions de défense mises en place par la plante (Davis *et al.*, 2004). En effet, la migration intra cellulaire provoque des nécroses conséquentes des cellules végétales et donc des réactions importantes de la plante hôte pouvant nuire au nématode (production de composés toxiques d'oxygène par exemple).



### **B. Sécrétions salivaires impliquées dans les stades précoces du parasitisme**

Les protéines les mieux caractérisées sont les enzymes de dégradation des parois des cellules végétales telles que les cellulases qui dégradent la cellulose, les pectates lyases et polygalacturonases qui dégradent la pectine, les xylanases qui dégradent l'hémicellulose ou encore les expansines dont le rôle est mal connu (Tableau 2) (Smant *et al.*, 1998; Popeijus *et al.*, 2000; De Boer *et al.*, 2002; Grenier *et al.*, 2002; Gao *et al.*, 2002). Ces enzymes auraient un rôle dans la lyse des parois végétales durant la première phase parasitaire qui ne dure que quelques heures (6-12h) mais qui est néanmoins nécessaire au parasitisme. La première cellulase ( $\beta$ -1,4-endoglucanase) identifiée chez un animal a été mise en évidence en 1998 par Smant *et al.* chez le nématode à kyste de la pomme de terre *G. rostochiensis*. Cette protéine a été identifiée aussi bien chez les nématodes à kyste que chez les nématodes à galle et a été localisée le long du sillon de migration du nématode et pas dans le syncytium. Elle serait donc bien impliquée dans les phases précoces. L'ensemble des enzymes de dégradation des parois identifiées ont leurs transcrits localisés dans les glandes ventrales. Leurs activités cellulolytiques ont été également identifiées (Wang *et al.*, 1999; Goellner *et al.*, 2001) confirmant le rôle de ces enzymes dans les dégradations des parois des cellules végétales au cours du parasitisme.

Il est envisageable que la nature protéique précise de ces enzymes puisse varier en fonction de l'hôte si l'on considère qu'il y a pu avoir une coévolution entre les nématodes et leur plante hôte. Les données moléculaires de séquences des gènes et la mise en évidence de variants de ces gènes devraient permettre d'élucider progressivement cette hypothèse.

Les données de séquences obtenues jusqu'ici montrent une forte homologie avec des enzymes pectino-lytiques de procaryotes. Ceci a conduit à suggérer pour la première fois un transfert horizontal de ces gènes de micro-organismes phytoparasites du sol vers les nématodes (Smant *et al.* 1998; Jones *et al.*, 2005). Néanmoins, des biais liés aux méthodes d'identification des évènements de transferts horizontaux de gènes (problème de reconstruction phylogénétique, recherche du meilleur homologue par Blasts) semblent avoir été souvent à l'origine de ces hypothèses (Kurland *et al.*, 2003). Il semblerait que ce type d'évènement soit extrêmement rare du fait que les populations semblent être davantage enclin à utiliser la dynamique de variabilité génétique à laquelle ils sont soumis afin de mettre en place des mécanismes d'évolution rapide des séquences pour s'adapter, plutôt que d'inclure dans leur génome de l'ADN exogène. Cette hypothèse devrait être confirmée par l'étude de la variabilité génétique des nématodes en particulier des gènes du parasitisme.

D'autres gènes dont la fonction est inconnue, ont été identifiés dans les glandes ventrales des nématodes suggérant une implication potentielle dans ces phases précoces. C'est le cas de glycoprotéines (Duncan *et al.*, 1997), de métalloprotéases (Robertson *et al.*, 1999) ou encore de protéines assimilées à des protéines de venin d'hyménoptères (Gao *et al.*, 2001).

Tableau 2 : Produits sécrétés par les nématodes à kyste et caractérisés.

Gène	Espèce	Fonction possible	Références
β-1,4-endoglucanases	<i>G. rostochiensis</i>	Dégradation des parois des cellules végétales	Goellner <i>et al.</i> , 2001 ; Smant <i>et al.</i> , 1998 ; Davis <i>et al.</i> , 2004 ; Yan <i>et al.</i> , 2001 ; Gao <i>et al.</i> , 2002
	<i>G. tabacum</i>		
	<i>H. glycines</i>		
	<i>H. schachtii</i>		
Pectate lyase	<i>G. rostochiensis</i> <i>H. glycines</i>	Dégradation des parois des cellules végétales	De Boer <i>et al.</i> , 2002 ; Popeijus <i>et al.</i> , 2000
Expansine	<i>G. rostochiensis</i>	Dégradation des parois	Qin <i>et al.</i> , 2004 ; Kudla <i>et al.</i> , 2005 ; Wieckzorek <i>et al.</i> , 2006
Chorismate mutase	<i>H. glycines</i>	Modifie la balance en auxine ; formation du syncytium	Bekal <i>et al.</i> , 2003 ; Jones <i>et al.</i> , 2003
	<i>G. rostochiensis</i>		
Protéase	<i>G. rostochiensis</i>	Digestion extra-corporelle	Robertson <i>et al.</i> , 1999
Ubiquitine	<i>H. schachtii</i>	Mise en place du syncytium	Tytgat <i>et al.</i> , 2004
Thioredoxin peroxidase	<i>G. rostochiensis</i>	Protection contre les défenses des plantes	Robertson <i>et al.</i> , 2000
Protéine venom allergen-like	<i>H. glycines</i>	?	Gao <i>et al.</i> , 2001

## 6.2. Les phases tardives du parasitisme : élaboration et maintien du site nourricier

### A. Modifications physiologiques de la plante hôte

Pour compléter leur cycle de développement, les nématodes phytoparasites sédentaires sont dépendant de leur capacité à induire et maintenir leur site nourricier. Lorsque le nématode atteint le cylindre central de la racine, il met en place une nouvelle batterie de protéines pour induire son site nourricier. Cette structure particulière, appelée syncytium, à l'interface entre l'hôte et le parasite est une des structures les plus complexes observées dans une interaction hôte/parasite. Cette étroite relation permettant un dialogue très fin et spécialisé entre les deux partenaires est le résultat d'une coévolution. Le syncytium, multinucléé et élaboré, est une source nutritive essentielle (eau et solutés) pour le nématode qui est dans l'incapacité d'achever son cycle parasitaire s'il n'est pas fonctionnel. Celui-ci induit des modifications physiologiques profondes chez la cellule choisie pour élaborer ce site nourricier. Par ce biais, il est capable de détourner le système hydrique et le système de transport des nutriments à son propre profit (connexion des cellules nourricières avec le xylème de la plante; Hofmann *et al.*, 2006). De nombreuses modifications cytologiques sont alors observées. Après avoir choisi puis piqué une cellule végétale, le nématode induit la dissolution des parois des cellules adjacentes (10 -15 cellules) qui fusionnent et deviennent capables de multiplier leur matériel nucléaire (polyploïdisation) sans subir de mitose: c'est le phénomène d'endoreduplication (Fig. 12 A et B). De nombreuses modifications cytologiques sont également observées dans ces cellules en fusion, telles que l'augmentation de la densité du cytoplasme, la réduction de la vacuole, l'augmentation du nombre d'organelles (mitochondries, ribosomes, plastes, reticulum endoplasmique) : les cellules du syncytium sont hypertrophiées. Toutes ces modifications interviennent rapidement après l'arrivée du nématode au niveau du

cylindre central afin qu'il puisse achever son cycle de développement. Le nématode doit également maintenir ce site nourricier en activité jusqu'à l'achèvement de son cycle de développement.

Le syncytium est une structure qui doit être métaboliquement très active pour subvenir aux besoins nutritifs importants des nématodes. Les cellules nourricières sont des cellules végétales entièrement reprogrammées, dont l'expression des gènes est totalement manipulée et détournée par le nématode à son profit. L'infection par le nématode provoque par exemple des surexpressions de gènes de plante, dans le syncytium, codant des endoglucanases et des polygalacturonases (Goellner *et al.*, 2001) ou encore un homologue de pectine acétyl estérase chez *Arabidopsis* (Vercauteren *et al.*, 2002). Ces études ont montré que les dégradations des parois des cellules végétales mobilisées pour former le syncytium sont plutôt dues à une manipulation du génome des cellules végétales par le nématode, *via* ses sécrétions probablement, plutôt qu'à une action directe des sécrétions du nématode sur les parois de la cellule végétale. La réponse de la plante à l'infection par les nématodes à kyste semble mettre également en jeu les balances hormonales avec une sur-expression d'auxine et d'éthylène. Enfin certains transporteurs tels que les transporteurs sucrose sont également surexprimés dans le syncytium (Juergensen *et al.*, 2003). Ces modifications cellulaires et d'expression des gènes dans le syncytium sont stimulées par le nématode et rendent les cellules nourricières métaboliquement très actives et donc très riches en substances nutritives. D'autres gènes sont au contraire sous régulés, par exemple les gènes impliqués dans les mécanismes de défense des plantes vis-à-vis du nématode ou encore des régulateurs négatifs de l'auxine (Gheysen et Fenoll 2002). Le nématode est ainsi capable de manipuler le génome des cellules nourricières de manière à former, mais aussi à maintenir son site nourricier.

Côté nématode, les déterminants de ces modifications cellulaires restent actuellement inconnus. L'implication des sécrétions salivaires est fortement soupçonnée puisque ces sécrétions sont directement injectées au cœur des cellules végétales.

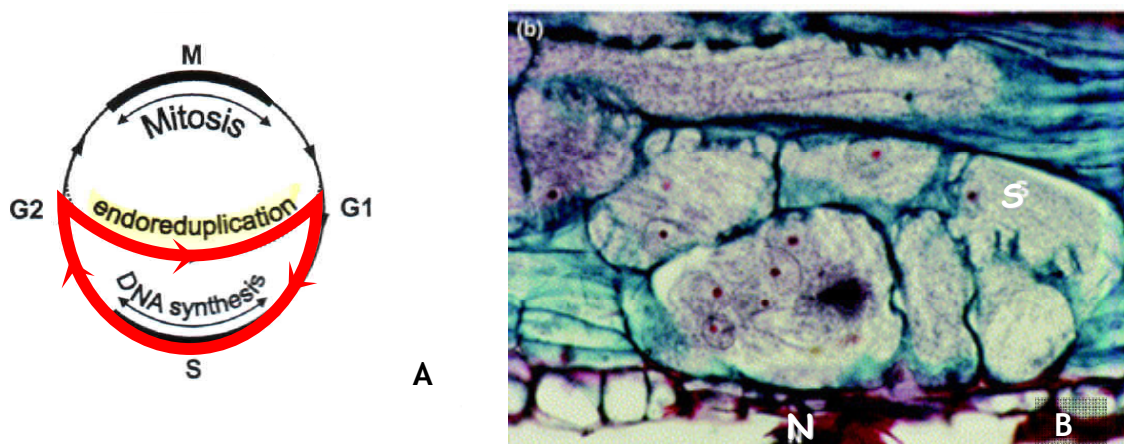


Figure 12 : Modifications cellulaires subies par les cellules végétales au cours du parasitisme par les nématodes à kyste. A : Schéma représentant les modifications du cycle cellulaire végétal. L'ADN nucléaire est répliqué durant la phase S (Synthèse). En rouge, le cycle cellulaire des cellules du syncytium dans lesquelles l'étape de mitose ne se fait pas (d'après Gheysen *et al.*, 2002). B : Photographie d'un syncytium induit par *H. glycines* (Davis *et al.* 2004). N : nématode, S : syncytium.

### **B. Sécrétions salivaires impliquées dans les phases tardives**

Les protéines sécrétées par le nématode et impliquées dans les étapes tardives présentent des transcrits avec des localisations cellulaires aussi bien dans les glandes ventrales que dans la glande dorsale. Certaines de ces protéines ont été identifiées comme intervenant dans les modifications des voies métaboliques de la cellule hôte. C'est le cas de la chorismate mutase. Cette enzyme, identifiée chez les nématodes à kyste (Tableau 2) et à galle, est une enzyme de la voie du shikimate, voie de biosynthèse des acides aminés aromatiques. Elle intervient notamment dans la synthèse des hormones végétales et également des composés liés à la défense des plantes tels que les phytoalexines ou l'acide salicylique. Il a été montré que le taux de phytohormones telles que l'auxine et l'éthylène augmente lors d'une infection par les nématodes. Ces hormones qui contrôlent aussi le cycle cellulaire végétal ont un rôle fondamental dans la mise en place du syncytium constitué de cellules végétales hypertrophiées (Goverse *et al.*, 2000). L'inhibition des défenses de l'hôte pourrait être un effet indirect des phytohormones (Davis *et al.*, 2004). Une hypothèse est que le nématode mobilise la voie de biosynthèse des hormones grâce à la sécrétion de chorismate mutase, au détriment de la voie de biosynthèse des composés de défense de la plante.

Les productions de la glande dorsale sont toutes aussi nombreuses que celles des glandes ventrales (Gao *et al.*, 2003) mais beaucoup moins connues (76% sont des nouveaux gènes). Des gènes de type régulateurs du cycle cellulaire ont été mis en évidence tels que *skp-1* (S-phase kinase-associated protein 1), RING-H2 et une ubiquitine (Gao *et al.*, 2003) qui sont des protéines formant un complexe de dégradation "ubiquitine-protéasome" intervenant dans les régulations cellulaires. Ces protéines sécrétées pourraient alors avoir un rôle dans les régulations cellulaires particulières du site nourricier et la régulation négative de l'expression de gènes impliqués dans les défenses des plantes (Davis *et al.*, 2004). Une protéine homologue d'une protéine exprimée dans le venin d'hyménoptère a été identifiée chez *H. glycines* (Tableau 2). Ce gène a un rôle non défini mais supposé dans l'induction du site nourricier (Davis *et al.*, 2004).

D'autres sécrétions semblent avoir un rôle dans les mécanismes d'induction du site nourricier, tels que des peptides CLE (CLAVATA3/ESR) localisés dans les glandes dorsales. En effet, des peptides ayant un rôle dans la signalisation (CLE) ont été mis en évidence chez les plantes et également chez les nématodes à kyste (Wang *et al.* 2005). Une hypothèse probable et fascinante serait que les nématodes à kyste ont évolué de façon à mimer les mécanismes moléculaires de leur plante hôte. Ces peptides pourraient être produits par le nématode pour stimuler les régulations du cycle cellulaire végétal et mettre en place leur site nourricier.

Enfin, les nématodes effectuent des digestions extracorporelles puisqu'ils ne possèdent pas de système de digestion. Des protéases sont sécrétées par le nématode pour digérer les composés cellulaires des cellules hôtes (Tableau 2). Ces enzymes sont essentielles à la survie du nématode dans la racine végétale pour se nourrir et se développer jusqu'aux stades adultes.

De nombreuses études sont actuellement menées pour identifier la fonction des protéines des sécrétions salivaires afin de mieux comprendre les interactions moléculaires engagées lors du processus parasitaire. Le volet à développer maintenant et qui n'a pas été abordé jusqu'ici est l'étude du potentiel évolutif de ces gènes afin notamment de mieux appréhender le développement de résistances artificielles. Ces données pourraient être également très utiles pour raisonner l'utilisation des résistances naturelles lorsque les interactions entre les gènes de résistance des plantes et les gènes du pouvoir pathogène des nématodes auront été élucidées. En effet, si les acteurs de l'interaction ainsi que leur mode évolutif sont connus, il sera alors plus aisé de transposer les données pour viser un maximum de nématodes avec les résistances dont nous disposons, tout en ayant une idée de la durabilité de celle-ci. Ces données devraient permettre de contourner les limites actuelles des méthodes de lutte, c'est-à-dire principalement éviter les contournements des résistances mises en place et utiliser les résistances les plus appropriées et efficaces en fonction des populations et espèces de nématodes présentes dans un sol.

Dans ce contexte, nous nous sommes attelés dans notre étude à identifier de nouveaux gènes du pouvoir pathogène comme cibles potentielles de lutte. De plus, nous avons développé une approche originale d'identification de la variabilité et de l'évolution moléculaire des gènes du pouvoir pathogène à l'échelle des nématodes à kyste, *Globodera* et *Heterodera*, qui n'avait jusqu'à présent jamais été entreprise.

# Résultats

Caractérisation de nouveaux gènes  
du pouvoir pathogène chez  
*Globodera pallida* et *G. "mexicana"*

*Partie II*

## PRESENTATION DE LA PARTIE II

La deuxième partie du manuscrit s'intègre dans la connaissance des mécanismes moléculaires régissant les interactions plante/nématode et en particulier les gènes impliqués dans le pouvoir pathogène des nématodes à kyste. L'introduction bibliographique montre que les mécanismes qui gouvernent le parasitisme sont encore mal connus et que des études sont encore nécessaires pour identifier davantage de gènes "clés" du parasitisme. Ces études devraient permettre d'élucider les voies du dialogue moléculaire fin qui régit les interactions entre le nématode et sa plante hôte.

Dans cette deuxième partie seront présentés les résultats de la caractérisation de gènes de nématodes du point de vue de leur structure, de leur expression et de leur variabilité, et ce, afin d'estimer leur implication dans le pouvoir pathogène.

Un premier chapitre est consacré aux gènes dont les transcrits ont été identifiés comme différentiellement exprimés entre deux nématodes à kyste très proches mais présentant des gammes d'hôte différentes. Ces gènes d'intérêt ont été caractérisés d'un point de vue séquence nucléotidique et expression cellulaire. Ces résultats font l'objet d'un article soumis à la revue *Journal of Heredity* et ont été présentés sous forme d'une communication affichée au congrès européen de nématologie à Rome en 2004 (annexe 4).

Les deux chapitres suivants se focalisent sur un gène (*rbp-1*) ayant des homologies significatives avec des RanBPM, et pouvant être lié à l'induction de la structure nourricière par les nématodes à kyste. Nous verrons dans ces chapitres comment nous avons isolé et décrit pour la première fois ce gène (chapitre 2), puis déterminé sa variabilité et identifié la famille multigénique auquel il appartient par le biais de l'exploitation d'une banque de transcrits de *G. pallida* (chapitre 3). Les résultats obtenus sur caractérisation du gène *rbp-1* ont été publiés dans la revue *Physiological and Molecular Plant Pathology* en 2005. Les premiers résultats acquis sur la caractérisation de la famille multigénique, la variabilité nucléotidique de celle-ci ainsi que sur l'étude de la variabilité du gène *rbp-1* au sein de différentes populations de *G. pallida* sont ici présentés sous la forme d'une short note qui pourrait être soumise à la revue *Molecular Plant Pathology* lorsque des données supplémentaires auront été obtenues.





## PARTIE II :

# CARACTERISATION DE NOUVEAUX GENES DU POUVOIR PATHOGENE

### INTRODUCTION

De nouveaux termes sont progressivement apparus pour tenter de regrouper des molécules en fonction de leur mode d'action (Greenbaum *et al.*, 2001). Ainsi, le parasitome regroupe toutes les protéines nécessaires à un organisme pour parasiter son hôte et le sécrétome définit toutes les molécules sécrétées. Tout au long de notre étude, nous avons ainsi défini comme gène du pouvoir pathogène tous les gènes qui permettent au nématode de parasiter la plante, de se développer jusqu'au stade adulte et ainsi de se multiplier. Ces gènes appartiennent au parasitome. Ils incluent ceux qui permettent au nématode de se protéger contre les défenses mises en place par la plante, de se nourrir du contenu cellulaire de la plante ou encore les gènes responsables des différentes étapes de pénétration, migration, mise en place et maintien du site nourricier. Parmi les gènes du parasitome, certains appartiennent également au sécrétome tels que les gènes dont les protéines sont localisées dans les sécrétions salivaires des nématodes à kyste. Nous nous sommes principalement focalisés sur ces types de gènes dont l'implication dans l'interaction est la plus probable. Tous ces gènes du pouvoir pathogène appartiennent également très souvent au "unknome" définit par Greenbaum *et al.* (2001) puisque la fonction de ces gènes est souvent inconnue.



# CHAPITRE 1

---

# CHAPITRE 1

## 1. ELEMENTS DE CONTEXTE : LE GENRE *GLOBODERA*

Les méthodes utilisées pour identifier les gènes du pouvoir pathogène sont diverses. Deux principales approches sont distinguées : les approches ciblées (exemple de l'AFLP à différents stades parasitaires, Qin *et al.*, 2001, ou de la stratégie gène candidat, Piotte *et al.*, 1999) ou les approches au hasard (exemple des ESTs : Etiquettes des séquences exprimées, Popeijus *et al.*, 2000; Dautova *et al.*, 2001). Dans ce dernier cas, la comparaison de nématodes génétiquement très proches mais présentant des différences de gamme d'hôtes ou de pathotypes doit permettre de mettre en exergue les gènes responsables de ces différences par des méthodes d'analyse génomique, transcriptomique ou protéique. Les nématodes à kyste du genre *Globodera* étant amphimictiques, une variabilité génétique relativement importante existe entre les individus d'une même population et *a fortiori* entre les populations et espèces. Néanmoins, deux espèces de *Globodera*, *G. pallida* et *G. "mexicana"* sont génétiquement proches, mais se distinguent par leur gamme d'hôtes. *G. pallida* se développe sur pomme de terre et pas sur morelle noire alors que *G. "mexicana"* a le comportement inverse (Bossis & Mugniéry, 1993; Thiéry *et al.*, 1997; Mugniéry *et al.*, 1992; voir tableau 1). La tomate est un hôte commun à ces deux espèces.

Tableau 1 : Gammes d'hôtes des différentes espèces de *Globodera*.

	Pomme de terre	Morelle noire	Tabac	Tomate
<i>G. tabacum</i>	-	+	+	+
<i>G. "mexicana"</i>	-	+	-	+
<i>G. pallida</i>	+	-	-	+
<i>G. rostochiensis</i>	+	-	-	+

Ces deux espèces peuvent se croiser et donner une descendance viable et féconde lorsque les croisements se font dans le sens : femelle *G. pallida* x mâle *G. "mexicana"*. Ces hybrides obtenus sur tomate sont incapables de se développer sur pomme de terre (Tableau 2). Cependant, une restauration rapide de la capacité à se développer sur cette plante est observée dès le premier croisement retour (ou backcross) avec *G. pallida*. L'exploitation de cette hérédité à se développer sur pomme de terre a été utilisée pour générer, à partir d'un croisement individuel *G. pallida* x *G. "mexicana"*, plusieurs lignées hybrides qui ont été sélectionnées pour leur aptitude à se développer

sur morelle noire et/ou sur pomme de terre (Tableau 3). Parmi les seize lignées pour lesquelles le pouvoir pathogène a été défini, six ont été utilisées dans notre étude.

L'incapacité de *G. "mexicana"* à parasiter la pomme de terre est liée à un défaut d'initiation du site nourricier. Les larves pénètrent les racines de pomme de terre mais aucune modification parasitaire n'est observée chez *Solanum tuberosum* en présence de *G. "mexicana"* qui n'induit pas de syncytium. Une hypothèse émise par Thiéry *et al.* (1997), est la non reconnaissance des sécrétions de *G. "mexicana"* par la plante hôte. La proximité génétique des deux espèces *G. pallida* et *G. "mexicana"* est apparue comme une originalité du modèle *Globodera* pouvant être très intéressante pour les études différentielles, notamment dans le but d'identifier de nouveaux facteurs génétiques impliqués dans le pouvoir pathogène voire plus particulièrement les différences de gamme d'hôtes observées. Certains éléments biologiques et moléculaires montrent que *G. "mexicana"* est bien une espèce à part entière même si la divergence génétique avec *G. pallida* est très faible (0,17) en comparaison avec les distance génétiques observées entre *G. pallida* et *G. rostochiensis* (0,50) ou entre *G. pallida* et *G. tabacum* (~ 0,6) (Bossis & Mugniéry, 1993). *G. "mexicana"* pourrait alors être plutôt considérée comme une sous espèces de *G. pallida* ayant perdu la capacité à se développer sur pomme de terre.

Tableau 2 : Capacité des hybrides issus de croisements entre *G. pallida* et *G. "mexicana"* à se développer sur tomate et sur pomme de terre.

% femelles formées		
Croisement	sur tomate	sur pomme de terre
GP	80	70
GM	90	0
GP x GM	71	0
<b>BC1 : GP x (GPxGM)</b>	70	<b>5</b>
<b>BC2 : GP x (GP x (GPxGM))</b>	45	<b>27</b>

GP : *G. pallida* ; GM : *G. "mexicana"*

Grenier *et al.* (2002) ont étudié les différences transcriptomiques entre *G. pallida* et *G. "mexicana"* par hybridation suppressive et soustractive. Trois transcrits partiels de la banque SSH n'ayant pas d'homologue identifié dans les bases de données et dont les analyses en reverse northern blot ont confirmé cette expression différentielle ont été conservés (*la7*, *lb3* et *IVg9*).

L'article du chapitre 1, rend compte des résultats d'étude d'expression des trois transcrits *IVg9*, *la7* et *lb3*, de la caractérisation de ces transcrits et des gènes correspondants ainsi que de leur variabilité au niveau moléculaire entre *G. pallida* et *G. "mexicana"*. Les trois questions posées étaient les suivantes :

- 1- Les transcrits différenciellement exprimés entre *G. pallida* et *G. "mexicana"* sont-ils impliqués dans le parasitisme ?

- 2- Si oui, les séquences des gènes sont elles variables entre les deux espèces ?  
 3- Si tel est le cas, certaines mutations peuvent-elles être reliées aux différences de gamme d'hôte observées ?

Tableau 3 : Caractérisation des lignées hybrides obtenues au laboratoire. Les hybrides de pouvoir pathogène simple ne se multiplient que sur *Solanum tuberosum*. Les doubles pathogènes se multiplient sur *S. tuberosum* mais également sur *Solanum nigrum* (morelle noire). Il existe quelques lignées qui ne se multiplient que faiblement sur *S. nigrum* (+/-). (Les lignées indiquées en rouge sont celles qui ont été utilisées pour les études du polymorphisme de gènes du parasitisme, voir article 2).

Lignée	Génération	Parents	Pouvoir pathogène	
			Morelle noire	Pomme de terre
<b>57</b>	<b>BC2</b>	<b>GP x GM (GM6)</b>	<b>+</b>	<b>+</b>
76	BC2	GP x GM (SA)	+	+
<b>80</b>	<b>BC2</b>	<b>GP x GM (SA)</b>	<b>+</b>	<b>+</b>
<b>86</b>	<b>BC2</b>	<b>GP x GM (SA)</b>	<b>-</b>	<b>+</b>
<b>88</b>	<b>BC2</b>	<b>GP x GM (SA)</b>	<b>-</b>	<b>+</b>
92	BC2	GP x GM (SA)	-	+
<b>96</b>	<b>BC2</b>	<b>GP x GM (SA)</b>	<b>+/-</b>	<b>+</b>
103	BC2	GP x GM (SA)	+	+
113	BC2	GP x GM (GM5)	+	+
124	BC1	GP x GM (GM6)	+	+
125	BC1	GP x GM (GM6)	+	+
126	BC1	GP x GM (SA)	+	+
128	BC2	GP x GM (Tlax)	+/-	+
129	BC1	GP x GM (GM5)	+	+
<b>131</b>	<b>BC2</b>	<b>GP x GM (GM6)</b>	<b>-</b>	<b>+</b>
132	BC2	GP x GM (SA)	+/-	+

GP : *G. pallida* Guiclan ; GM : *G. "mexicana"*. GM6, GM5, Tlax (Tlaxcala) et SA (Santa Ana) sont des populations de *G. "mexicana"*. BC : Backcross.

## 2. ARTICLE 1

**ENCADRÉ N° 1 : SEQUENCE POLYMORPHISM OF TWO GLOBODERA PIONEER GENES: STUDY OF THEIR PUTATIVE IMPLICATION IN PATHOGENICITY**

**IMPLICATION DES GENES *IA7* ET *IVG9* DANS LE PARASITISME ?**

**1. Localisation cellulaire**

- ***IVG9*** (localisation dans la glande dorsale) et ***IA7*** (localisation dans les glandes ventrales) sont des protéines probablement sécrétées *in planta* via le stylet.
- ***IB3*** (localisation intestinale) est écarté même si cette localisation ne peut être une preuve de non implication dans le parasitisme.

**2. Etude des séquences nucléotidiques**

- Identification d'un signal de sécrétion dans la séquences des transcrits *la7* et *IVg9*
- Fort taux de mutation au niveau génomique (8% pour *la7* et 17% pour *IVg9*) entre *G. pallida* et *G. "mexicana"*.

La localisation dans les glandes salivaires et la présence d'un signal peptide sont des indications fortes de la sécrétion des protéines *IA7* et *IVG9*. L'hypothèse d'un rôle de ces protéines dans l'interaction plante-nématode est renforcée par le fort taux de variabilité de ces deux gènes entre les deux espèces de nématode. Ces résultats suggèrent que les gènes *la7* et *IVg9* pourraient avoir accumulé des mutations à des fins adaptatives des espèces à leurs différents hôtes.

**SPECIFICITE D'HOTE ?**

Quel que soit le gène étudié, les séquences obtenues chez les hybrides (issus du croisement *G. pallida* X *G. "mexicana"* et ayant des gammes d'hôte différentes) sont identiques à celles du parent *G. pallida*. Par conséquent, aucune spécificité d'hôte n'a pu être révélée par l'étude du polymorphisme chez les lignées hybrides. Cependant, le fort taux de mutation conforte l'hypothèse d'une implication dans le pouvoir pathogène d'autant que ces gènes présentent un fort taux de mutation non synonyme suggérant qu'ils ont été soumis à une sélection forte générant de la diversité protéique au cours de l'évolution des nématodes.

**SEQUENCE POLYMORPHISM OF TWO PIONEER  
GENES IN PHYTOPARASITIC NEMATODES SHOWING DIFFERENT  
HOST RANGES**

Alexandra BLANCHARD, Didier FOUVILLE, Magali ESQUIBET, Didier MUGNIERY, Eric GRENIER.

Institut National de la Recherche Agronomique (INRA) - Agrocampus, unité mixte de recherche  
Biologie des organismes et des populations appliquée à la protection des plantes (UMR BiO3P),  
Domaine de la Motte au Vicomte, BP 35327, 35653 Le Rheu cedex France.

Tel: +33-2-23.48.51.91 Fax: +33-2-23.48.51.50

Corresponding author e-mail addresses: [alexandra.blanchard@rennes.inra.fr](mailto:alexandra.blanchard@rennes.inra.fr),  
[eric.grenier@rennes.inra.fr](mailto:eric.grenier@rennes.inra.fr)

*Soumis à Journal of Heredity*



## ABSTRACT

In order to identify genes involved in pathogenicity, we compared the closely related species *G. pallida* (GP) and *G. "mexicana"* (GM) that have different host ranges and are able to produce viable and fertile hybrids. Three pioneer genes were previously identified as differentially expressed between GP and GM: GPLIA7 and GPLIB3 were found more highly expressed in GP whereas GMLIVG9 was found over-expressed in GM. In this study, we showed that *la7* and *IVg9* genes probably encode products secreted by the nematode salivary glands, respectively the ventral oesophageal glands and the dorsal oesophageal gland. No Blast homolog was found in the databases, but a metridin-like ShK (*Stygodactyla helianthus*) toxin domain was identified in the *la7* sequence. Analysis of the full length sequences of these two genes between GP and GM revealed a high level of interspecies variability (8% for the *la7* transcript and 17% for the *IVg9* transcript) and a high proportion (90%) of non-synonymous mutations among the substitutions observed. This suggested that these two pioneer genes are under strong diversifying selection pressures and therefore may be involved in pathogenicity. Further investigations of the sequence polymorphism of *la7* and *IVg9* genes were conducted in GP x GM hybrid lines that were selected in laboratory conditions for their different ability to develop on potato and blacknight shade. As similar sequences were obtained for all the hybrid lines tested independently of their pathogenicity status, no correlation could be established between IA7 and IVG9 amino acids changes and the host range differences observed between GP and GM.

Keywords: parasitism genes, *Globodera pallida*, *Globodera "mexicana"*, potato cyst nematode, positive selection.

Accession numbers: DQ493453, DQ493454

## 1. INTRODUCTION

Plant parasitic nematodes such as cyst nematodes produce an assortment of parasitism proteins in order to infect plants (Davis *et al.*, 2004; Vanholme *et al.*, 2004). Both host range specificity and suppression of host plant resistance are thought to be controlled by some of these nematode parasitism genes. Progress has been made in recent years on the isolation of parasitism genes that may also constitute virulence genes (Semblat *et al.*, 2001, Bekal *et al.*, 2003, Neveu *et al.*, 2003), but few is known about which nematode parasitism genes are responsible for the success of the interaction. In plant parasitic nematodes no sequenced genome is yet available but several initiatives are underway. We can therefore expect that considerable progress in determining the genetic basis of the parasitism in nematodes will be made in the future as it was achieved on this question using a comparative genomic approach regarding other pathogens (Da Silva *et al.*, 2002, Paulsen *et al.*, 2002). However, these studies have also strengthened that apart some unique genes restricted to particular species, point mutations may also contribute to the differences in the host specificity observed. The role of point mutations that change the amino acids sequence of a protein should therefore be considered; for example a single amino acid change in *E. coli* FimH adhesion leads to the loss of activity and an altered host specificity for this bacteria (Pouttu *et al.*, 1999).

Cyst nematodes of agronomic interest in the genus *Globodera* infest only Solanaceous plants. Among them, *Globodera pallida* (GP) and *Globodera "mexicana"* (GM) - which was described in the unpublished Campos-Vela's thesis in 1967 - are two closely related species (Bossis and Mugniéry, 1993, Thiéry and Mugniéry, 1996) that display distinct host ranges. *G. pallida* is able to develop on potato (*Solanum tuberosum*) but not on blacknight shade (*Solanum nigrum*), whereas *G. "mexicana"* is able to develop on blacknight shade but not on potato. The tomato (*Lycopersicon esculentum*) is a common host for the two species. The incapacity of GM to develop on *S. tuberosum* resides in its inability to induce a feeding site, even after penetrating the roots (Thiéry *et al.*, 1997). Interestingly, GP and GM are able to produce viable and fertile hybrids (Mugniéry *et al.*, 1992). The progeny obtained in controlled conditions lose their ability to develop on potato, but this ability can be quickly restored after one or two backcrosses with the GP parent (Thiéry *et al.*, 1997). Depending on the plant host used to rear these hybrids, we were able to select in the laboratory inbred populations that are able to develop on both potato and blacknight shade or only on potato.

Grenier *et al.* (2002) investigate the transcriptome differences between *G. pallida* and *G. "mexicana"* by SSH and identified several specific sequences of potato cyst nematode. Three of them, GPLIA7, GPLIB3 and GMLIVG9, were pioneer sequences without homolog in databases and were then classified as potato cyst nematode specific genes. The GPLIA7 and GPLIB3 transcripts were detected as over-expressed in GP whereas GMLIVG9 transcript was over-expressed in GM. These results were confirmed in reverse Northern blots. Establishing the function of the proteins encoded by these pioneer genes is a difficult task. It is currently admitted that the genes expressed in nematode salivary glands have a role in parasitism as they correspond to secreted products that can interact with plant proteins (Williamson *et al.*, 2003; Davis *et al.*, 2004).

Thus, in this study which is complementary to the quantitative SSH approach performed on the GP and GM transcriptomes, we carried out a qualitative investigation of the three pioneer genes, *GPLIA7*, *GPLIB3* and *GMLIVG9*. In a first step, experiments including *in situ* hybridisation and analysis of the full length cDNA sequence were conducted to assign a function to these sequences. In a second step, the genes structure and sequence variability among GP, GM and the hybrid lines selected in the laboratory were compared. Putative implications of the mutations observed in host range specificity were investigated according to these data.

## 2. MATERIAL & METHODS

### 2.1. Biological material

The *G. pallida* population Guiclan and the *G. "mexicana"* population Santa Ana used in this study were respectively reared on *Solanum tuberosum* and *Solanum nigrum*. Juveniles (J2s) were obtained by soaking cysts first in water for 4 days at 4°C and then in potato root diffusate at room temperature. The J2s were stored in water at 4°C for *in situ* hybridisation or at -80°C for nucleic acid extraction.

As sexe determination for these nematodes is epigenetic, females and males were independently produced. Hybrid lines were obtained in controlled conditions following the selection schema presented in figure 1. A first cross was generated with GP female and GM male to produce hybrid lines of first generation (G1). Then successive backcrosses using GP females and several hybrid males were performed to generate hybrid lines of generations 2 to 3 on tomato. For each cross, females were produced after *in vitro* inoculation of one juvenile in one root of tomato, and then one male of GM or G1 or G2 was added to each plate containing a single female. At the end, half of the progeny obtained in G3 was multiplied on potato, the other half on blacknight shade. Some hybrid lines were obtained on potato but never on blacknight shade. Each potato line generated was tested for its ability to develop on *S. nigrum*.

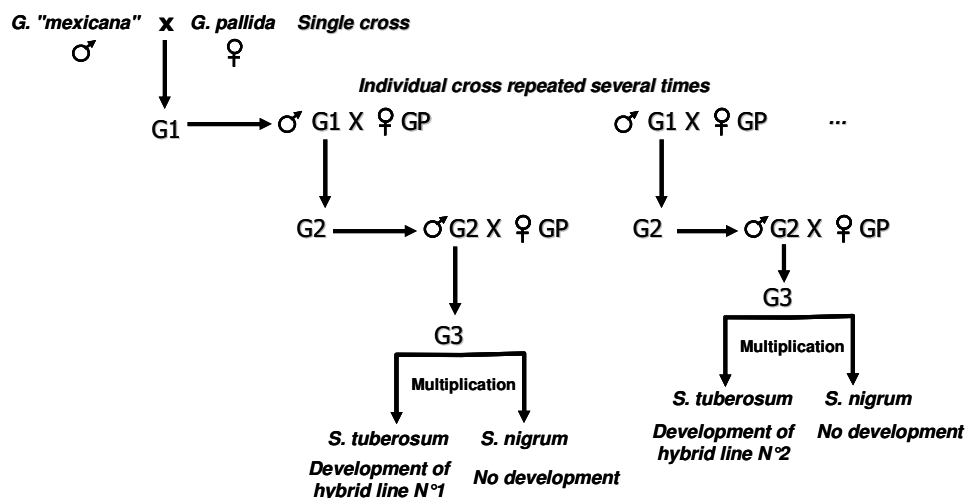


Figure 1: Hybrid lines selection from crossings between *Globodera pallida* (GP) and *Globodera "mexicana"* (GM) species. G1 to G3: Generation 1 to 3.

## 2.2. mRNA *in situ* hybridisation

The clones GPLIA7, GPLIB3 and GMLIVG9 were used to design specific primers (Table 1) to generate *in situ* hybridisation probes. The first step was to amplify the insert of each clone before dig-labelling each probe (sense and anti-sense) by asymmetric PCR procedure. Insert amplification was performed in 50µl with 30ng of insert, 0.4µM of each primer, 1X buffer, 0.2mM of each dNTP, 2U of Taq DNA polymerase (Promega) and 1.5mM MgCl<sub>2</sub>, for 4 min at 94°C, 40 cycles at 94°C for 30s, Tm for 30s, 72 for 1min, and a final elongation step of 4 min at 72°C. Asymmetric PCR was performed in 40µl with 2µl of the previous purified PCR product, 2.5µM of a primer (Fwd or Rev), 1X buffer, 1.5 mM MgCl<sub>2</sub>, 0.75 X Dig-dUTP/dNTP mix (Roche Applied Science) and 2U of Taq DNA polymerase (Promega) for 4 min at 94°C, 34 cycles at 94°C for 30s, Tm for 1min, 72°C for 90s, and a final elongation step of 4 min at 72°C. *In situ* hybridisation was performed as described by De Boer *et al.* (1998) with slight differences. Freshly hatched J2s were fixed in fixation buffer (including paraformaldehyde) for 18h at 4°C and 4h at room temperature. The nematodes were cut with a razor blade and the sections were permeabilised with proteinase K for 30 min at 22°C before freezing at -80°C. Hybridisations were performed at 45°C over night with purified sense or anti-sense single strand cDNA probes. The signal was detected using alkaline phosphatase immunostaining (NBT-BCIP, Boehringer).

Table 1: Characteristics of the SSH clones identified by Grenier *et al.*, 2002, and primers used in this study for each clone.

SSH clone	Insert size	PolyA tail	Over expression	Primers names	Sequences (5'→ 3')
GPLIA7	201 bp	Yes	<i>G. pallida</i>	IA7 Fwd2	CGTCGGCCTCTAGCTGTA
				IA7 Rev	TGCCTCTTTATCTTTGAAAATG
				5'IA7	AGCCGTCTGAGTCCAATG
				3'IA7	TTGATTTCTGTCTTCAGCAAAA
				GPLIA7Rev2	GTTTTTGGGCACATGGTTG
GPLIB3	343 bp	No	<i>G. pallida</i>	IB3 Fwd	TGATATGCCGACAGCTCAAC
				IB3 Rev	ACCGACCACACCATCATCTT
GMLIVG9	419 bp	Yes	<i>G. "mexicana"</i>	IVG9 Fwd	AAACAACGTTTGCCGAATG
				IVG9 Rev	GCGCAACATTCAACGACA
				5'IVG9.2	ATGACAAAAATTGTGTTTTTCCT
				3'IVG9.2	CGAATGTACACATATCACCAAT

## **2.3. Isolation of the full length of the Ia7 and IVg9 transcripts**

### **A. RNA isolation**

Total RNA was extracted from 50 to 100 mg of J2s with the RNeasy Mini kit (Sigma) according to the manufacturer's instructions. The RNAs were treated with DNase I (10 U, Ambion) for 15 min at 37°C and then purified by phenol/chloroform/isoamyl alcohol (25/24/1) extraction. The RNAs were precipitated with 3M sodium acetate and absolute ethanol, and stored at -80°C.

### **B. mRNA reverse transcription**

Reverse transcription was performed in a mix containing 2µg of total RNA, 0.25µM of d(T)<sub>25</sub> primer and 8.5µl of RNase-free water. The samples were heated for 10 min at 70°C and set on ice. 1X superscript buffer, 2.5mM MgCl<sub>2</sub>, 0.25µM of each dNTP and 0.01µM of DTT were then added. The sample was heated for 1 min at 42°C. Superscript III (Invitrogen, 200U/µl) was added and the mix was incubated at 42°C for 50 min, and at 70°C for 15 min. One microlitre of RNase mix (Boehringer) was added and the mixture was incubated for 30 min at 37°C.

### **C. 5' RACE**

Transcripts 5' ends were isolated using the 5' RACE System for Rapid Amplification of cDNA Ends (Invitrogen) according to the manufacturer's instructions. The first strand cDNA was obtained from total RNAs using 0.4µM of specific primer GPLIA7rev2 or IVG9fwd (Table 1). Nested PCR was performed in 50µl with 1X buffer, 1.5mM of MgCl<sub>2</sub>, 0.2mM of each dNTP, 2.5 units of Taq DNA polymerase (Promega), 5µl of cDNA, 0.4µM of AAP (5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3'), and 0.4µM of the internal specific primer (GPLIA7rev2 or IVG9fwd; Table 1). PCR conditions were 96°C for 1min, 35 cycles at 96°C for 20s, 53°C for 30s and 72°C for 30s, with a final elongation step at 72°C for 5 min.

The PCR products were purified (GeneElute PCR Clean-Up kit, Sigma) and cloned in pGEM-T Easy Vector System I (Promega) according to the manufacturer's instructions.

## **2.4. Characterisation of Ia7 and IVg9 genomic sequences**

### **A. DNA isolation**

Genomic DNA was extracted from the eggs (separated from the cyst) of 500 rehydrated cysts. The frozen eggs were crushed in eppendorf tubes with a piston. The DNA was extracted with lysis buffer (0.1M Tris, pH 8, 10mM EDTA, 2% SDS) and proteinase K, and then incubated at 65°C for 1 hr. The DNA was centrifuged and purified with 140µl of 5M NaCl and 64µl of 10% CTAB, incubated for 10 min at 65°C and then with a phenol/chloroform (1:1) mix. The DNA was precipitated with ammonium acetate 5M overnight at 4°C. A mix containing 30% of PEG 6000 and isopropanol (99.8%) was added to achieve DNA precipitation. The sample was treated with 1µl of RNase A (500µg/mL) for 2 hr at 37°C.

### B. *la7* and *IVg9* gene amplifications

The 5'IA7/3'IA7 and 5'IVG9.2/3'IVG9.2 (Table 1) primers designed from the *la7* and *IVg9* cDNA sequences were used to respectively amplify the *la7* and *IVg9* genes in *G. pallida*, *G. "mexicana"* and in the hybrid populations.

The *la7* gene was amplified by PCR using 50 ng of DNA, 1X buffer, 1.5mM of MgCl<sub>2</sub>, 0.2mM of each dNTP, 0.4μM of each primer, and 1.25 U of GoTaq flexi (Promega). PCR was performed following the program: 96°C for 1min, 35 cycles at 94°C for 20s, 59°C for 5s, 57°C for 5s, 55°C for 10s and 68°C for 3 min, with a final elongation step at 68°C for 10 min. The *IVg9* gene was amplified in the same conditions except for the MgCl<sub>2</sub> concentration which was 3 mM.

All the PCR products were purified using Sephadex G50 and then directly sequenced by MacroGen (Korea, <http://www.macrogen.com>).

### C. Sequence analysis

WU-Blast2 Parasite Genomes Database Query on EMBL-EBI ([www.ebi.ac.uk/blast2/parasites.html](http://www.ebi.ac.uk/blast2/parasites.html)), and Blast on NCBI ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)) or on nemaBLAST (<http://www.nematode.net/>) were used to search for sequence homologies. The ExPasy translate tool (<http://us.expasy.org/tools/dna.html>) was used to predict amino acid sequences. Sequence alignments were performed with Multalin algorithm (Corpet, 1988; <http://prodes.toulouse.inra.fr/multalin/multalin.html>). The sequence signatures were detected using SignalP (Bendtsen *et al.*, 2004; [www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), and InterProScan (Quevillon *et al.*, 2005; <http://www.ebi.ac.uk/InterProScan/>) and the properties of the proteins were identified using ProtParam (Gasteiger *et al.*, 2005; <http://us.expasy.org/tools/protparam.html>).

## 3. RESULTS

### 3.1. Spatial expression of transcripts *GPLIA7*, *GPLIB3* and *GMLIVG9*

Grenier *et al.* (2002) identified *GPLIA7*, *GPLIB3* and *GMLIVG9* as differentially expressed by SSH (Table 1). The spatial expression patterns of these three transcripts were studied using *in situ* hybridisation. Approximately 70 to 80 % of the nematodes were labelled in each experiment. One of them, *GPLIB3*, was located in the intestine (Fig. 2A) of the GP nematodes using the IB3Fwd digoxigenin-labelled probe (see primers in Table 1).

The *GPLIA7* transcript was located in oesophageal glands of the GP nematodes (Fig. 2B) using the IA7Rev digoxigenin-labelled probe. After 3 days of incubation the labelling extended to the median bulb of the J2s (Fig. 2C) where the valve of the ventral glands is connected to the lumen of the oesophagus, clearly indicating a localisation in the ventral oesophageal glands.

Using the IVG9Fwd digoxigenin-labelled probe an accumulation of *GMLIVG9* transcripts was clearly observed in the dorsal gland of the GM nematodes (Fig. 2D).

The three complementary probes generated using IB3Rev, IA7Fwd2 and IVG9Rev primers were used as negative control (data not shown) and produced no signal. Following these

observations, we decided to focus further investigations on only the GPLIA7 and GMLIVG9 transcripts that showed a localisation in the nematode salivary glands and therefore correspond to secreted products that are probably implicated in the parasitism process.

### 3.2. Isolation of full-length *la7* and *IVg9* cDNA

As a polyA tail was observed in both GPLIA7 and GMLIVG9 inserts, specific primers (Table 1) were designed in order to obtain the full length cDNAs using 5' RACE PCR technique. The GPLIA7Rev2 and IVG9Fwd primers (Table 1) were used to obtain the 5' end of respectively the *la7* and *IVg9* transcripts in respectively GP and GM. The PCR products were cloned and sequenced to build the full length transcripts *in silico* using the SSH clone insert sequences.

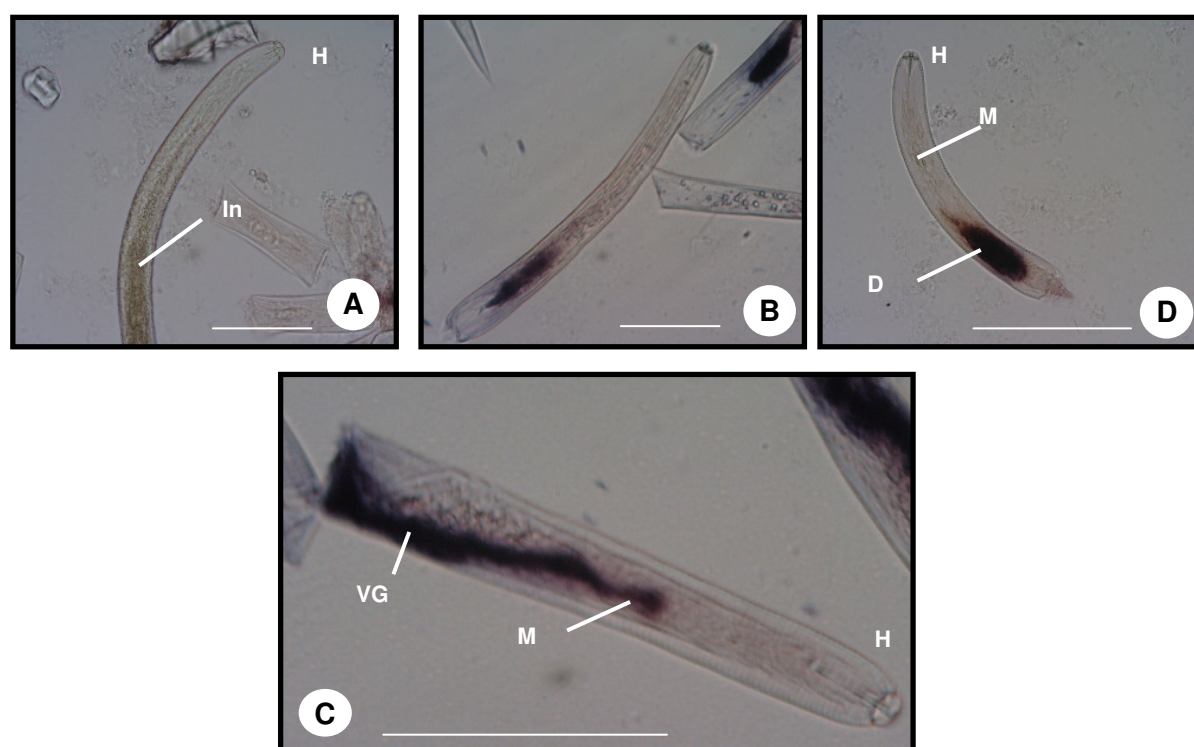


Figure 2: Localisation by *in situ* hybridisation of the *Ib3* (A), *la7* (B and C) and *Ivg9* (D) transcripts. A: *Ib3* transcripts localisation in the intestine of the *G. pallida* J2s using the IB3Fwd digoxigenin-labelled probes. B and C: ventral glands localisation of the *la7* transcripts in *G. pallida* infective juveniles using anti-sense IA7Rev digoxigenin-labelling probes. A longer incubation of the sections with the phosphatase alkaline enzyme showed a signal up to the metacarpus of J2s (C). D: dorsal oesophageal gland localisation of the *Ivg9* transcripts of the *G. "mexicana"* infective juvenile using the sense (IVG9 Fwd) digoxigenin-labelled probe. H, head; M, metacarpus; DG, dorsal gland; VG; ventral gland; In, intestine; Bar, 70µm.

The full length *la7* cDNA was then estimated to be 389 bp in length for *G. pallida* with an identified open reading frame (ORF) of 222 bp. Two primers (5'IA7 and 3'IA7, Table 1) flanking the ORF were designed to check the *in silico* construction. A fragment of approximately 250 bp was amplified confirming the construction. The *la7* cDNA encoded a deduced protein of 73 aa (Fig. 3A) with a calculated molecular weight of 8 kD. A peptide signal of 25 aa was predicted by SignalP and a metridin-like ShK toxin domain was identified by the InterProScan program from the amino acids 39 to 73 (Fig. 3A).

The full length *IVg9* cDNA built *in silico* represented in GM a product of 471 bp with an ORF of 285 bp encoding a putative 95 aa protein with a molecular weight of 11kD and a peptide signal of 24 aa (Fig. 3B). In order to check the *in silico* construction, the 5'IVG9.2 and the 3'IVG9.2 primers were designed (Table 1). A PCR product of more than 300 bp was obtained in agreement with predicted sequence. No specific domains were identified using InterProScan program.

BLAST searches revealed no significant homology (E value threshold  $10^{-4}$ ) in databases for both full length transcripts.

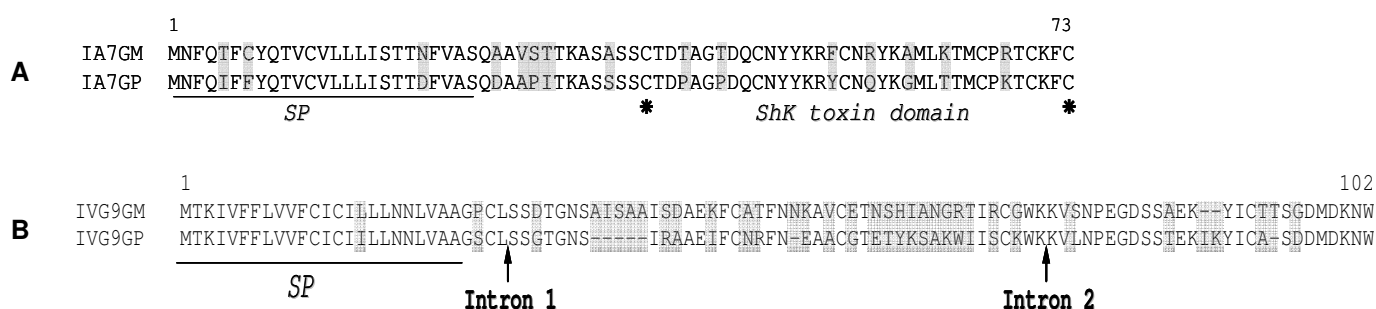


Fig. 3: Alignment of the predicted proteins of *la7* (A) and *IVg9* (B) genes in *Globodera pallida* (GP) and *G. "mexicana"* (GM). Variable sites were highlighted in grey boxes. The *la7* ShK toxin domain is delimited with asterisks and the signal peptides (SP) are underlined for both genes. The intron positions of the *IVg9* gene are indicated using arrows.

### 3.3. Sequence variability of *la7* and *IVg9* genes in *G. pallida* and *G. "mexicana"*

Amplifications of the *la7* and *IVg9* genes were obtained from GP and GM genomic DNAs. For *la7*, a band of about 250 bp was observed on agarose gel from both GP and GM suggesting that this gene was intronless. The high quality of the sequences enabled to analyse them without a cloning step. Sequences obtained for GP and GM were of the same size. Out of the 222 bp of the ORF, 18 substitutions (8% of sites were variable) were observed between GP and GM (Fig. 4A) resulting in 15 variable aa in the predicted protein sequence between the two species (89% of non-synonymous substitutions) (Fig. 3A). The mutations are equally dispatched along the amino acid sequence with only one accumulation of 3 non-synonymous substitutions from the amino acid 29 to 31.



The *IVg9* gene was found to be 830 bp long for GP and 712 bp long for GM. The ORFs were respectively of 285 bp for GP and 300 bp for GM, encoding respectively predicted proteins of 95 and 100 aa (Fig. 3B). Comparing GP and GM sequences, 50 variable sites were identified (17.5 % of sites were variable) resulting in 26 altered amino acid positions in the predicted protein sequences. Similarly to *la7*, 90% of the substitutions observed are non-synonymous substitutions. Three insertion-deletions were observed between the GM and GP *IVg9* sequences (Fig. 4B): two insertions of 15 and 3 nucleotides were observed in the GM *IVg9* sequence and one deletion of 3 nucleotides was observed in the C-terminal part of the GM gene. Compared to the transcript sequence, this gene appeared to contain two introns. The introns had the same distribution along the *IVg9* gene sequence from GP and GM (Fig. 4B) indicating that these are orthologs. The first intron was 102 bp in length for both species and displayed 4 % of variability. The second was 428 bp in length for GP and 295 bp in length for GM and showed 8.5% of variability and 6 indels (4 indels of only one nucleotide).

### 3.4. Various pathogenicities of the hybrid lines

We obtained 9 G3 hybrid lines all derived from the same individual cross between GM males and GP females. All these hybrid lines were obtained from nematodes that were reared and multiplied on potato. Hybrid lines generated on black night shade were all lost during the multiplication process. Hybrid lines derived from potato G3 displayed various pathogenicities: 3 were unable to develop on blacknight shade and were called simple pathogens (HLs), the others were able to develop on *S. nigrum* but with various levels (Table 2).

Considering the level of development of GM on *S. nigrum* during this test (30%), and the fact that the hybrid lines only contained 20% of GM AFLP markers (Grenier, pers. com.). Then the hybrid lines development will at best reach 6% on blacknight shade. Among the hybrid lines, we observed various pathogenicities represented by the percentage of developed females that ranged from 0.7% to 4.4% (Table 2). Then we chose two lines that represent the extremities: the line that showed development at the inferior limit was called LHi and the one that showed the superior limit was called LHd. Finally, only three of these lines (80, 86, 96) were used in the following investigations.

Table 2: Parental and hybrid lines pathogenicity on *Solanum nigrum*. GP: *G. pallida*; GM: *G. "mexicana"*; HL: hybrid line; HLs: hybrid lines that only develop on *S. tuberosum*; HLd: hybrid lines that develop both on *S. tuberosum* and *S. nigrum* with the higher percentage of females; HLi: hybrid lines that develop both on *S. tuberosum* and *S. nigrum* with the lowest percentage of females.

Parental lines		Pathogenicity	% of female on <i>S. nigrum</i>
GP		-	0
GM		+	30
Backcross	Hybrid line	Pathogenicity	% of female on <i>S. nigrum</i>
BC2	76	HLd	4
BC2	80	HLd	3.5
BC2	86	HLs	0
BC2	88	HLs	0
BC2	92	HLs	0
BC2	96	HLi	0.7
BC2	103	HLd	3
BC2	126	HLd	4.4
BC2	132	HLi	0.7

### 3.5. Characterisation of *la7* and *IVg9* genes in hybrid lines

As the sequence variations observed between GP and GM for both *la7* and *IVg9* genes strongly affected the amino acid sequences and therefore putatively the function of these genes in GM and GP, we decided to investigate the sequence variability of these two genes in the hybrid lines derived from these species. The *la7* and *IVg9* genes were amplified in 3 different hybrids lines (HL) originating from the same initial cross between a GM male and a GP female using the same primers as previously. For the *la7* gene, a PCR product of about 250 bp was amplified (Fig. 5A). Sequences obtained for the 3 HL were also of high quality enabling to analyse them without cloning previously step. These sequences were the same and strictly correspond to the GP sequence (Fig. 4A). For the *IVg9* gene, a PCR product of about 850 bp was amplified for the 3 hybrid lines (Fig. 5C). In the coding region of this gene, no variation was found among the hybrid lines. The *IVg9* sequences of the HL correspond to the GP sequence except for 2 nucleotides at the position 267 and 831 (Fig. 4B) that appeared to be synonymous substitutions in the deduced proteins.

Consequently, we can consider that for both genes the hybrid sequences were mainly inherited from de GP parent and appeared independent of the ability for the hybrids to develop on potato or on potato and blacknight shade.

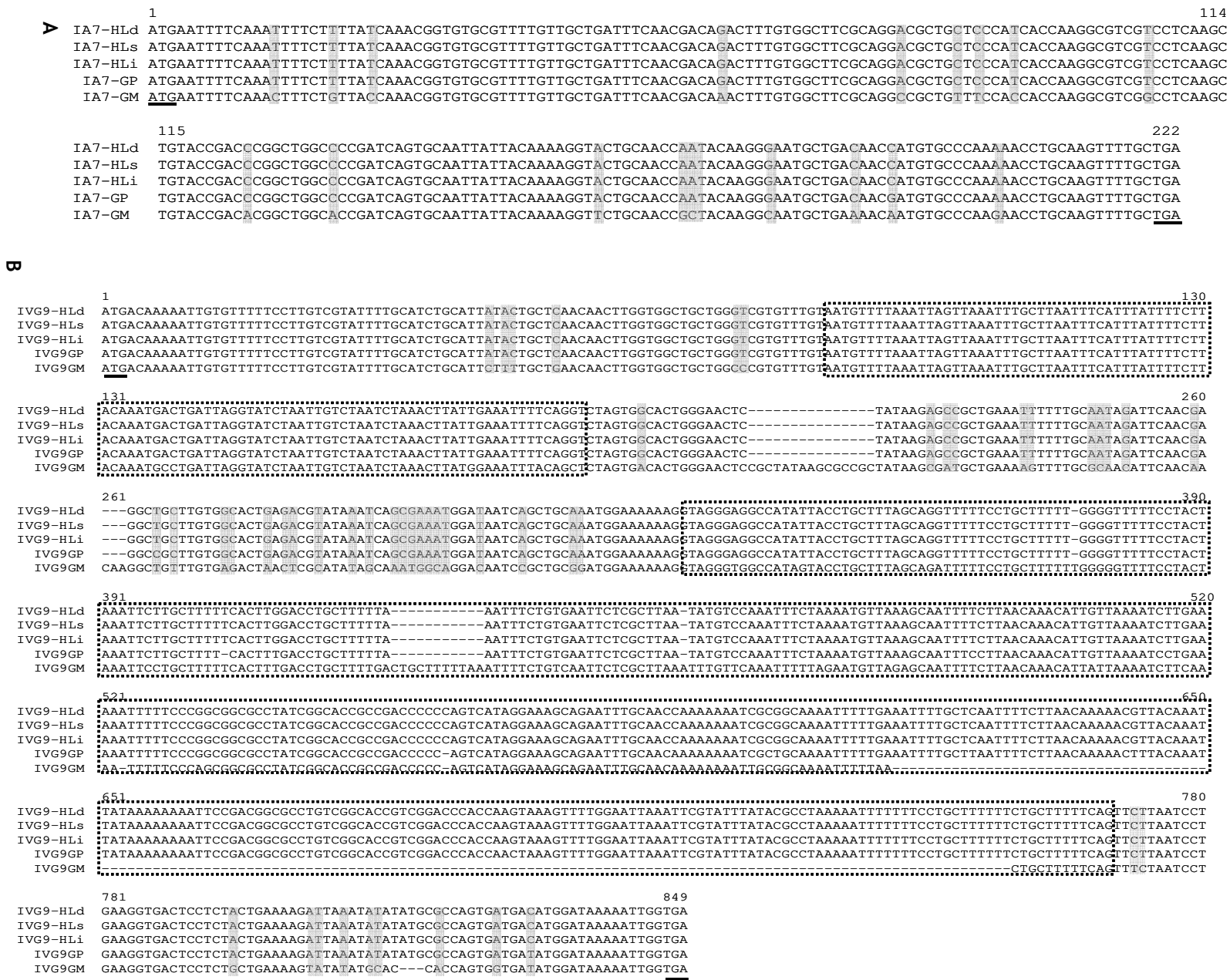


Fig. 4: Alignment of the *Ia7* and *Iv9* genes sequences for the GP, GM and hybrid lines populations. Variables sites are indicated in grey boxes. Introns were indicated in dot open boxes. Variable sites in coding regions were represented in grey boxes; start and stop codons were underlined. GP: *G. pallida*, GM: *G. "mexicana"*; HLd: hybrid line developing on *S. tuberosum* and *S. nigrum*; HLi: hybrid line that only develops on *S. tuberosum*; HLi: hybrid line that develops on *S. tuberosum* and weakly on *S. nigrum*.

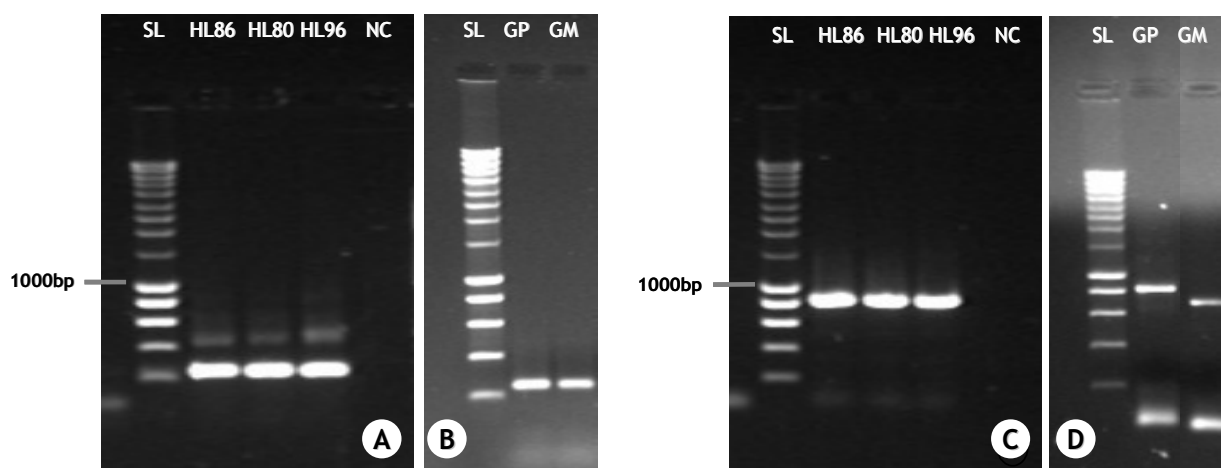


Fig. 5: PCR amplification of the *la7* (A, B) and *IVg9* (C,D) genes in genomic DNA of *G. pallida*, *G. "mexicana"* and hybrid lines obtained after two backcrosses. HLs: hybrid lines that only develop on *S. tuberosum*; HLd: hybrid lines that develop both on *S. tuberosum* and *S. nigrum* with the higher percentage of females; HLi: hybrid lines that develop both on *S. tuberosum* and *S. nigrum* with the lowest percentage of females, GP : *G. pallida*, GM: *G. "mexicana"*. NC: negative control; SL: Smart Ladder (Eurogentec).

#### 4. DISCUSSION

The study of the genetic differences between two closely related species of *Globodera* that have different host-ranges is an opportunity to reveal potential pathogenicity factors that act specifically against an host during the plant invasion by the nematode. Two genes were described in this study as potential pathogenicity factors of *Globodera* nematodes. These genes encode putative secreted proteins. However no significant matches in databases to homologs of these genes were found. So their function in host specificity remained unknown.

The *in situ* hybridisation experiments localised *la7* and *IVg9* to oesophageal glands and *lb3* to the intestine. This suggests that the *IA7* and *IVG9* proteins are secreted during the plant invasion and thus play a role in the plant-nematode interaction whereas the role of the *IB3* protein in the parasitism remains more uncertain. The accumulation of *la7* transcripts was clearly observed in the ventral glands and not in the dorsal gland. In longer exposure times, the transcripts were also localised in the valve of the glands which is connected to oesophagus in the median bulb of the J2s. In contrast, the *IVg9* transcripts were localised in the dorsal gland of the nematodes. The accumulation of those transcripts in the two different types of salivary gland could be an indication that these two genes act at different stages of parasitism as Von Mende (1997) reported that ventral glands are metabolically active earlier than the dorsal one. The sequences of the *la7* and *IVg9* transcripts indicated the presence of a secretion signal in the corresponding predicted proteins as a supplementary indication that these two proteins are secreted. Moreover the small sizes (MW: 8kD

for IA7 and 11kD) of the predicted proteins are compatible with the exclusion size of the lumen of the nematode stylet.

Database searches with the full length sequences of *la7* and *IVg9* did not reveal significant homologies with any sequences. The function of these genes remained unknown, although the bioinformatics analyses did identify specific motifs in *la7* sequence. A metridin-like ShK toxin (*Stichodactyla helianthus* K channel toxin) domain was identified in the IA7 predicted protein whose function is unknown but found in some toxins (Castadena *et al.*, 1995; Tudor *et al.*, 1998). This domain is also found in several *Caenorhabditis elegans* proteins without any function identified. Thus, considering that IA7 could be related to a toxin, we could speculate that it could have similar role with some fungi compounds implicated in detoxification of antimicrobial compounds produced by plants or by interfering with the induced defence response to suppress the fundamental resistance process (George *et al.*, 2001; Bouarab *et al.*, 2002).

The orthology of the GP and GM *IVg9* sequences was confirmed by the findings of same introns/exons boundaries and a low variability of the introns observed between GP and GM sequences (4 to 8%). Indeed, numerous studies sustained the high conservation of the intron/exon structure of orthologous genes as well at genus level (Rokas *et al.*, 1999; Wada *et al.*, 2002), as at the phyla level (Rogozin *et al.*, 2003), even if in some cases this rule is not exact (Goetze *et al.*, 2006). Considering the orthology of the GP and GM *la7* sequences, the question remains as the IA7 gene is intron less. However, as the variability of the *la7* coding region (8%) is less than half of that observed for *IVg9* (17%), we assume that we also compared orthologs in the case of *la7*.

The variability of the *la7* and *IVg9* genes observed between GP and GM reached to approximately 8% for *la7* and 18% for *IVg9*. This inter species variability appeared very high compared to other parasitism genes such as a cellulase or a RanBPM genes that showed 1 to 2% of variability between the same species (Grenier *et al.* 2002; Blanchard *et al.*, 2005). As a high proportion of the mutations observed correspond to non synonymous substitutions (nearly 90%), we supposed that the *la7* and *IVg9* genes have both evolved under strong selective pressure. Some recent studies have reported that genes encoded surface proteins of parasitic bacteria that were directly in contact with the host were under positive selection (Tsai *et al.*, 2006, Sawires *et al.*, 2006). Moreover, it was also shown that the hypervariability of some *hrp* genes (hypersensitive response and pathogenicity genes) involved in pathogenicity could be found at the intra specific level for the plant parasitic bacteria of various *Xanthomonas campestris* pathovars (Weber *et al.*, 2005) suggesting a potential role in host specificity. Genes under positive selection are the key of the adaptation to a habitat or niche (Chen *et al.*, 2006). We can therefore hypothesize that this could also occur for hosts and that nematode parasitism genes encoding proteins implicated in plant nematode interactions, will be under positive selection to facilitate the adaptation of the species on different host plants.

In this study, several results indicated that the pioneer genes identified could be pathogenicity genes: differential expression between GP and GM, their putative secretion, the absence of homologs in the extensive *C. elegans* databases, the high sequence variability and the high rate of non synonymous substitutions. In an attempt to further characterise the function of these genes and the impact of the mutations observed, we used GP x GM hybrid lines with various

pathogenicities, and postulated that the variability of the *la7* and *IVg9* genes in these hybrid lines can be correlated to their host specificity. The *la7* and *IVg9* sequences obtained in all the hybrid lines tested were mainly the same as that of GP that was the recurrent parent during the crossing experiment. Some PCR performed in other hybrid lines obtained in G2 instead G3 (data not shown) displayed the same results. Then, as no sequence difference was observed among hybrid lines tested, we were unable to show any implication of the mutations observed in these two genes in host range specificity. However as none of the hybrid lines were unable to develop on potato then we cannot yet exclude that the mutations observed in *la7* or *IVg9* GM genes preclude the development of GM on potato or that the mutations observed in *la7* and *IVg9* GP genes allow the nematodes to develop on potato. Clearly, a new crossing experiment should be now designed to obtain new hybrid lines representing all the expected various pathogenicities in blacknight shade in order to answer these last questions.

We described an original method to study the parasitism genes through the evaluation of their polymorphism in populations and lines that differentially develop on various plants. We strongly support this kind of investigation using different putative parasitism genes to test the hypothesis of a host specificity implication. These data could be complementary to functional analyses of corresponding proteins.

## REFERENCES

- Bekal S, Niblack TL, Lambert KN. 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Mol Plant Microbe Interact.* 16:439-446.
- Bendtsen JD, Nielsen H, Von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.* 340:783-795.
- Blanchard A, Esquibet M, Fouville D, Grenier E. 2005. Ranbpm homologue genes characterized in the cyst nematodes *Globodera pallida* and *Globodera 'mexicana'*. *Physiol Mol Plant Pathol.* 67:15-22.
- Bossis M, Mugniéry D. 1993. Specific status of six *Globodera* parasites of Solanaceous plants studied by means of two-dimensional gel electrophoresis with a comparison of gel patterns by a computed system. *Fundam Appl Nematol.* 16:47-56.
- Bouarab K, Melton R, Peart J, Baulcombe D, Osbourn A. 2002. A saponin-detoxification enzyme mediates suppression of plant defences. *Nature.* 418:889-892.
- Campos-Vela A. 1967. Taxonomy, life cycle and host range of *Heterodera mexicana*. Sp. (Nematoda: Heteroderidae), University of Wisconsin : 70p.
- Castadena O, Sotolongo V, Amor AM, Stöcklin R, Anderson AJ, Harvey AL, Engström A, Wernstedt C, Karlson E. 1995. Characterization of a potassium channel toxin from the caribbean sea anemone *Stichodactyla helianthus*. *Toxicon.* 33:603-613.
- Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D, Bieri T, Meyer RR, Ozersky P, Armstrong JR, Fulton RS, Latreille JP, Spieth J, Hooton TM, Mardis ER, Hultgren SJ, Gordon J I. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. *Proc Natl Acad Sci.* 103:5977-5982.
- Corpet F. 1998. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16:10881-10890.
- Da Silva ACR, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello CB, Sluys MAV, Almeida NF, Alves LMC, do Amaral AM, Bertolini MC, Camargo LEA, Camarotte G, Cannavan F, Cardozo J, Chambergo F, Ciapina LP, Cicarelli RMB, Coutinho LL, Cursino-Santos J R, El Dorry H, Faria JB, Ferreira AJS, Ferreira RCC, Ferro MIT, Formighieri EF, Franco MC, Greggio CC, Gruber A, Katsuyama AM, Kishi LT, Leite RP, Lemos EGM, Lemos MVF, Locali EC, Machado MA, Madeira AMBN, Martinez-Rossi NM, Martins EC, Meidanis J, Menck CFM, Miyaki C Y, Moon DH, Moreira LM, Novo MTM, Okura VK, Oliveira MC, Oliveira VR, Pereira HA, Rossi A, Sena JAD, Silva C, de Souza RF, Spinola LAF, Takita MA, Tamura RE, Teixeira EC, Tezza RID, Trindade dos Santos M, Truffi D, Tsai SM, White FF, Setubal JC, Kitajima JP. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature.* 417:459-463.

Davis EL, Hussey RS, Baum TJ. 2004. Getting to the roots of parasitism by nematodes. Trends Parasitol. 20:134-141.

De Boer JM, Yan YT, Smant G, Davis EL, Baum TJ. 1998. In-situ hybridization to messenger RNA in *Heterodera glycines*. J Nematol. 30:309-312.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy Server. In: John M. Walker, editors. The proteomics protocols handbook. Humana Press. p. 571-607.

George HL, VanEtten HD. 2001. Characterization of pisatin-inducible cytochrome p450s in fungal pathogens of pea that detoxify the pea phytoalexin pisatin. Fungal Genet Biol. 33: 37-48.

Goetze E. 2006. Elongation factor 1-alpha in marine copepods (Calanoida: Eucalanidae): phylogenetic utility and unique intron structure. Mol Phylogenet Evol.40:880-886.

Grenier E, Blok VC, Jones JT, Fouville D, Mugniéry D. 2002. Identification of gene expression differences between *Globodera pallida* and *G. mexicana* by suppression subtractive hybridization. Mol Plant Pathol. 3:217-226.

Mugniéry D, Bossis M, Pierre JS. 1992. Hybridations entre *Globodera rostochiensis* (Wollenweber), *G. pallida* (Stone), *G. virginiae* (Miller & Gray), *G. solanacearum* (Miller & Gray) et *G. mexicana* (Campos-Vela). Description et devenir des hybrides. Fundam Appl Nematol. 15:375-382.

Neveu C, Abad P, Castagnone-Sereno P. 2003. Molecular cloning and characterization of an intestinal cathepsin L protease from the plant-parasitic nematode *Meloidogyne incognita*. Physiol Mol Plant Pathol. 63:159-165.

Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, Read TD, Dodson RJ, Umayam L, Brinkac LM, Beanan MJ, Daugherty SC, Deboy RT, Durkin AS, Kolonay JF, Madupu R, Nelson W C, Ayodeji B, Kraul M, Shetty J, Malek J, Van Aken SE, Riedmuller S, Tettelin H, Gill SR, White O, Salzberg SL, Hoover DL, Lindler LE, Halling SM, Boyle SM, Fraser CM. 2002. The *Brucellus* genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc Natl Acad Sci. 99:13148-13153.

Pouttu R, Puustinen T, Virkola R, Hacker J, Klemm P, Korhonen TK. 1999. Amino acid residue Ala-62 in the FimH fimbrial adhesin is critical for the adhesiveness of meningitis-associated *Escherichia coli* to collagens. Mol Microbiol. 31:1747-1757.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. Nucleic Acids Res. 33:116-120.



**Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV.** 2003. Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. *Curr Biol.* 13:1512-1517.

**Rokas A, Kathirithamby J, Holland PW.** 1999. Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. *Insect Mol Biol.* 8:527-530.

**Sawires YS, Songer JG.** 2006. *Clostridium perfringens*: Insight into virulence evolution and population structure. *Anaerobe.* 12:23-43.

**Semblat JP, Rosso MN, Hussey RS, Abad P, Castagnone-Sereno P.** 2001. Molecular cloning of a cDNA encoding an amphid-secreted putative avirulence protein from the root-knot nematode *Meloidogyne incognita*. *Mol Plant Microbe Interact.* 14:72-79.

**Thiéry M, Mugniéry D.** 1996. Interspecific rDNA restriction fragment length polymorphism in *Globodera* species, parasites of Solanaceous plants. *Fundam Applied Nematol.* 19:471-479.

**Thiéry M, Mugniéry D, Bossis M, Sosa-Moss C.** 1997. Résultats de croisements entre *Globodera pallida* Stone et *G. mexicana* Campos-Vela : hérabilité du développement sur pomme de terre et notion d'espèce. *Fundam Applied Nematol.* 20:551-556.

**Tsai YH, Orsi RH, Nightingale KK, Wiedmann M.** 2006. *Listeria monocytogenes* internalins are highly diverse and evolved by recombination and positive selection. *Infect Genet Evol.* 6:378-389.

**Tudor JE, Pennington MW, Norton R.** 1998. Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin. *Eur J Biochem.* 251:133-141.

**Vanholme B, De Meutter J, Tytgat T, Van Montagu M, Coomans A, Gheysen G.** 2004. Secretions of plant-parasitic nematodes: a molecular update. *Gene.* 332:13-27.

**Von Mende N.** 1997. Invasion and migration behaviour of sedentary nematodes. In Fenoll C, Grundler FMW, Ohl SA, editors. Cellular and molecular aspects of plant-nematodes interactions. Kluwer Academic Publishers. p. 51-64.

**Wada H, Kobayashi M, Sato R, Satoh N, Miyasaka H, Shirayama Y.** 2002. Dynamic insertion-deletion of introns in deuterostome EF-1 alpha genes. *J Mol Evol.* 54:118-128.

**Weber E, Koebnik R.** 2006. Positive selection of the Hrp pilin HrpE of the plant pathogen *Xanthomonas*. *J Bacteriol.* 188:1405-1410.

**Williamson VM, Gleason CA.** 2003. Plant-nematode interactions. *Curr Opin Plant Biol.* 6:327-333.

# CHAPITRE 2

---

## CHAPITRE 2

### 1. ELEMENTS DE CONTEXTE : LES RANBPM

Le transcrit (IC5), autre candidat issu de l'hybridation suppressive et soustractive entre *G. pallida* et *G. "mexicana"*, présentait une homologie avec le facteur A18 identifié par cDNA-AFLP chez *G. rostochiensis* (Qin *et al.*, 2001). Lorsque nous avons débuté notre étude, ce facteur n'avait pas été décrit et présentait une homologie avec des protéines de type RanBPM.

Les protéines RanBPM (protéines Ran de liaison aux microtubules) présentent toutes un domaine très spécifique, SPRY, qui est retrouvé dans les protéines de type récepteur à la Ryanodine. Ce domaine est très conservé même entre des organismes aussi éloignés que l'Homme, la souris et le xénope (Wang *et al.*, 2002). Cette forte conservation au sein de la famille des RanBPM laisse penser que le domaine SPRY joue un rôle important dans la fonction de ces protéines, notamment en tant que domaine de liaison. En effet la bibliographie concernant les domaines de type SPRY mentionne un rôle important dans les interactions protéine-protéine. Plusieurs rôles des RanBPM tels que des dérégulations du cycle cellulaire ou des réorganisations du cytosquelette cellulaire ont été suggérés dans d'autres modèles biologiques, notamment chez les mammifères. Ce type de protéine aurait la capacité de se lier aux protéines Ran (Ras-like nuclear small GTPase) impliquées dans le transport nucléoplasmique, l'assemblage des microtubules et la formation des membranes nucléaires chez l'homme (Nishitani *et al.*, 2001). Les RanBPM seraient localisées dans les centrosomes des cellules au niveau du centre d'organisation des microtubules (MTOC). La surexpression de ces protéines a été observée dans les cellules cancéreuses dans les cas de cancer du sein (Emberley *et al.*, 2002). En extrapolant le rôle de ces RanBPM, on peut imaginer que le site nourricier puisse être apparenté à un cancer des cellules végétales où l'on observerait une dérégulation du cycle cellulaire au même titre que dans les cellules cancéreuses de mammifères. Ainsi on comprend que la surexpression de *rbp-1* pourrait être impliquée dans la mise en place du site nourricier. Les RanBPM ont également été détectées dans un complexe de 670 kDa associé à des RanGTPases dans différents compartiments cellulaires (Hafizi *et al.*, 2005) et ont été définies comme des protéines modulables ayant la capacité à interagir avec les protéines hôtes et donc à avoir un rôle de régulateur à différents niveaux. Les Ran ont notamment un rôle dans la régulation du cycle cellulaire mitotique. Etant donné que les cellules du site nourricier ne présentent pas de phase de mitose, les RanBPM de nématodes pourraient intervenir dans le phénomène d'endoreduplication en se liant aux RanGAP et en favorisant la production de RanGTP (Fig. 1A). Or, Heald & Weis (2000) ont montré que l'augmentation du niveau de RanGTP dans la cellule provoque le désassemblage des chromosomes et des microtubules provoquant la formation d'asters ectopiques de microtubules (Fig. 1B & C).

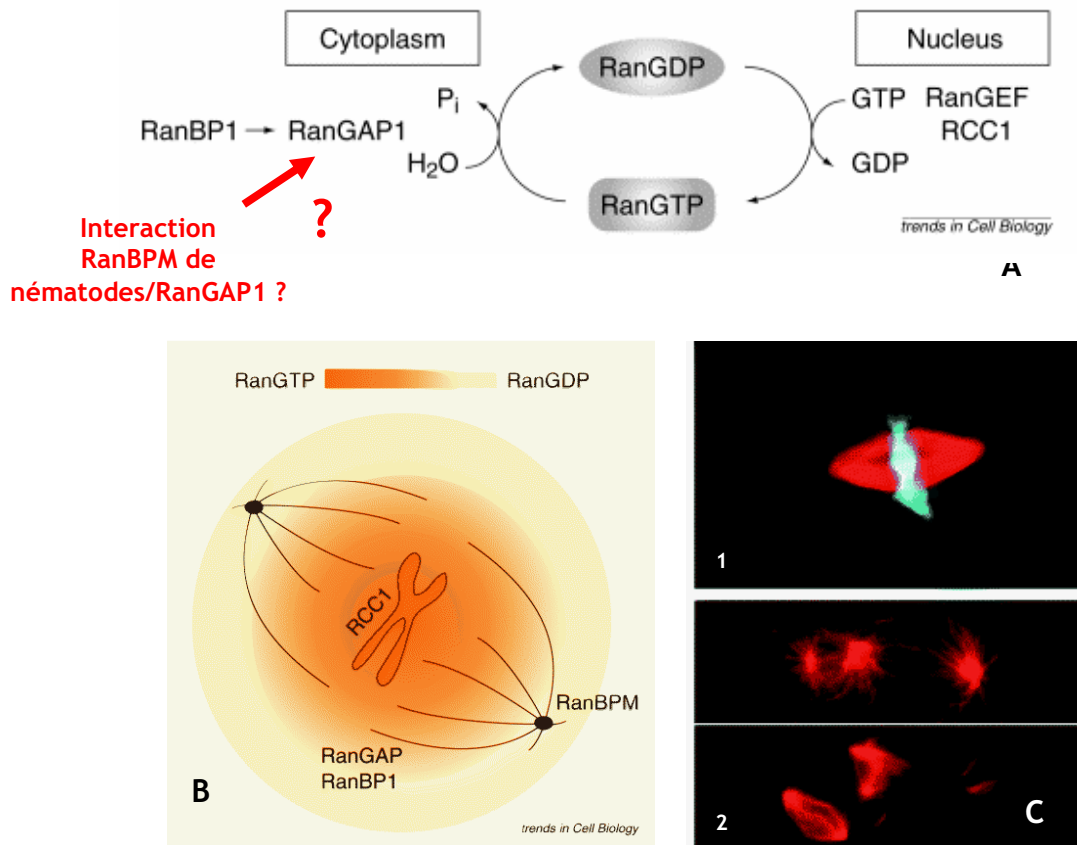


Figure 1 : Intervention des protéines de type RanBPM dans les régulations du cycle cellulaire (Heald & Weis 2000, Trends in Cell Biology). A : Cycle des RanGTP. B : Modèle montrant le rôle des régulations possibles du cycle cellulaire. Les RanGAP et RanBP1 conduisent à l'hydrolyse du GTP par les Ran, générant un gradient de RanGTP qui favorise l'assemblage des microtubules grâce aux protéines de type RanBPM. C: Assemblages des chromosomes (vert) et microtubules (rouge) lors de la mitose, dans des œufs de xénopes (1). (2) Assemblage ectopique des microtubules en augmentant le niveau de RanGTP.

L'article de ce chapitre synthétise les résultats obtenus sur l'identification et la caractérisation, pour la première fois chez *G. pallida*, d'un gène homologue de RanBPM que nous avons nommé *rbp-1*. Ces résultats sont les premiers publiés sur les RanBPM de nématodes phytoparasites. Les questions centrales de cet article étaient :

- 1- Le transcrite homologue de RanBPM identifié est-il impliqué dans le parasitisme ?
- 2- Quelle est la fonction de cette protéine dans le parasitisme ?

## 2. ARTICLE 2

**ENCADRE N ° 2 : RANBPM HOMOLOGUE GENES CHARACTERISED IN THE CYST NEMATODES**

**GLOBODERA PALLIDA AND GLOBODERA "MEXICANA"**

**IMPLICATION DU GENE IC5 DANS LE PARASITISME ?**

- Localisation du transcrit dans la glande dorsale des seconds stades larvaires des nématodes.
- Identification d'un signal de sécrétion dans les séquences des transcrits chez *G. pallida* et *G. "mexicana"*.
- Expression du gène *rbp-1* spécifique des stades larvaires. Aucune amplification aux stades adultes.
- Identification d'un domaine conservé SPRY, ayant un rôle dans les interactions protéine-protéine.

La protéine codée par le gène est probablement sécrétée dans les cellules végétales et aurait un rôle dans les phases tardives du parasitisme, la mise en place et le maintien du site nourricier. Les RanBPM pourraient être liées à la dérégulation du cycle cellulaire, sans phase de mitose, observée dans les cellules nourricières.

**VARIABILITE DU GENE *RBP-1* AU SEIN DES NEMATODES DE LA FAMILLE DES HETERODERIDAE ?**

- Pas d'amplification de la copie *rbp-1* hormis chez *G. pallida* et *G. "mexicana"* avec le même jeu d'amorces.
- Plusieurs copies du gène *rbp-1* chez *G. pallida* identifiées par Southern blot.
- Divergence importante entre *rbp-1* et la copie identifiée chez *G. rostochiensis* (44% de similarité).

Le gène *rbp-1* appartient probablement à une famille multigénique qu'il convient de caractériser avant de comparer des séquences de RanBPM entre différentes espèces de nématodes à kyste afin de s'assurer au préalable de leur orthologie.

**ENCADRE N° 2 : RANBPM HOMOLOGUE GENES CHARACTERISED IN THE CYST NEMATODES**  
**GLOBODERA PALLIDA AND GLOBODERA "MEXICANA"**

**IMPLICATION DU GENE IC5 DANS LE PARASITISME ?**

- Localisation du transcrit dans la glande dorsale des seconds stades larvaires des nématodes.
- Identification d'un signal de sécrétion dans les séquences des transcrits chez *G. pallida* et *G. "mexicana"*.
- Expression du gène *rbp-1* spécifique des stades larvaires. Aucune amplification aux stades adultes.
- Identification d'un domaine conservé SPRY, ayant un rôle dans les interactions protéine-protéine.

La protéine codée par le gène est probablement sécrétée dans les cellules végétales et aurait un rôle dans les phases tardives du parasitisme, la mise en place et le maintien du site nourricier. Les RanBPM pourraient être liées à la dérégulation du cycle cellulaire, sans phase de mitose, observée dans les cellules nourricières.

**VARIABILITE DU GENE *RBP-1* AU SEIN DES NEMATODES DE LA FAMILLE DES HETERODERIDAE ?**

- Pas d'amplification de la copie *rbp-1* hormis chez *G. pallida* et *G. "mexicana"* avec le même jeu d'amorces.
- Plusieurs copies du gène *rbp-1* chez *G. pallida* identifiées par Southern blot.
- Divergence importante entre *rbp-1* et la copie identifiée chez *G. rostochiensis* (44% de similarité).

Le gène *rbp-1* appartient probablement à une famille multigénique qu'il convient de caractériser avant de comparer des séquences de RanBPM entre différentes espèces de nématodes à kyste afin de s'assurer au préalable de leur orthologie.

**RANBPM HOMOLOGUE GENES CHARACTERISED IN THE CYST  
NEMATODES *GLOBODERA PALLIDA* AND *GLOBODERA "MEXICANA"***

Alexandra BLANCHARD, Magali ESQUIBET, Didier FOUVILLE, Eric GRENIER.

Institut National de la Recherche Agronomique (INRA) - Agrocampus, unité mixte de recherche  
Biologie des organismes et des populations appliquée à la protection des plantes (UMR BiO3P),  
Domaine de la Motte au Vicomte, BP 35327, 35653 Le Rheu cedex France.

Tel: +33-2-23.48.51.91 Fax: +33-2-23.48.51.50

Corresponding author e-mail addresses: [alexandra.blanchard@rennes.inra.fr](mailto:alexandra.blanchard@rennes.inra.fr),  
[eric.grenier@rennes.inra.fr](mailto:eric.grenier@rennes.inra.fr)

*Publié dans Physiological & Molecular Plant Pathology*



## *Ranbpm* homologue genes characterised in the cyst nematodes *Globodera pallida* and *Globodera 'mexicana'*<sup>☆</sup>

Alexandra Blanchard\*, Magali Esquibet, Didier Fouville, Eric Grenier

Institut National de la Recherche Agronomique (INRA)—Agrocampus, Unité Mixte de Recherche Biologie des Organismes et des Populations appliquée à la Protection des Plantes (UMR BIO3P), Domaine de la Motte, BP 35327, 35653 Le Rheu Cedex France

Accepted 8 September 2005

### Abstract

A homologue of a Ran binding protein to microtubules (*ranbpm*) gene, initially known as IC5, was identified and characterised in second-stage juveniles (J2s) of the potato cyst nematode *Globodera pallida*. The full-length cDNA (937 bp) was obtained by 5' and 3' rapid amplification of cDNA ends (RACE) and specific primers were designed to amplify the genomic sequences of 2396 bp containing six introns. The ORF (798 bp) encodes a putative 265 amino acid sequence with a predicted SPRY domain and a signal peptide of 23 amino acids on the N-terminal part of the protein. In situ hybridisation experiments showed that the transcript is located in the dorsal gland of the J2s, suggesting that the encoded protein has an extracellular function and can be involved in the late stages of parasitism such as feeding site establishment. This gene was specifically over-expressed in the juveniles before and during parasitism, but not in adult developmental stages. As this gene is presumed to be involved in plant–nematode interaction, particularly in the development and maintenance of the feeding structure that allows the nematode to achieve parasitic development, homologous genes were sought in other cyst nematode species. One of them was cloned and sequenced in the closely related species *Globodera 'mexicana'*.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Pathogenicity; In situ hybridisation; Oesophageal gland; RT-PCR; Nematode; SPRY

### 1. Introduction

Potato cyst nematodes are responsible for severe damage in cultivated Solanaceous plants throughout the world. This damage represents 300 million dollars in potato crop yield in Europe every year [1]. Plant resistance is an effective and environmentally friendly way of controlling potato cyst nematodes. However, its application to *Globodera pallida* is hampered by the few sources of resistance yet to be identified [2]. In order to circumvent this problem, artificial resistances can be developed [3] but require good knowledge of the genes implicated in plant–nematode interaction.

These quarantine nematodes develop an intimate relationship with their host via the induction of a complex structure, a feeding site otherwise known as the syncytium, in the vascular cylinder of the potato roots. This syncytium is induced 48 h p.i. and used by the cyst nematode to withdraw nutrients from its host. The three oesophageal glands of the cyst nematodes are presumed to be involved in parasitism, because they are the production site of the secretory products that can be injected into plant cells via the nematode stylet [4]. Secretions from the two subventral glands are often involved in the penetration and migration steps, while the dorsal gland secretions seem to be involved in feeding site formation [4,5]. Indeed, the development and functioning of the syncytium are among the most complex responses induced by a nematode in plant tissues. The cyst nematodes choose one cell and then induce cell wall collapse in adjacent cells. Syncytium induction involves both proliferation of endoplasmic reticulum, ribosomes, mitochondria and plastids in the cell, and enlargement of the nucleus without mitosis, a phenomenon known as endoreduplication [6].

Thus, secretory products are preferential targets for studying plant–nematode interactions due to their key role in parasitism. Several genes corresponding to such secretory products have

**Abbreviations** J2, second juvenile stage; GP, *Globodera pallida*; GM, *Globodera 'mexicana'*; Dig, digoxigenin.

<sup>☆</sup> The nucleotide sequence data reported in this paper is available in the GenBank™ database under accession number AY769949.

\* Corresponding author. Tel.: +33 2 23485173; fax: +33 2 23485150.

**E-mail addresses:** alexandra.blanchard@rennes.inra.fr (A. Blanchard), eric.grenier@rennes.inra.fr (E. Grenier).

0885-5765/\$ - see front matter © 2005 Elsevier Ltd. All rights reserved.  
doi:10.1016/j.pmpp.2005.09.001



already been identified in cyst nematodes. Many correspond to proteins with cellulolytic activity such as cellulase [7–9], pectate lyase [10] and expansin [11] and are involved in the migratory process. Few correspond to proteins putatively involved in the development and maintenance of the feeding site, one example being chorismate mutase [12].

Among the genus *Globodera*, *Globodera pallida* and *Globodera 'mexicana'* are genetically close species that can be differentiated through their host range [13,14]. *G. pallida* develops on *Solanum tuberosum* whereas *G. 'mexicana'* develops on *Solanum nigrum*, but not on *S. tuberosum* [14]. In a previous study, the transcriptome of these two species was compared by SSH [15]. One of the SSH fragments, designated as GPLIC5, was identified after subtraction of *G. pallida* transcripts by *G. 'mexicana'* transcripts and was found to match the A18 factor. This A18 factor was isolated in *Globodera rostochiensis* by cDNA-AFLP and had homology with Ran binding protein to microtubules (RanBPM) genes [16].

The purpose of this study was to gain new insight into the IC5 gene. In situ hybridisation was carried out to localise the transcript and its full-length sequence was isolated by RACE-PCR. Data presented in this paper sheds new light on the expression pattern and secretion of the protein derived from this gene that appears to encode a RanBPM. Subsequently, a homologue was cloned in *G. 'mexicana'* and potential functional differences were sought by comparing *G. pallida* and *G. 'mexicana'* *ranbpm* genes.

## 2. Materials and methods

### 2.1. Nematodes

*G. pallida* population Guiclan (France) was mainly used in this study. The cysts were obtained from potato crops (*S. tuberosum* var. Désirée) in laboratory conditions. In order to stimulate juvenile hatching (J2s), the cysts were soaked in water for 4 days and then in potato root diffusate at room temperature. The J2s were stored in water at 4 °C for in situ hybridisation, or at –80 °C after centrifugation for nucleic acid extraction.

### 2.2. Isolation of 5' and 3' cDNA ends

#### 2.2.1. RNA isolation

Total RNA was extracted from 50 to 100 mg of J2 with the RNeasy Mini kit (Sigma) according to the manufacturer's instructions. The RNAs were treated with DNase I (10 U, Ambion) for 15 min at 37 °C and then purified by phenol/chloroform/isoamyl alcohol (25/24/1) extraction. The RNAs were precipitated with sodium acetate 3 M and absolute ethanol, and stored at –80 °C.

#### 2.2.2. mRNA reverse transcription

Reverse transcription was performed in a mix containing 2 µg of total RNA, 0.25 µM of d(T)<sub>25</sub> primer and 8.5 µl of RNase-free water. The samples were heated for 10 min at 70 °C and set on ice. 1X superscript buffer, 2.5 mM MgCl<sub>2</sub>,

0.25 µM of each dNTP and 0.01 µM of DTT were then added. The sample was heated for 1 min at 42 °C. Superscript III (Invitrogen, 200 U/µl) was added and the mix was incubated at 42 °C for 50 min, and at 70 °C for 15 min. One microlitre of RNase mix (Boehringer) was added and the mixture was incubated for 30 min at 37 °C.

#### 2.2.3. 5' and 3' RACE

5' and 3' RACE [17] were performed using the 3' and 5' RACE System for rapid amplification of cDNA ends (Invitrogen) according to the manufacturer's instructions. The 3' end of the transcript was amplified by PCR using two microlitres of the first strand cDNA, 1X buffer, 1.5 mM of MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.4 µM of AUAP primer (Table 1), 0.4 µM of an internal specific primer of the SSH fragment (GPLIC5 Fwd; Table 1) and 1 U of Taq polymerase (Promega). PCR was performed with 30 cycles at 96 °C for 20 s, 58 °C for 20 s and 72 °C for 30 s, with a final elongation step at 72 °C for 5 min.

5' RACE reaction: the first strand cDNA was obtained from total RNAs using 0.4 µM of specific primer GPLIC5Rev2 (Table 1). Nested PCR was performed in 50 µl with 1X buffer, 1.5 mM of MgCl<sub>2</sub>, 0.2 mM of each dNTP, 2.5 units of Taq DNA polymerase, 5 µl of cDNA, 0.4 µM of AAP, and 0.4 µM of the internal specific primer (GPLIC5Rev2 or GPLIC5Rev3; Table 1). PCR conditions were 35 cycles at 96 °C for 20 s, 53 °C for 30 s and 72 °C for 30 s, with a final elongation step at 72 °C for 5 min.

The PCR products were purified (GeneElute PCR Clean-Up kit, Sigma) and cloned in pGEM-T Easy Vector System I (Promega) according to the manufacturer's instructions.

### 2.3. Cellular localisation and developmental expression

#### 2.3.1. mRNA in situ hybridisation

The GPLIC5Fwd and GPLIC5Rev2 primers (Table 1), designed from the GPLIC5 SSH clone, were used to amplify the insert and to synthesise digoxigenin (DIG)-labelled sense and anti-sense probes by asymmetric PCR amplification.

Insert amplification was performed in 50 µl with 30 ng of insert, 0.4 µM of each primer, 1X buffer, 0.2 mM of each

Table 1  
The primers used in this study are summarised in the table

Name	Sequences (5'→3')
RACE primers	
GPLIC5fwd	GGCTATTTTGGGGTCAC
GPLIC5Rev2	CGTATGAGTAAGTGCCCT
GPLICFRev3	TCAC TTCATAGTAGAAAATGCC
AUAP	GGCCACGCGTCGACTAGTAC
AAP	GGCCACGCGTCGACTAGTACGGGHHGGGHHGGGHH
Other primers	
5'IC5	TTTTTATTGCCCCCAAATG
5'IC5.2	AAATGCGCACCTTTCTCTTC
3'IC5.3	GCAAACCCATCATAAATCTCG
3'IC5.2	CATGATCAATAGGCAAACCG
GPFARF	TCGCTGCCGCCATTGAT
3'GPFAR	GGCGCCTCAGTCTTCAGCA

dNTP, 2 U of Taq DNA Polymerase (Promega) and 1.5 mM MgCl<sub>2</sub>, for 4 min at 94 °C, 40 cycles at 94 °C for 30 s, 56 °C for 30 s, 72 °C for 1 min, and a final elongation step of 4 min at 72 °C. Asymmetric PCR was performed in 40 µl with 2 µl of the purified insert, 2.5 µM of a primer (Fwd or Rev), 1X buffer, 1.5 mM of MgCl<sub>2</sub>, 0.75X Dig-dUTP/dNTP mix (Roche Applied Science) and 2 U of Taq DNA polymerase (Promega) for 4 min at 94 °C, 34 cycles at 94 °C for 30 s, 55 °C for 1 min, 72 °C for 90 s, and a final elongation step of 4 min at 72 °C. In situ hybridisation was performed as described by De Boer et al. [18] with slight differences. Freshly hatched J2s were fixed in fixation buffer (including paraformaldehyde) for 18 h at 4 °C and 4 h at room temperature. The nematodes were cut with a razor blade and the sections were permeabilised with proteinase K for 30 min at 22 °C before freezing at –80 °C. Hybridisations were performed at 45 °C with purified sense or anti-sense single strand cDNA probes. The signal was detected using alkaline phosphatase immunostaining (NBT-BCIP, Boehringer).

### 2.3.2. *ranbpm* gene expression analysis

One to five nematodes were used to extract the RNAs. They were crushed with a modified Pasteur pipette, and recovered in a mix containing RNase-free water, 1 µl of d(T)<sub>25</sub> primer at 10 µM and 1 µl of RNasin. The mix containing the crushed nematodes was incubated for 10 min at 70 °C and then a solution containing 1X buffer, 2.5 µM MgCl<sub>2</sub>, 0.4 µM of each dNTP and 0.01M of DTT was added and incubation was resumed for 1 min at 42 °C before adding 200 U of Superscript III (Invitrogen). The reaction was incubated for 50 min at 42 °C, and then for 15 min at 70 °C. DNase I was used as previously described in RNA isolation.

The IC5 cDNA was amplified using 5 µl of the reverse-transcribed mRNAs in a mix containing 1X buffer, 1.5 mM of MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.4 µM of each specific primer (5' IC5.2 and 3' IC5.2; Table 1) and 2.5 units of Taq DNA Polymerase (Promega). As a positive control, the *Gp-far-1* gene reference was amplified in the same conditions with the specific primers GPFARF and 3' GPFAR (Table 1). PCR conditions were 5 min at 96 °C, 35 cycles at 96 °C for 20 s, 58 °C for 30 s, 72 °C for 30 s with an additional second per cycle, and a final elongation step for 5 min at 72 °C.

## 2.4. Identification of the genomic sequence

### 2.4.1. DNA isolation

Genomic DNA was extracted from the eggs of 500 rehydrated cysts. The frozen eggs were crushed with a piston. The DNA was extracted with lysis buffer (0.1 M Tris, pH 8, 10 mM EDTA, 2% SDS) and proteinase K, and then incubated at 65 °C for 1 h. The DNA was precipitated and purified with 5M NaCl and 10% CTAB, incubated for 10 min at 65 °C and then with a phenol/chloroform (1:1) mix. The DNA was precipitated with ammonium acetate 5 M overnight at 4 °C. A mix containing 30% of PEG 6000 and isopropanol (99.8%) was added to achieve DNA precipitation. The sample was treated with 1 µl of RNase A (500 µg/mL) for 2 h at 37 °C.

### 2.4.2. *ranbpm* gene amplification

The 5' IC5.2 and 3' IC5.3 primers (Table 1) designed from the IC5 cDNA sequence were used to amplify the *ranbpm* gene from 100 ng of genomic DNA. The PCR products were checked on 1% agarose gel and purified using GenElute PCR product kit (Sigma). The purified PCR products were cloned in the pCR2.1 vector using the TA cloning kit (Invitrogen) according to the manufacturer's instructions. The minipreps were performed using the GenElute Plasmid Miniprep kit (Sigma) according to the manufacturer's instructions.

## 2.5. Sequence analysis

The cloned inserts were sequenced by Macrogen (Korea) or using our own facilities (ABI310 machine, Perkin). WU-Blast2 Parasite Genomes Database Query on EMBL-EBI ([www.ebi.ac.uk/blast2/parasites.html](http://www.ebi.ac.uk/blast2/parasites.html)), and Blast on NCBI ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)) or on nemaBLAST (<http://www.nematode.net/>) were used to search for sequence homologies. The ExPasy translate tool (<http://us.expasy.org/tools/dna.html>) was used to predict amino acid sequences. Sequence alignments were performed with Multalin algorithm [19] (<http://prodes.toulouse.inra.fr/multalin/multalin.html>). The sequence signatures were detected using SignalP algorithm [20] ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), and ELM program [21] (<http://elm.eu.org/>). The Radar program [22] (<http://www.ebi.ac.uk/Radar/>) was used to identify repeats in amino acid sequences.

## 3. Results

### 3.1. Isolation of the full-length *ranbpm* cDNA in *G. pallida* and sequence analysis

Specific RACE-PCR primers were designed from the GPLIC5 SSH fragment. The GPLIC5Fwd primer was used to amplify the 3' end of the GPLIC5 transcript. A single band of approximate size 500 bp was visualised. GPLIC5Rev2 was used to amplify the 5' end of the GPLIC5 transcript but did not produce the entire cDNA. Another primer, GPLIC5Rev3, was therefore used to obtain the entire 5' end sequence. Once again, a single band of 500 bp was visualised on agarose gel after electrophoresis of the PCR products. These PCR products were cloned and sequenced. Using overlapping regions, a putative full-length cDNA of 937 bp was built in silico (Fig. 1). An ORF of 798 bp was identified and two specific primers (5' IC5 and 3' IC5.3) flanking this ORF were designed to check the in silico construction. The product obtained was about 800 bp, in keeping with the expected size of 824 bp. The corresponding sequence perfectly matched the IC5 transcript sequence built in silico and encoded a predicted protein of 265 amino acid with a predicted molecular weight of 29 kD. Marked similarities (E-value:  $7.10^{-43}$ , of 2 817 606 sequences available in Genbank) to the A18 factor from *G. rostochiensis* [16] and *ranbpm* genes from different mammalian, insect or fungal species (E-value  $<10^{-9}$ , of 2 817 606 sequences available in Genbank) were obtained using BlastP. Similar results were obtained with the WU-Blast2 or NemaBase databases that



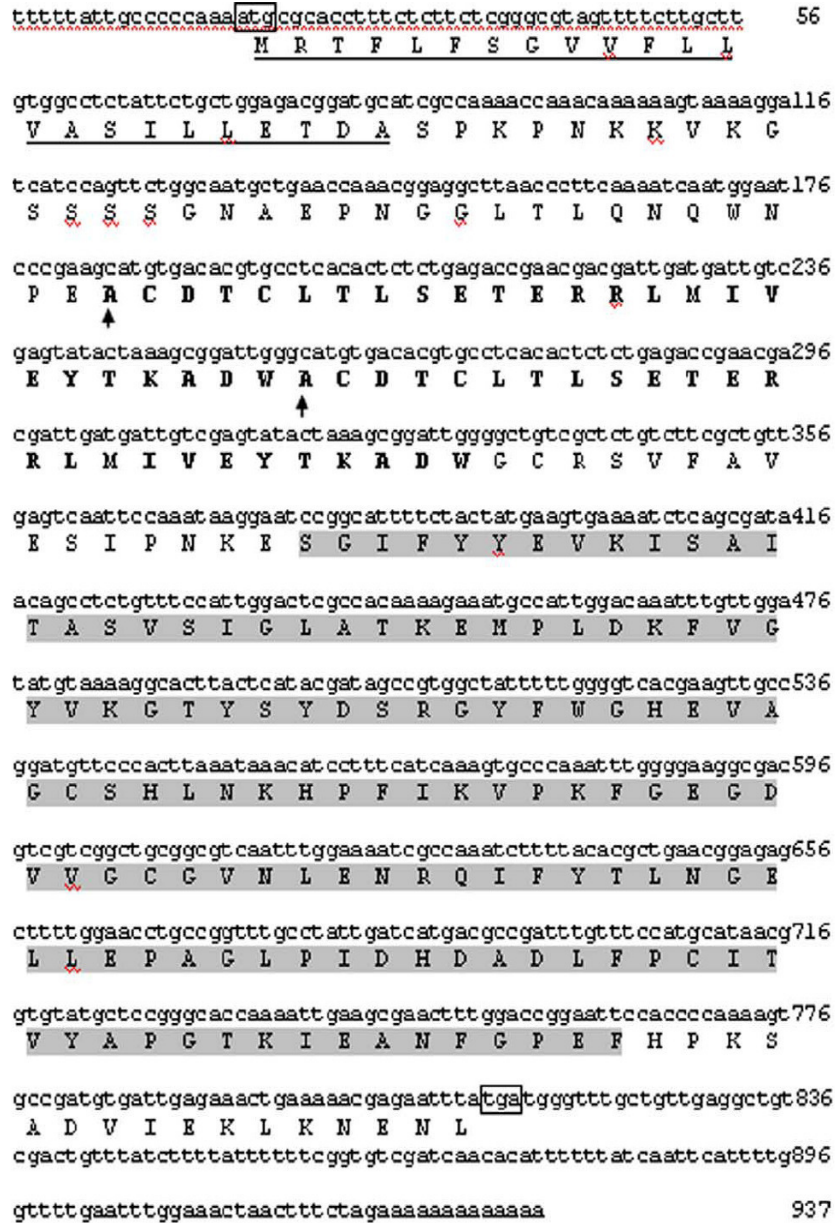


Fig. 1. Nucleic acid and amino acid sequence representation of *Globodera pallida ranbpm* cDNA. The start and stop codons are identified in the nucleotidic sequence by two boxes. The putative peptide signal of 23 amino acids is underlined on the 5' side of the transcript, and the SPRY domain is identified by a grey box. The tandem repeated sequence is shown in bold characters with arrows indicating each beginning of the repetition.

confirm the sturdiness of the blasts. SignalP 3.0 predicted the presence of a signal peptide of 23 amino acid in the N-terminal end of the protein (Fig. 1) indicating that the product of this gene is most probably secreted. The Radar program identified a tandem repeated sequence of 25 amino acids downstream of the start codon (Fig. 1). A motif search in the ELM database identified a SH3 ligand domain ('KPNKKVKG') in this region. A SPRY domain (domain in SPLa and the RYanodine receptor) was identified beginning at position 121 and ending at position 249 (E-value: 3.9E-20). Although of unknown function, this domain was identified in different proteins particularly, in association with LisH motifs that are involved in the regulation of microtubule dynamics.

### 3.2. *In situ* localisation of *G. pallida ranbpm* transcripts

The tissue localisation of the *ranbpm* transcript was analysed in *G. pallida* J2s by *in situ* hybridisation. The anti-sense digoxigenin-labelled DNA probe clearly showed a specific accumulation of *ranbpm* transcripts in the dorsal gland of the infective juveniles (Fig. 2(A)). No signal was observed using the GPLIC5fwd probe as a negative control (Fig. 2(B)).

### 3.3. Expression pattern of the *G. pallida ranbpm* gene

The developmental expression pattern of the *ranbpm* gene was characterised by RT-PCR amplification of part of

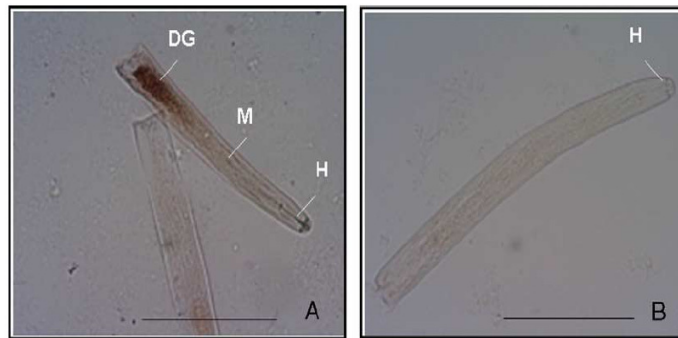


Fig. 2. Localisation by in situ hybridisation of the *rbp-1* transcript in the dorsal oesophageal gland of *Globodera pallida* infective juveniles. J2s were hybridised by anti-sense GPLIC5Rev2: (A) and sense GPLIC5Fwd; (B) digoxigenin-labelled DNA probes. H, head; M, metacarpus; DG, dorsal gland; Bar, 70 µm

the corresponding transcripts in five developmental stages of *G. pallida* (eggs, hatched J2s, J2s after penetration into roots, males, and females). RNAs from each stage were extracted and reverse transcribed and the cDNAs were amplified with specific primers (5' IC5.2 and 3' IC5.2). Gene expression was detected in eggs, pre-parasitic J2s and hatched J2s for which the strongest band was visualised at about 700 bp (Fig. 3(A)). No such IC5 gene expression was detected in the male and female stages whereas the constitutive GPFAR-1 cDNA fragment (492 bp) used as a positive control was amplified (Fig. 3(B)) in all the developmental stages.

### 3.4. Characterisation of the *ranbpm* gene in *G. pallida* and other Heteroderidae species

Genomic DNA of three *Globodera*, four *Heterodera* and one *Meloidogyne* species was used to amplify the genomic ORF of the *ranbpm* gene. Amplification products were obtained only with the *G. pallida* and *G. 'mexicana'* DNA. A band of about 2500 bp (Fig. 4) was obtained, suggesting the presence of introns in the *ranbpm* (*rbp*) gene with an ORF of only 798 bp. A second band was observed in *G. 'mexicana'* at about 2400 bp, suggesting that the *rbp* gene belongs to a multigenic family.

The two genes thereafter designated as *Gp-rbp-1* and *Gm-rbp-1* were cloned and sequenced. Their sizes were, respectively, 2396 and 2544 bp. The intron/exon boundaries were determined by aligning *Gp-rbp-1* and *Gm-rbp-1* with the corresponding *G. pallida* cDNA sequence. Six introns were identified in the two species at the same positions. Intron number 3 was located between the two 25 amino acid repeated motifs identified (Fig. 1). Intron size ranged from 42 to 890 bp (Fig. 5) and was relatively conserved between *G. pallida* and *G. 'mexicana'* except for intron number 2 which was 100 bp smaller in *Gp-rbp-1* than in *Gm-rbp-1*. The sequences of the corresponding introns were highly conserved, ranging from 95.2% for intron number 2 to 100% identity for intron number 4. *Gp-rbp-1* and *Gm-rbp-1* showed only two indels of three nucleotides (one in *Gp-rbp-1* and one in *Gm-rbp-1*) and one non-synonymous substitution resulting in 98.9% similarity between the coding regions of these two genes. Amino acid alignment of *G. pallida*, *G. 'mexicana'* and the closest *ranbpm* homologue found in the databases (AJ251757) for *G. rostochiensis* is shown in Fig. 6. Only 43.7% sequence similarity was observed between *Gp-rbp-1* and the *G. rostochiensis* RanBPM derived protein. Focusing on the signal peptide and SPRY domains, the comparison of the two sequences showed, respectively, 52% (18 aa) and 58% (128 aa)

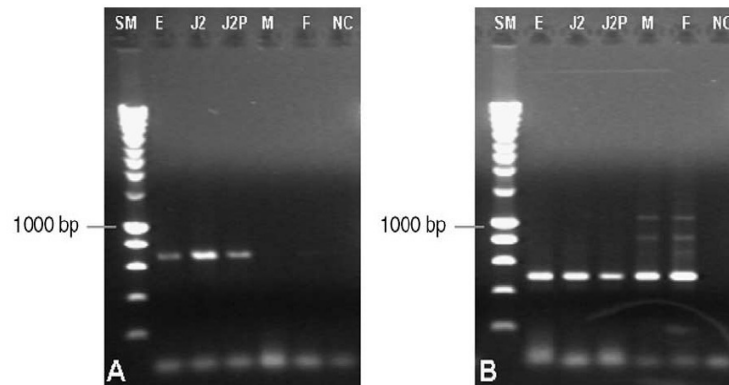


Fig. 3. Expression pattern of the *rbp-1* transcripts in different developmental stages of *Globodera pallida*: (A) PCR products obtained with *rbp-1* specific primers (5' IC5.2 and 3' IC5.2); (B) PCR products obtained as a control with the constitutively expressed gene *Gp-far-1* (specific primers: GPFARF and 3' GPFAR). E, eggs; J2, hatched juveniles; J2P, juveniles after penetration into roots; M, males; F, females; NC, negative control; SL, Smart ladder (Eurogentec).



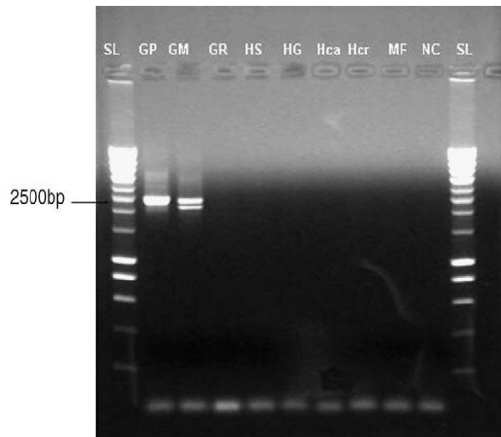


Fig. 4. PCR amplification of the *rbp-1* gene in different species of *Globodera*, *Heterodera*, and *Meloidogyne*. GP, *Globodera pallida*; GM, *G. 'mexicana'*; GR, *G. rostochiensis*; HS, *Heterodera schachtii*; HG, *H. glycines*; Hca, *H. carotae*; H. cr, *H. cruciferae*; MF, *Meloidogyne fallax*; NC, negative control; SL, smart ladder (Eurogentec). The DNAs were from INRA collection except the *H. glycines* DNA which was provided by K. Lambert.

sequence identity, indicating that both domains tend to be more conserved compared with the rest of the sequence.

#### 4. Discussion

Few genes with a potential role in the late events of parasitism, such as establishment of the syncytium, have been identified. Although *ranbpm* homologues sequences found in cyst nematodes exist in GenBank databases, this is the first study of the isolation and characterisation of *ranbpm* genes in *G. pallida* (*Gp-rbp-1*) and *G. 'mexicana'* (*Gm-rbp-1*). Candidate genes for pathogenicity are expected to correspond to products that can be secreted by the nematode. The predicted proteins from *Gp-rbp-1* and *Gm-rbp-1* both had a small size (MW: 29 kDa) compatible with the exclusion size of the lumen

of the nematode stylet. Protein secretion was supported by the identification of a signal peptide motif on the N-terminal side, as usually found in other secreted proteins [23]. Localisation of the *Gp-rbp-1* transcripts in the dorsal oesophageal gland cell of the J2s was obtained by in situ hybridisation. As this oesophageal gland is activated after the subventral ones [24], we can assume that the *Gp-rbp-1* gene is more likely involved in the latest stages of parasitism, such as feeding site development and maintenance. Other genes such as chorismate mutase or the *Skp1* genes [12,23] display similar transcript localisation in accordance with their putative role in the sedentary stages of parasitism. All these results provide indirect evidence of RanBPM protein secretion in the host plant by the nematode. It will be of considerable importance to develop of antisera capable of recognising this RanBPM protein to be able to confirm that it is secreted in planta and identify its plant cellular target.

In *G. pallida*, the *Gp-rbp-1* gene is expressed in the egg and J2 stages, but never detected in adult stages. This expression pattern has already been observed with other genes implicated in plant–nematode interactions such as the pectate lyases identified in *M. incognita* [25], or one copy of chorismate mutase identified in *M. incognita* and in *G. pallida* [26,27]. However, our data differ slightly from those published by Qin et al. [16] who indicated that the expression pattern of the *G. rostochiensis* A18 factor is J2-stage specific. Expression of the *Gp-rbp-1* gene obtained in the eggs is surprising but cannot be explained by genomic DNA amplification because the amplification product would have been larger due to the presence of introns. However, as a few J2s can also be found in non-stimulated eggs, this can lead to misinterpretation of the egg expression pattern. Further investigations will be needed to determine *Gp-rbp-1* gene expression pattern more precisely in quantitative terms.

During parasitism, cyst nematodes induce a feeding site which is characterised by genome multiplication with polytene

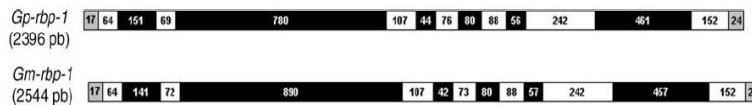


Fig. 5. Organisation of the *rbp-1* genes isolated in *Globodera pallida* and *G. 'mexicana'* species, *Gp-rbp-1* and *Gm-rbp-1*. Introns (black boxes), exons (open boxes) and non-transcribed sequences (grey boxes) are represented with the corresponding lengths (in bp) indicated in each box.

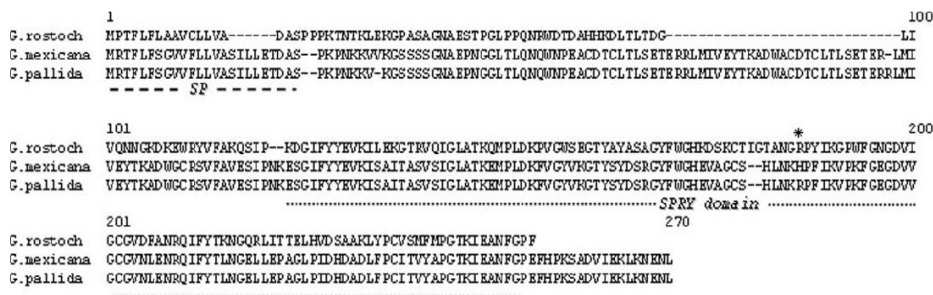


Fig. 6. Alignment of the predicted RBP-1 proteins in *Globodera pallida*, *G. 'mexicana'* and *G. rostochiensis* (*G. rostock*) (AJ251757). Conserved amino acids are in grey boxes, the signal peptide is dash underlined and the SPRY domain is dot underlined. The unique substitution observed between the *G. pallida* and *G. 'mexicana'* proteins is indicated by an asterisk.

nuclei where chromocentre size enlarges due to the endoreduplication mechanism [28]. Several sequence data converge to strongly support a specific role of the RBP-1 protein secreted by *G. pallida* and *G. 'mexicana'* in the development and maintenance of the feeding site as a key step of the parasitic cycle. Firstly, RanBPM-like proteins are ubiquitous and can be found independently of parasitic interactions. This suggests that the proteins RanBPM (RBP-1) normally associates with are available in plants. Secondly, weak similarities were identified between *rbp-1* and RanBPM-like proteins of the free-living organisms and particularly, in the nematode *C. elegans* (E-value:  $10^{-8}$ ), whereas strong similarities (E-value:  $10^{-42}$ ) were identified with RanBPM of other cyst nematode species. Furthermore, a SPRY domain was identified in an RBP-1 predicted protein. In some cases, this domain appears to correspond to a RYanodine receptor with an RNA binding role [29]. In other cases, the SPRY domain is associated with a LisH motif involved in the regulation of microtubule dynamics. No LisH domain was identified in predicted RBP-1, but a domain linked to a SH3 protein with a potential role in cytoskeleton organisation has been identified in both *Gp-rbp-1* and *Gm-rbp-1*. Clearly, functional investigations are now needed to assess the exact role of *rbp-1*. In this way, the RNAi method will be of interest, if accurate observations of the syncytium can be achieved in order to affirm that the phenotype observed is well induced by the inactivated gene.

Among the Heteroderidae family different species exist and have different host ranges. Consequently, it is worthwhile to study the variability of the *rbp-1* gene in different species. The primers used allowed us to detect the gene only in *G. pallida* and the closely related species, *G. 'mexicana'*. The high conservation of the protein sequence observed between *Gp-rbp-1* and *Gm-rbp-1* (98.9%) and their common intron distribution strongly suggest that they are orthologous copies. Six introns of 42–890 bp were observed in *rbp-1* genes. In the opposite, a high degree of variability was observed between *G. rostochiensis* and *G. pallida* protein sequences. Compared to the cellulase genes, the similarity observed (56.3%) for *ranbpm* genes is quite low as 93% similarity was found between *G. pallida* and *G. rostochiensis* cellulases [15], and 71.4% between *G. rostochiensis* (GR-eng1) and *Heterodera glycines* (HG-eng2) cellulases [30]. Therefore, the *ranbpm* isolated in *G. rostochiensis* could well be a paralogous copy of *rbp-1*. *Ranbpm* genes identified in *G. pallida* and *G. rostochiensis* may have different roles in plant–nematode interaction as was shown by Gao et al. [23,31] for cellulases HG-ENG5 and HG-ENG6 which have only 42–47% similarity and different enzymatic properties.

No amplification was observed with the other nematode species tested despite the fact that RanBPM-like ESTs were already sequenced in *G. rostochiensis* and *H. glycines*. Assuming that *rbp-1* is involved in feeding site initiation, there is no reason why this gene should be absent from *Globodera* and *Heterodera* species that all display the same syncytium phenotype regardless of the host plant. In order to obtain homologues, alignment of sequences and designing primers to conserved regions or an alternative approach such as

library screening would be required. However, the presence of *rbp-1* in *Meloidogyne* remains questionable as root-knot nematodes induce morphologically similar feeding sites (giant cells) but using different mechanisms. For example, in root-knot nematodes, feeding site development is characterised by an increase in the number of nuclei through polyploidisation mitosis that involves mitosis without cell division but not endoreduplication like in cyst nematodes [28]. We can assume that if *ranbpm*-like genes exist in *Meloidogyne*, they probably have a different role in plant parasitic interaction and may also be quite divergent in sequence compared to *rbp-1*. Further investigations are underway to find *ranbpm*-like genes in *Heterodera* and *Meloidogyne* species. This will give new insights into the specificity of *rbp-1* in plant–nematode interaction.

### Acknowledgements

We are grateful to Kris Lambert for kindly providing us with *H. glycines* DNA, and Marie-Noëlle Rosso and Stéphanie Jaubert for training in the in situ hybridisation technique.

### References

- [1] Mulholland V, Carde L, O'Donnel KJ, Fleming CC, Powers TO. Use of polymerase chain reaction to discriminate potato cyst nematode at the species level. Diagnostics in crop protection, British crop council 1996; 65:247–52.
- [2] Mugniéry D, Fouville D, Dantec JP, Pellé R, Rousselle-Bourgeois F, Ellissèche D. Résistance à *Globodera pallida* Pa2/3 chez *Solanum sparsipilum*. Nematology 2001;3:619–26.
- [3] Nolke G, Fischer R, Schillberg S. Antibody-based pathogen resistance in plants. Journal of Plant Pathol 2004;86:5–17.
- [4] Fenoll C, Grundler FMN, Ohl SA. Invasion and migration behaviour of sedentary nematodes. Cellular and molecular aspects of plant–nematodes interactions. New York, NY: Kluwer Academic Publishers; 1997:51–64.
- [5] Davis EL, Hussey RS, Baum TJ, Bakker J, Schots A, Rosso MN, et al. Nematode parasitism genes. Annu Rev Phytopathol 2000;38:365–96.
- [6] Gheysen G, Fenoll C. Gene expression in nematode feeding sites. Annu Rev Phytopathol 2002;40:191–219.
- [7] Smant G, Stokkermans JP, Yan Y, De Boer JM, Baum TJ, Wang X, et al. Endogenous cellulases in animals: Isolation of beta -1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc Natl Acad Sci 1998;95:4906–11.
- [8] Wang X, Meyers D, Yan T, Baum T, Smant G, Hussey R, et al. In planta localization of a beta-1,4-endoglucanase secreted by *Heterodera glycines*. Mol Plant–Microbe Interact 1999;12:64–7.
- [9] Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS. Identification of a new beta-1,4-endoglucanase gene expressed in the oesophageal subventral gland cells of *Heterodera glycines*. J Nematol 2002;34:12–15.
- [10] De Boer JM, McDermott JP, Davis EL, Hussey RS, Popeijus H, Smant G, et al. Cloning of a putative pectate lyase gene expressed in the subventral oesophageal glands of *Heterodera glycines*. J Nematol 2002;34:9–11.
- [11] Qin L, Kudla U, Roze EHA, Goversc A, Popeijus H, Nieuwland J, et al. A nematode expansin acting on plants. Nature 2004;427:30.
- [12] Doyle EA, Lambert KN. *Meloidogyne javanica* chorismate mutase 1 alters plant cell development. Mol Plant–Microbe Interact 2003;16: 123–31.
- [13] Bossis M, Mugniéry D. Specific status of six *Globodera* parasites of solanaceous plants studied by means of two-dimensional gel

- electrophoresis with a comparison of gel patterns by a computed system. *Fund Appl Nematol* 1993;16:47–56.
- [14] Thiery M, Mugniéry D, Bossis M, Sosa-Moss C. Résultats de croisements entre *Globodera pallida* Stone et *G. mexicana* Campos-Vela: héritabilité du développement sur pomme de terre et notion d'espèce. *Fundam appl Nematol* 1997;20:551–6.
- [15] Grenier E, Blok VC, Jones JT, Fouvillle D, Mugniéry D. Identification of gene expression differences between *Globodera pallida* and *G. mexicana* by suppression subtractive hybridization. *Mol Plant Pathol* 2002;3: 217–26.
- [16] Qin L, Overmars B, Helder J, Popeijus H, Rouppe van der Voort JN, Groenink W, et al. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. *Mol Plant–Microbe Interact* 2000;13: 830–6.
- [17] Frohman MA, Dush MK, Martin GR. Rapid production of full length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Nat. Acad. Sci. USA* 1988;85:8998–9002.
- [18] De Boer JM, Yan YT, Smant G, Davis EL, Baum TJ. In-situ hybridization to messenger RNA in *Heterodera glycines*. *J Nematol* 1998;30:309–12.
- [19] Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988;16:10881–90.
- [20] Bendtsen DJ, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: signalP 3.0. *J Mol Biol* 2004;340:783–95.
- [21] Puntervoll P, Linding R, Gemund C, Chabanis-Davidson S, Mattingsdal M, Cameron S, et al. ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res* 2003;31:3625–30.
- [22] Pearson WR. Rapid and sensitive sequence comparison with FastaP and Fasta. *Methods Enzymol* 1990;183:63–98.
- [23] Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS. The parasitome of the phytonematode *Heterodera glycines*. *Mol Plant–Microbe Interact* 2003;16:720–6.
- [24] Jones JT, Robertson WM. *Nematode secretions*. C.Fenoll et al. (ed.), Kluwer Academics, Netherlands 1997: 98-106.
- [25] Huang G, Dong R, Allen R, Davis EL, Baum TJ, HusseyRS. Developmental expression and molecular analysis of two *Meloidogyne incognita* pectate lyase genes. *Int J Parasitol* 2005;35:685–92.
- [26] Huang G, Dong RUIH, Allen REX, Davis EL, Baum TJ, Hussey RS. Two chorismate mutase genes from the root-knot nematode *Meloidogyne incognita*. *Mol Plant Pathol* 2005;6:23–30.
- [27] Jones JT, Furlanetto C, Bakker E, Banks B, Blok VC, Chen Q, et al. Characterization of a chorismate mutase from the potato cyst nematode *Globodera pallida*. *Mol Plant Pathol* 2003;4:43–50.
- [28] Govere AJ, Almeida Engler J, Verhees S, Van der Krol J, Helder J, Gheysen G. Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol* 2000;43:747–61.
- [29] Ponting C, Schultz J, Bork P. SPRY domains in ryanodine receptors (Ca<sup>2+</sup>-release channels). *Trends in Biochem Sci* 1997;22:193–4.
- [30] Yan Y, Smant G, Stokkermans J, Qin L, Helder J, Baum T, et al. Genomic organization of four [beta]-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. *Gene* 1998; 220:61-70.
- [31] Gao BL, Allen KD, Davis EL, Baum TJ, Hussey RS. Developmental expression and biochemical properties of a beta-1,4-endoglucanase family in the soybean cyst nematode. *Heterodera glycines*. *Mol Plant Pathol* 2004;5:93–104.





# CHAPITRE 3

---

## CHAPITRE 3 :

### 1. ELEMENTS DE CONTEXTE : LA FAMILLE MULTIGENIQUE DES RANBPM

Le chapitre 2 montre que *rbp-1*, joue certainement un rôle important dans la mise en place du parasitisme. Les connaissances acquises suggèrent l'appartenance de ce gène à une famille multigénique (article 2). Dans ce chapitre, nous apportons des éléments de réponse à la question suivante :

Quelle est la variabilité horizontale (taille et nature de la famille de gène) et verticale (variabilité d'une copie au sein de différentes populations et espèces) des RanBPM chez le nématode à kyste du genre *G. pallida* ?

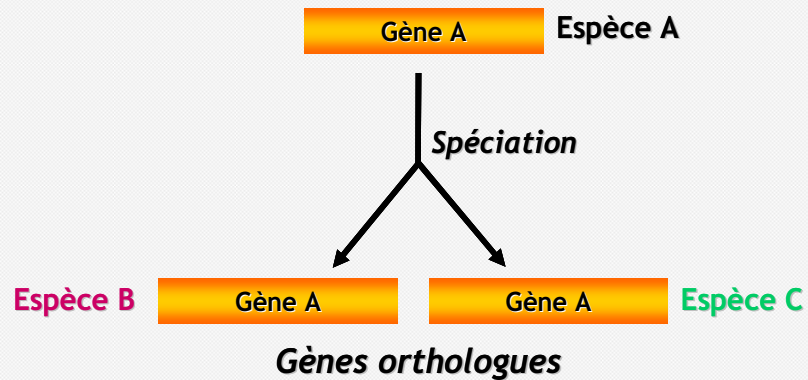
Une meilleure connaissance de la famille multigénique devrait permettre de savoir combien de gènes de ce type sont impliqués dans le parasitisme, d'étudier la variabilité de ces copies et d'identifier leur rôle dans l'interaction plante/nématode. Ces données devraient permettre de définir la ou les copies sur lesquelles il faut focaliser les efforts en matière d'acquisition de données à des fins cognitives - meilleure compréhension des mécanismes d'interaction - et appliquées - en matière d'identification cibles contre lesquelles il faudrait diriger les résistances artificielles ou de gestion des résistances naturelles qui interagiraient avec ce gène de nématode. Certaines études ont d'ores et déjà permis d'identifier différentes copies de gènes ayant des expressions différentes suggérant des fonctions différentes, comme les cellulases chez *H. glycines* (Gao *et al.*, 2004). Si la variabilité de la structure intron/exon a été étudiée, l'étude du polymorphisme de ces copies n'a pas encore été menée.

Dans le cadre d'une collaboration entre le laboratoire de Nématologie du Scottish Crop Research Institut (SCRI, Dundee, Ecosse) et le laboratoire de biologie et génétique des nématodes phytoparasites de Rennes ayant pour but d'aboutir à la production et à l'exploitation de nouvelles ressources génétiques pour comprendre les mécanismes de la pathogénie chez le nématode *G. pallida*, une banque d'ADNc de ce nématode a été générée. Nous avons criblé cette banque à la recherche de différentes copies de gènes homologues de RanBPM (Matériel & Méthode détaillé dans l'annexe 2). Un premier séquençage aléatoire de la banque avait révélé 8 clones homologues des séquences RanBPM disponibles dans les bases de données. Les clones de la banque identifiés suite au criblage et ayant une homologie significative ( $<10^{-5}$ ) avec d'autres RanBPM ont été désignés sous le terme "gènes RanBPM-like". Les résultats principaux sont présentés dans l'encadré 4 et l'analyse de ces séquences est présentée dans l'article qui suit.

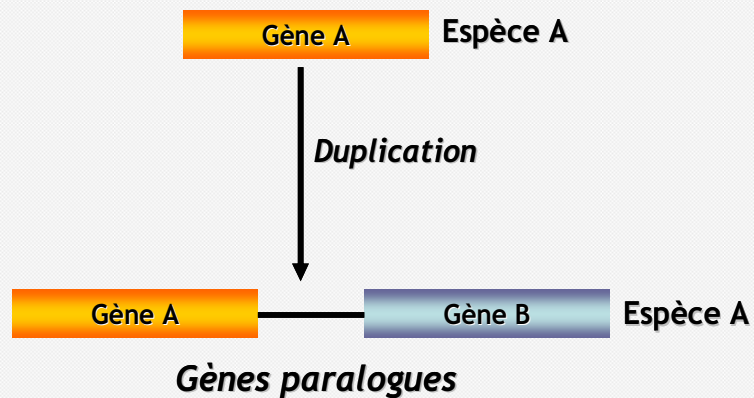
### 2. ARTICLE 3

**ENCADRE N° 3 : NOTIONS D'ORTHOLOGIE ET DE PARALOGIE**

- Gènes orthologues : gènes homologues qui ont divergé entre eux après un événement de spéciation.



- Gènes paralogues : gènes homologues qui ont divergé entre eux après duplication génique.



#### **ENCADRÉ N° 4 : CHARACTERISATION OF THE RANBPM-LIKE GENE FAMILY OF THE PLANT**

##### **PARASITIC NEMATODE *GLOBODERA PALLIDA***

#### **CRIBLAGE DE LA BANQUE D'ADNC DE *G. PALLIDA***

- 27 clones ont été identifiés avec les 4 sondes utilisées (voir annexe 2). Les inserts séquencés font 350 à 1100 pb.
- 20 clones sont homologues à des RanBPM de nématodes et correspondent à 12 copies différentes dont 10 ont un domaine SPRY tel que celui décrit pour *rbp-1*.
- 2 copies présentent une ORF complète avec un signal de sécrétion (peptide signal). Ces deux transcrits, codent respectivement des protéines de 283 et 208 acides aminés.

#### **ANALYSE DE LA VARIABILITE DES SEQUENCES RANBPM-LIKE**

##### *- Variabilité horizontale*

- Les 12 copies des gènes RanBPM-like sont très divergentes (distance génétique moyenne : 0,463, calculée sous Mega en utilisant le modèle Kimura 2 paramètres).
- L'arbre issu de l'alignement des ces séquences avec celles de *G. rostochiensis* disponibles dans les bases de données ne montre pas de copies orthologues chez *G. rostochiensis*. Seule une séquence se groupe avec *rbp-1* mais montre une diversité génétique importante (37% de sites variables au niveau acide aminé).

##### *- Variabilité verticale*

Les distances génétiques montrent une diversité plus importante au sein des populations péruviennes de *G. pallida* (0,2074), en comparaison avec les populations européennes (0,0284). La variabilité est répartie pour plus de 60% en première et deuxième bases de codon, suggérant un fort taux de substitutions non-synonymes. Ceci indique une pression sélective forte sur ce gène qui tend à créer de la diversité et montre ainsi une potentielle accélération de son tempo d'évolution.

Des analyses sont en cours pour déterminer 1- les patrons d'expression et la localisation cellulaire des transcrits des différentes copies de RanBPM, et 2- la variabilité du gène *rbp-1* sur un éventail de populations et d'espèces de *Globodera* plus important.

**ENCADRÉ N° 4 : CHARACTERISATION OF THE RANBPM-LIKE GENE FAMILY OF THE PLANT**

**PARASITIC NEMATODE *GLOBODERA PALLIDA***

**CRIBLAGE DE LA BANQUE D'ADNC DE *G. PALLIDA***

- 27 clones ont été identifiés avec les 4 sondes utilisées (voir annexe 2). Les inserts séquencés font 350 à 1100 pb.
- 20 clones sont homologues à des RanBPM de nématodes et correspondent à 12 copies différentes dont 10 ont un domaine SPRY tel que celui décrit pour *rbp-1*.
- 2 copies présentent une ORF complète avec un signal de sécrétion (peptide signal). Ces deux transcrits, codent respectivement des protéines de 283 et 208 acides aminés.

**ANALYSE DE LA VARIABILITE DES SEQUENCES RANBPM-LIKE**

*- Variabilité horizontale*

- Les 12 copies des gènes RanBPM-like sont très divergentes (distance génétique moyenne : 0,463, calculée sous Mega en utilisant le modèle Kimura 2 paramètres).
- L'arbre issu de l'alignement des ces séquences avec celles de *G. rostochiensis* disponibles dans les bases de données ne montre pas de copies orthologues chez *G. rostochiensis*. Seule une séquence se groupe avec *rbp-1* mais montre une diversité génétique importante (37% de sites variables au niveau acide aminé).

*- Variabilité verticale*

Les distances génétiques montrent une diversité plus importante au sein des populations péruviennes de *G. pallida* (0,2074), en comparaison avec les populations européennes (0,0284). La variabilité est répartie pour plus de 60% en première et deuxième bases de codon, suggérant un fort taux de substitutions non-synonymes. Ceci indique une pression sélective forte sur ce gène qui tend à créer de la diversité et montre ainsi une potentielle accélération de son tempo d'évolution.

Des analyses sont en cours pour déterminer 1- les patrons d'expression et la localisation cellulaire des transcrits des différentes copies de RanBPM, et 2- la variabilité du gène *rbp-1* sur un éventail de populations et d'espèces de *Globodera* plus important.

## CHARACTERISATION OF THE *RANBPM*-LIKE GENE FAMILY OF THE PLANT PARASITIC NEMATODE *GLOBODERA PALLIDA*

Alexandra BLANCHARD<sup>1</sup>, John JONES<sup>2</sup>, Eric GRENIER<sup>1</sup>.

<sup>1</sup> Institut National de la Recherche Agronomique (INRA) - Agrocampus, unité mixte de recherche Biologie des organismes et des populations appliquée à la protection des plantes (UMR BiO3P), Domaine de la motte, BP 35327, 35653 Le Rheu cedex France.

Tel: +33-2-23.48.51.91 Fax: +33-2-23.48.51.50

Corresponding author e-mail address: [alexandra.blanchard@rennes.inra.fr](mailto:alexandra.blanchard@rennes.inra.fr),  
[eric.grenier@rennes.inra.fr](mailto:eric.grenier@rennes.inra.fr)

<sup>2</sup> PPI Programme, Scottish Crop Research Institute, Invergowrie, Dundee. DD2 5DA  
Tel: +44 1382 562731 x2205

*En préparation*

## ABSTRACT

Sedentary endoparasitic nematodes have evolved an intimate relationship with their host plant who bring into play salivary secretions. Few are currently known on the molecular mechanisms involved in the manipulation of the host plant by the nematode for its own benefit. Some salivary proteins are suspected to play a role in the extraordinary cell modifications that occurred during elaboration and maintenance of the feeding structure by the nematode into the vascular cylinder of the plant. Among them, RanBPM were recently described in *Globodera pallida* and *Globodera "mexicana"* species. Previous results suggested that these genes are members of a multigenic family. In order to get a first knowledge of the size and diversity in the *G. pallida* RanBPM genes family, a cDNA library of this nematode was screened with different probes that allowed us to identify 12 new RanBPM-like genes. These copies were extremely divergent with a mean genetic distance of 0.38 and also appeared strongly different to the copies available in databases and issued from the other potato cyst nematode *Globodera rostochiensis*. Among the 12 copies characterised, two were full length copies possessing a secretion signal and a SPRY domain as described in previously isolated RanBPM like genes. The study of the sequence variability of one of these paralogs, the *rbp-1* copy, in different populations of *G. pallida* showed features that suggest that this gene may be under selection pressure that tend to generate more diversity and then suggest a fast evolutionary tempo.

## INTRODUCTION

Potato cyst nematodes of the genus *Globodera* are micro-organisms that infect plants of the Solanaceous family. The second larval stages of nematode penetrate the roots in the apical region and migrate intracellularly to the vascular cylinder. The nematodes choose one cell of the vascular system that will undergo strong modifications and will result in the elaboration of the feeding site, called syncytium. During the initiation of the syncytium, the cells become metabolically very active and are able to increase their nuclear material without mitosis. It is now admitted that nematode secretions play a central role in the initiation and maintenance of feeding structure. If numerous genes implicated in parasitism processes have already been described (Davis *et al.*, 2004; Vanholme *et al.*, 2004), few seemed to be implicated in the cell cycle modifications observed. Blanchard *et al.* (2005) have described one pathogenicity gene, *rbp-1*, that was probably involved in this process. RBP-1 displayed homology to RanBPM proteins that have the capacity to link to other proteins such as RanGTPases to form a complex to regulate de cell cycles (Heald & Weis 2000; Hafizi *et al.*, 2005). Moreover, *rbp-1* possessed a SPRY domain (domain in SPl $\alpha$  and the Ryanodine receptor; Pontig *et al.*, 1997) that seemed to play a key role in protein-protein interaction and appeared strongly conserved among distantly related species as mouse, xenope or human (Nakamura *et al.*, 1998; Wang *et al.*, 2002, Masters *et al.*, 2006). But, contrary to *rbp-1*, the other RanBPM genes such as the ones identified in *Caenorhabditis elegans* also present a LisH-CTHL (Lis1-homology-Cterminal to LisH) domain that is implicated in microtubule dynamics, cellular migration, nucleokinesis and

chromosomes segregation (Umeda *et al.*, 2003). This suggests that plant parasitic nematodes and free living nematodes probably possess different copies of RanBPM that displayed various biological roles. The existence of this multigenic family was reinforced both by the fact that no *rbp-1* gene copy was found in the *G. rostochiensis* and *G. tabacum* cyst nematodes (Blanchard *et al.*, 2005), and by the weak sequence similarity (44%) observed with the *G. rostochiensis* RanBPM genes characterised by Qin *et al.* (2000). It appeared therefore necessary to get a better knowledge of the RanBPM family diversity and polymorphism in order to be able to distinguish the various copies, a prerequisite before further works on the evolution of such orthologous genes or on the function of such genes using techniques as RNAi that require also to precisely target the gene that will be silenced.

The cDNA library used was generated in the SCRI from mRNA of second larval stages of the *G. pallida* Bedale population and using the SMART cDNA library construction kit from Clontec. For this library, the mRNAs were extracted, reverse transcribed and cloned into the bacterial vector pBluescript II. After transformation, 12672 clones were obtained. The <sup>32</sup>P-labelled probe used to screen this library that was previously transferred onto a nylon membrane, was generated from *rbp-1* and three distantly related sequences of RanBPM like genes identified in *G. pallida* in a previous EST sequencing project.

The screening of the cDNA library revealed 27 positive clones. Each clone was cultivated over night in LB and ampicilline (0.1µg/µL) and the plasmids were purified. The inserts of the 27 clones were sequenced (macrogen) and aligned using Multalin software (Corpet, 1988). The sizes of the insert sequences, ranged from 352 bp to 1101 bp. Blast searches revealed 20 sequences with significant homologies (E-value < 10<sup>-5</sup>) in WUBlast2 parasite ([www.ebi.ac.uk/blast2/parasites.html](http://www.ebi.ac.uk/blast2/parasites.html)) with Ranbpm-like genes from potato cyst nematodes (*G. pallida* or *G. rostochiensis*), and 7 sequences without homologies with RanBPM like transcripts of plant parasitic nematodes (Table 1).

Alignment of these 20 sequences allowed us to identify redundant clones that displayed inserts of different sizes but with sequences that have 100 % of homology. When such sequences were identified, they were grouped under the same sequence name. Finally, we identified 12 copies of RanBPM-like genes in *G. pallida* named *Gp-rbp-2* to *Gp-rbp-13* (Table 1). Ten of these various copies possessed a SPRY domain (Table 1) that was previously identified in *rbp-1*, but no LisH-CTHL domain was identified in any of these sequences. This means that the *G. pallida* copies probably have different implication in nematode biology than the *C. elegans* RanBPM. We noted that the three copies that did not possess a SPRY domain also corresponded to the shortest fragments analysed. When we interrogated the SMART program, the SPRY domain was identified but without strong E-value. We therefore assume that the SPRY is truncated in those three copies. Indeed this domain seems to be present in all the RanBPM-like copies identified in cyst nematodes.



Table 1: Identification of the RanBPM-like gene family in the *G. pallida* cDNA library. For each clone sequenced, the insert size and the results of the BLAST search are indicated. Detection of a SPRY domain and a signal peptide are shown by a (+). The two new full lengths RanBPM like genes found are highlighted in grey.

New RanBPM-like name	Clone	Insert size	Homolog accession number	Homology in WUblast2Parasite	E-value	SPRY	PS
<b>Ranbpm homologs</b>							
<i>Gp-rbp-2</i>	MP2-J1 MP2-J15	628 bp	<a href="#">BM414942</a>	EST from <i>G. pallida</i> similar to <i>G. rostochiensis</i> clone A18 mRNA sequence	8.8e-54	+	-
<i>Gp-rbp-3</i>	MP4H23	947 bp	<a href="#">BM344163</a>	rr47e07.y1 <i>G. rostochiensis</i>	8.6e-52	+	-
<i>Gp-rbp-4</i>	MP4-017 M05-M5 MP5-M6 MP33-L19	436 bp	<a href="#">BM414942</a>	EST from <i>G. pallida</i> similar to <i>G. rostochiensis</i> clone A18 mRNA sequence	4.8e-26	-	-
pseudogene	MP6-J22 *	465 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	8.7e-29	-	-
<i>Gp-rbp-5</i>	MP12-L6	688 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	3.9e-43	+	-
<i>Gp-rbp-5</i>	MP24-L21	966 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	1.6e-53	+	-
<i>Gp-rbp-5</i>	MP30-N10 M31-M21	770 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	1.8e-45	+	-
<i>Gp-rbp-6</i>	MP13-G11	965 bp	<a href="#">BM354257</a>	rr14d06.y1 <i>G. rostochiensis</i> similar RANBPM PROTEIN.	2.9e-47	+	+
<i>Gp-rbp-7</i>	MP15-K4	352 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	2.5e-20	-	-
<i>Gp-rbp-8</i>	MP18-I13	956 bp	<a href="#">BM356126</a>	<i>G. rostochiensis</i> J2 pcDNAII similar to RANBPM PROTEIN	6.6e-72	+	-
<i>Gp-rbp-9</i>	MP21-O12	824 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	6.7e-167	+	-
<i>Gp-rbp-10</i>	MP22-E10	827 bp	<a href="#">BM344163</a>	rr47e07.y1 <i>G. rostochiensis</i> J2 pcDNAII Smart v1	6.4e-44	+	+
<i>Gp-rbp-11</i>	MP26-L6	1101 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	9.2e-41	+	-
<i>Gp-rbp-12</i>	MP33-L19b	781 bp	<a href="#">AJ251757</a>	<i>G. rostochiensis</i> mRNA for for hypothetical protein (clone A18)	1.1e-29	+	-
<i>Gp-rbp-13</i>	MP33-B3	1063 bp	<a href="#">AY769949</a>	<i>G. pallida</i> IC5 mRNA	6.3e-49	+	-
<b>Others</b>							
-	MP1-N14	615 bp	<a href="#">BM345335</a>	rr54c12.y1 <i>G. rostochiensis</i> similar CYTOCHROME P450 3A29	2.4e-23	-	/
-	MP7-F18	801 bp	<a href="#">CA940623</a>	rq53c12.y1 <i>H. glycines</i> similar OXYGEN REGULATED PROTEIN	3.2e-25	-	/
-	MP14-L12	491 bp	<a href="#">CN576884</a>	rc52f02.x1 <i>M. hapla</i> similar MOLYBDOPTERIN-SYNTHASE LARGE SUBUNIT	7.4e-15	-	/
-	MP18-K8	816 bp	<a href="#">CV577941</a>	kf24f12.y1 <i>G. pallida</i> similar to PUTATIVE CYSTEINE DIOXYGENASE	3.8e-68	-	/
-	MP33-E10	807 bp	<a href="#">AW505666</a>	pal354 <i>G. pallida</i> gt11	1.1e-124	-	/
-	MP33-G14	813 bp	<a href="#">BM356200</a>	rr34g11.y1 <i>G. rostochiensis</i>	6.4e-37	-	/
-	MP33-H17	549 bp	<a href="#">CV579084</a>	kf07h08.y3 <i>G. pallida</i> similar PROBABLE 60S RIBOSOMAL PROTEIN	1.3e-106	-	/

\* insert with a degenerate ORF.

Among the copies identified, two (MP13-G11 and MP22-E10) were full length sequences, one (MP6-J22) had a degenerated ORF and was therefore considered as a pseudogene, and 10 possessed a partial ORF. We tried to obtain new full length cDNA using the SL1 sequence that is present in 70% of *C. elegans* transcripts (Blumenthal, 2005) as a primer in RT PCR. However no amplification was obtained for the four transcripts tested (three used to generate the probes and *rbp-1*) which suggest that these transcripts were not trans-spliced. The two full length transcripts we obtained had a signal peptide of 25 amino acids, as an indication of the secretion of the corresponding protein. The MP13-G11 transcript encoded a predicted protein of 283 aa with a molecular weight of 31kD and a predicted pI of 8.81 (ProtParam). The InterProScan interface allowed identifying an SPRY domain from aa 125 to 274, and a RanBP9 related domain from aa 214 to 283. The MP22-E10 transcript encoded a predicted protein of 208 aa with a molecular weight of 22.4kD and a pI of 9.09. A shorter SPRY domain was identified from aa 135 to 208. The predicted amino acids sequences were aligned with *Gp-rbp-1* (Fig. 1). Few aa positions were conserved among these three copies that displayed also several indels. However, some motifs of at least 7 aa were found conserved in the signal peptide or the SPRY domain and may be useful to design consensus primers for RanBPM genes.

In order to align as correctly as possible all these extremely divergent copies, the 12 transcript sequences and the *Gp-rbp-1* transcript were first aligned in amino acids using Tcoffee (Poirot *et al.*, 2003) and then the Revtrans 1.3 (Wernersson and Pederson, 2003) software was used to align the corresponding nucleotidic sequences. An average genetic divergence of 0.3837 was obtained using the Kimura 2 parameters model of the Mega software (complete deletion), which means that about 38% of the nucleotids are different.

All the data available in data bases on *G. rostochiensis* RanBPM like sequences were added to our data set to 1- evaluate the representativeness of our data among the *ranbpm* family of potato cyst nematodes, and 2- to identify some putative orthologous copies between *G. pallida* and *G. rostochiensis*. The tree based on the amino acids alignment of this new data set was generated using the Neighbour-Joining algorithm of MEGA in pairwise deletion conditions (Fig. 2). It showed that the *G. rostochiensis* copies grouped rather within each other than with *G. pallida* copies, except for BM343869.1 which grouped with *rbp-1* and *rbp-9* and BM356126.1 which grouped with *rbp-8*. *Rbp-1* and BM343869.1 amino acids sequences displayed 37% of variable sites (average p-distance: 0.3296 calculated in pairwise deletion with Mega) with some strongly conserved domains and the *rbp-8* and the BM356126.1 sequences displayed 24.4% of variable sites (aa p-distance: 0.258) and shared two large indels of 6 aa. The fact that only two copies identified in *G. rostochiensis* are related to some *G. pallida* sequences could be explained through two hypotheses: 1- the two nematodes possessed variable set of RanBPM genes, 2- the screening of the library was incomplete or some copies were not represented in the library. To test these hypotheses, one possibility could be to generate probes from ESTs of *G. rostochiensis* to screen one more time the *G. pallida* cDNA library and try to identify more copies of RanBPM-like genes.



Figure 1: Amino acids alignment (T-coffee software) of the three full length RanBPM-like genes (*Gp-rbp-6*, *Gp-rbp-10*, *Gp-rbp-1*) identified in the potato cyst nematode *G. pallida*. The signal peptide is underlined, the SPRY domain is indicated in dot boxes and the common amino acids are grey highlighted.

Nonetheless, the RanBPM gene family here described constitutes the first such large family of parasitism genes described to date in *G. pallida*. For example, the family members described to date for the cellulase, count only 5 copies described in *G. rostochiensis* (Yan *et al.*, 1998; Yan *et al.*, 2001) and 6 in *H. glycines* (Gao *et al.*, 2004). Moreover, taking into consideration the variability observed among the copies identified in this gene family, we can hypothesize that the different copies of RanBPM in *G. pallida* displayed various roles in the nematode parasitism. Clearly it will be now interesting to look at the differences in the patterns and levels of expression during nematode parasitism of some of these copies.

As the horizontal variability of RanBPM-like family revealed that the 12 copies identified were sufficiently divergent to be sure to target by PCR only one copy and study the variability among orthologs, we investigated the variability of *rbp-1* in different *G. pallida* populations. A part of the *rbp-1* transcript was amplified by RT-PCR on one cyst of 10 *G. pallida* populations originating from Europe and Peru and one population of *G. "mexicana"* (see table 2). We amplified a product of about 750 bp in *G. pallida* and of about 700 bp in *G. "mexicana"* that were sequenced (Macrogen). Using Blasts searches, we showed that all the sequences matched with *rbp-1* (AY699949), but

surprisingly the sequences obtained for the GPS5 and GPS6 populations displayed a best hit with another RanBPM like sequence (BM344321). The GPS5 and GPS6 sequences appeared distinct from the other as only 0.0461 of genetic distance (using Kimura 2 parameters model, in complete deletion and including the *G. "mexicana"* sequence) was observed in the dataset excluding these two sequences against 0.1611 when we included them in the dataset. Because of their high divergence compared to the other sequences, the GPS5 and GPS6 sequences were eliminated of further analyses since they were considered as paralogous copies of *rbp-1*.

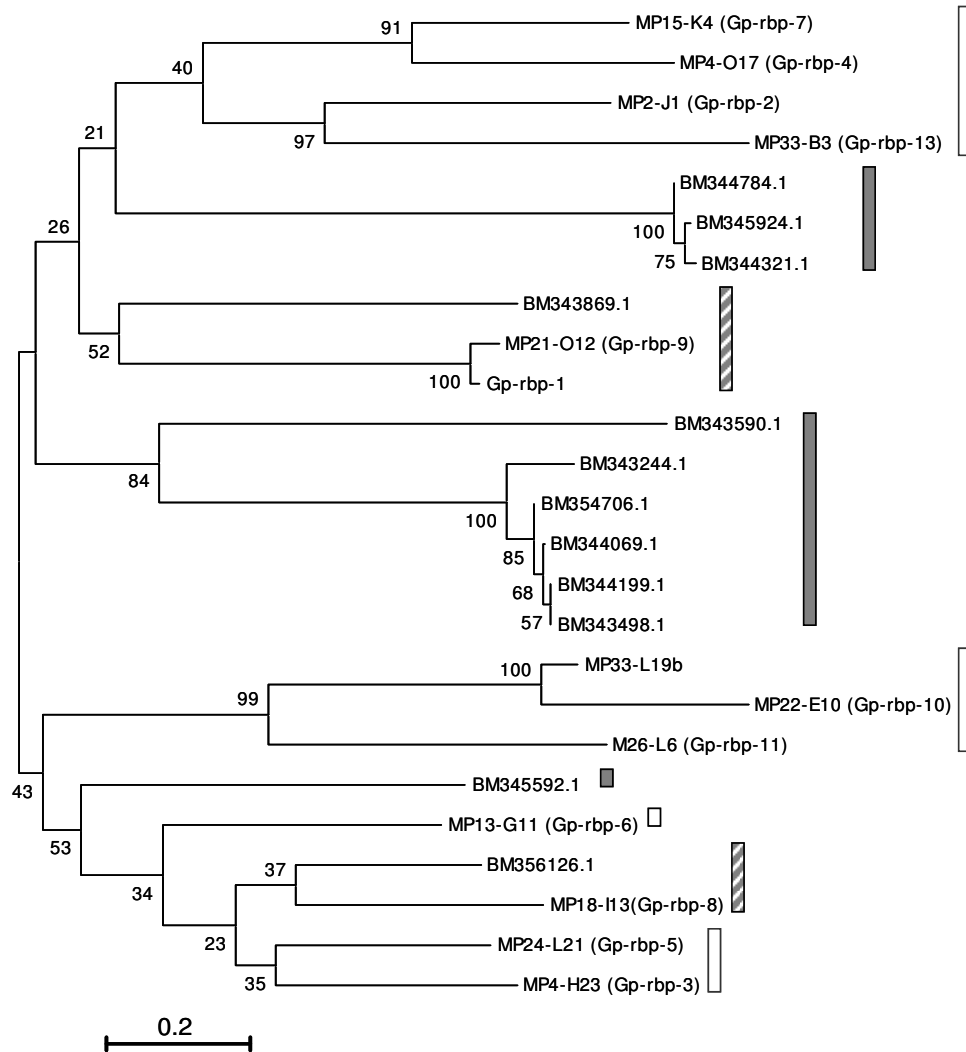


Figure 2: Clustering of the RanBPM-like gene copies identify to date in *G. pallida* and *G. rostochiensis*. The tree was generated using Neighbour Joining algorithm in pairwise deletion with Mega software. White bars represent groups in which only *G. pallida* copies were found, grey bars groups composed of only *G. rostochiensis* copies, grey and white bars represent groups in which both *G. pallida* and *G. rostochiensis* copies were found.

The selected sequences were translated in amino acids and aligned using T-coffee. The nucleotidic alignment was obtained using RevTrans. The genetic distances calculated using the

Kimura 2 parameters models indicated that genetic variability is 2 times higher inside the Peruvian populations (0.069) than inside the European populations (0.031). These data are in agreement with the high levels of variability described by Picard (2005) in the Peruvian populations of this nematode.

Table 2: Populations and species used to study the polymorphism and the evolutionary history of the *rbp-1* gene.

	Population name	Geographic origin	Code
<b>Genus <i>Globodera</i></b>			
<b><i>G. pallida</i> from Europe</b>	Luffness (Pa2/3)	Scotland	GPE1
	Ouessant (Pa2/3)	France	GPE2
	Chavornay (Pa2/3)	Switzerland	GPE3
<b><i>G. pallida</i> from South America</b>	Amantani	South Peru	GPS3
	Cutzco	Center Peru	GPS5
	Andahuaylas	Central Peru	GPS6
	Huancayo	North Peru	GPS7
	Chocon (P4A)	Peru	GPS8
	Otuzco (P5A)	Peru	GPS9
	Huamacucho (P6A)	Peru	GPS10
<b><i>G. "mexicana"</i></b>	Popocatepetl	Mexico	GM1

The repartition of the variable sites was studied in the data set. Out of the 13.7% of variable sites, 33.6% were on the first codon position, 30.4% on the second position and 34.4% on the third position. This means that approximately 2/3 of the mutations correspond to non synonymous substitutions that will modify the amino acids composition of the protein encoded by this gene. If we compare with similar data obtained on a housekeeping gene, the elongation factor 1 $\alpha$ , only 30% (at the *Globodera* species level) of the mutations were found on the first and second positions of the codons (Grenier com. pers.). The high rate of variability observed at the intra-specific level suggests that strong selection pressures act on the *rbp-1* gene to generate more diversity as it was suggested for other pathogenicity genes that displayed 20% of variable site between *G. pallida* and *G. "mexicana"* (Blanchard *et al.*, submitted).

Positive selection in a parasitism gene is therefore considered as an adaptation sign of the pathogen to its environment (Weber *et al.*, 2006). In a co-evolutionary context of a plant gene against a pathogen gene, we can assume that the *RBP-1* protein could interact with plant protein showing a high specificity in the interaction. Recent works showed that transient expression of *RBP-1* protein in *Nicotiana benthamiana* leaves also expressing the potato resistance gene *Gpa2*, results in a hypersensitive response. This suggests that *rbp-1* can also act as a determinant of the incompatible plant/nematode interaction (Moffet, pers. com.). In this interaction, both nematode *RanBPM* like genes and plant resistance genes have probably evolved to avoid the pathogen if we focused on the plant side, or to adapt to its plant if we look at the nematode side. Therefore, *RanBPM* like genes could present, as it was described for the *H. glycines* chorismate mutase genes (Bekal *et al.*, 2003; Lambert *et al.*, 2005), alleles that would display sequence polymorphism and

various expression in virulent/avirulent populations. This will be the next challenge in the studying of these genes.

## **ACKNOWLEDGMENTS**

This work was supported by a PAI Alliance in the frame of collaboration between INRA UMR BiO3P and the Scottish Crop Research Institute of Dundee. I would like to gratefully thank Patrick LHOMME for his contribution to this work.

## REFERENCES

- Bekal S, Niblack TL, Lambert KN.** 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Mol Plant-Microbe Interact.* 16:439-446.
- Blanchard A, Esquibet M, Fouville D, Grenier E.** 2005. Ranbpm homologue genes characterised in the cyst nematodes *Globodera pallida* and *Globodera 'mexicana'*. *Physiol Mol Plant Pathol.* 67:15-22.
- Blumenthal T.** 2005. Trans-splicing and operons. In *Wormbook* (eds. The *C. elegans* research community). doi/10.1895/worbook.1.5.1, <http://www.wormbook.org>.
- Corpet F.** 1988. Multiple sequence alignment in hierarchical clustering. *Nucleic Acids Res.* 16:10881-10890.
- Davis EL, Hussey RS, Baum TJ.** 2004. Getting to the roots of parasitism by nematodes. *Trends Parasitol.* 20:134-141.
- Gao BL, Kallen KD, Davis EL, Baum TJ, Hussey RS.** 2004. Developmental expression and biochemical properties of a beta-1,4-endoglucanase family in the soybean cyst nematode, *Heterodera glycines*. *Mol Plant Pathol.* 5:93-104.
- Hafizi S, Gustafsson A, Stenhoff J, Dahlback B.** 2005. The Ran binding protein RanBPM interacts with Axl and Sky receptor tyrosine kinases. *Int J Biochem Cell Biol.* 37:2344-2356.
- Heald R, Weis K.** 2000. Spindles get the Ran around. *Trends Cell Biol.* 10:1-4.
- Lambert KN, Bekal S, Domier LL, Niblack TL, Noel GR, Smyth CA.** 2005. Selection of *Heterodera glycines* chorismate mutase-1 alleles on nematode-resistant soybean. *Mol Plant-Microbe Interact.* 18:593-601.
- Masters SL, Yao S, Willson TA, Zhang JG, Palmer KR, Smith BJ, Babon JJ, Nicola NA, Norton RS, Nicholson SE.** 2006. The SPRY domain of SSB-2 adopts a novel fold that presents conserved Par-4-binding residues. *Nature Struct Mol Biol.* 13:77-84.
- Nakamura M, Masuda H, Horii J, Kuma KI, Yokoyama N, Ohba T, Nishitani H, Miyata T, Tanaka M, Nishimoto T.** 1998. When Overexpressed, a Novel Centrosomal Protein, RanBPM, Causes Ectopic Microtubule Nucleation Similar to gamma -Tubulin. *J Cell Biol.* 143:1041-1052.
- Picard D.** 2005. Génétique des populations et phylogéographie du nématode à kyste de la pomme de terre (*Globodera pallida*) au Pérou. INRA-Agrocampus Rennes. Thesis. p. 185.
- Poirot O, O'Toole E, Notredame C.** 2003. Tcoffee@igs: a web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res.* 31:3503-3506.

**Ponting C, Schultz J, Bork P.** 1997. SPRY domains in ryanodine receptors (Ca<sup>2+</sup>-release channels). *Trends Biochem Sci.* 22:193-194.

**Qin L, Overmars B, Helder J, Popeijus H, Rouppe van der Voort JN, Groenink W, Van Koert PH, Schots A, Bakker J, Smant G.** 2000. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. *Mol Plant-Microbe Interact.* 13:830-836.

**Umeda M, Nishitani H, Nishimoto T.** 2003. A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM. *Gene.* 303:47-54.

**Vanholme B, De Meutter J, Tytgat T, Van Montagu M, Coomans A, Gheysen G.** 2004. Secretions of plant-parasitic nematodes: a molecular update. *Gene.* 332:13-27.

**Wang D, Li Z, Messing EM, Wu G.** 2002. Activation of Ras/Erk pathway by a novel MET-interacting protein RanBPM. *J Biol Chem.* 277:36216-36222.

**Weber E, Koebnik R.** 2006. Positive selection of the Hrp pilin HrpE of the plant pathogen *Xanthomonas*. *J Bacteriology.* 188:1405-1410.

**Wernersson R, Pederson AG.** 2003. Revtrans- Constructing alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res.* 31:3537-3539.

**Yan Y, Smant G, Stokkermans J, Qin L, Helder J, Baum T, Schots A, Davis E.** 1998. Genomic organization of four [beta]-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. *Gene.* 220:61-70.

**Yan Y, Smant G, Davis E.** 2001. Functional screening yield a new beta-1,4-endoglucanase gene from *Heterodera glycines* that may be the product of recent gene duplication. *Mol Plant-Microbe Interact.* 14:63-71.



## CONCLUSION DE LA DEUXIEME PARTIE

Dans cette étude, nous sommes parti des transcrits candidats identifiés par SSH pour caractériser des nouveaux gènes impliqués dans le pouvoir pathogène des nématodes à kyste de la pomme de terre. Parmi les 3 candidats différentiels entre *G. pallida* et *G. "mexicana"*, deux semblent être liés au parasitisme: *la7* et *IVg9*. Malgré le fait qu'aucune homologie n'ait été identifiée dans les bases de données, régulièrement interrogées, certaines particularités font de ces gènes des candidats intéressants. C'est notamment aussi le cas du candidat *lc5 (rbp-1)* qui n'avait pas été identifié comme différentiel entre les deux espèces de nématodes mais pour lequel l'homologie identifiée dans les bases de données a fait de lui un candidat particulièrement intéressant.

Les données présentées dans cette première partie montrent que les 3 gènes sélectionnés présentent tous des caractéristiques indiquant leur implication dans les interactions qui régissent le dialogue moléculaire entre le nématode et sa plante hôte. Il faudrait maintenant développer des méthodes d'étude de la fonction de ces gènes. La complexité du modèle biologique rend l'accès aux protéines impliquées relativement difficile. Nous n'avons notamment pas la possibilité de générer des mutants chez les nématodes à kyste à l'image de ce qui est réalisé sur le nématode modèle *C. elegans*. De plus, les données génomiques, transcriptomiques, et protéomiques issues du séquençage complet du génome de ce nématode ne sont pas transposables à notre modèle phytoparasite notamment du fait de leur éloignement phylogénétique qui rend les comparaisons difficiles et également du fait de leurs modes de vie différents (libre *versus* parasite). Il est donc nécessaire de mettre au point des techniques d'étude particulières telles le RNA interférence (RNAi) pour identifier le rôle des gènes du parasitisme.

Si le rôle putatif des RanBPM est partiellement éclairci par les données que nous avons obtenues, il faudrait valider les hypothèses que nous avons émises par des techniques permettant d'identifier clairement la fonction de cette protéine. La fonction des gènes du pouvoir pathogène peut être recherchée notamment par des approches de type RNAi qui permettent d'identifier les conséquences de l'absence d'expression du gène cible sur la physiologie de la plante ou sur la multiplication du nématode (nombre de femelles formées). Si cette technique semble être très prometteuse, elle n'est pour le moment qu'au stade de mise au point chez les nématodes à kyste. De plus, cette technique nécessite de connaître la famille multigénique du gène ciblé si l'on ne souhaite viser qu'une copie en particulier (Chen *et al.*, 2005). Enfin, dans le cas des RanBPM, s'il serait relativement simple d'observer le nombre de femelles qui se développent lorsque le gène est éteint, il serait beaucoup plus ardu de s'attacher à observer les conséquences au niveau cellulaire puisque cette protéine semble être active au niveau du site nourricier. Il faudrait envisager alors de faire des coupes de syncytium. Une méthode alternative pour avoir une meilleure idée de l'implication des RanBPM dans le parasitisme serait de localiser la protéine au niveau cellulaire, dans la plante. Ceci permettrait de confirmer l'hypothèse d'une implication des RanBPM de nématode dans les dérégulations du cycle cellulaire végétal au niveau des cellules syncytiales. Ces données pourraient être complétées par une approche quantitative d'expression du gène *rbp-1* à

## **ENCADRE N° 5 : INFORMATIONS COMPLEMENTAIRES SUR LES GENES *la7* ET *IVg9***

### **Le gène *la7***

L'article 1 montre la présence d'un domaine "ShK toxine" (*Stichodactyla helianthus*, Castaneda *et al.*, 1995) dans la séquence du gène *la7*. C'est un domaine initialement identifié chez une métridine, une toxine d'anémone de mer brune (*Metridium senile*). Ce domaine ShK dont la fonction est inconnue est identifié chez de nombreuses protéines de *Caenorhabditis elegans* dont certaines sont des metallopeptidases, dont les fonctions biologiques sont variées. Ce domaine chez *C. elegans* est constitué de 30 acides aminés et est riche en résidus cystéine. Chez *G. pallida* et *G. "mexicana"* ce domaine contient 35 acides aminés dont 6 cystéines. La toxine ShK fait partie de la famille des inhibiteurs de canaux potassiques chez les anémones de mer et s'avère avoir un rôle immunosuppresseur en bloquant certains canaux potassiques impliqués dans la prolifération des lymphocytes T et la production de lymphokine chez le rat (Tudor *et al.*, 1998). Si cette toxine affecte le système immunitaire des mammifères, on peut imaginer que ce type de protéine pourrait avoir un effet similaire chez les plantes de suppression des réactions de défenses mises en place par la plante hôte lors d'une infection par les nématodes à kyste. Ceci pourrait expliquer le fait que nous n'ayons pas mis en évidence de relation entre les variations de séquence de ce gène et les différences de gammes d'hôtes présentées par les lignées hybrides (voir article 1), puisque ce mécanisme pourrait être alors généralisé à tous les couples nématode endoparasite/plante.

### **Le gène *IVg9***

Même si aucune homologie n'a encore été identifiée pour ce gène, il présente un intérêt particulier du fait que ce soit le seul gène identifié dont le transcrit est surexprimé chez *G. "mexicana"*. De plus, comme décrit dans l'article 1, ce gène présente des différences de structure entre *G. pallida* et *G. "mexicana"*. Etant donné que les hybrides issus du croisement entre ces deux nématodes ont tous en commun le développement sur pomme de terre, on peut toujours supposer que les différences de séquences observées pour la copie du gène *IVg9* de *G. "mexicana"* puisse être responsable du fait que cette espèce ne se développe pas sur pomme de terre. Une stratégie pour tester cette hypothèse pourrait être de générer des hybrides issus de nouveaux croisements *G. pallida* x *G. "mexicana"* où la capacité à se développer sur morelle serait cette fois toujours conservée au détriment parfois d'un développement sur pomme de terre.

différents stades larvaires de manière à identifier les stades où ce gène est préférentiellement exprimé et donc nécessaire au nématode.

Concernant les gènes *la7* et *IVg9*, nous avons choisi une démarche originale d'étude de la variabilité de ces gènes au sein de lignées hybrides dont le pouvoir pathogène est variable. Cette démarche est pour autant relativement logique du fait que les deux gènes d'intérêt avaient été identifiés chez *G. pallida* et *G. "mexicana"*. Nous avons donc voulu tester l'hypothèse d'une relation entre les mutations accumulées sur ces gènes et les différences de pathogénicité observées au sein des hybrides.

Enfin, les approches très fonctionnelles ont beaucoup à gagner du développement d'approches évolutives. En effet, la variabilité d'un gène au sein d'une population, d'une espèce ou d'un genre est le reflet des pressions de sélection auxquels ils ont été soumis au cours de l'évolution et doit donc pouvoir nous renseigner également sur les types d'interaction et éventuellement sur les variations responsables des différences de gamme d'hôtes, des pathotypes ou encore du statut virulent ou avirulent des individus. Néanmoins, à ce jour aucune étude n'a été menée sur la variabilité et l'évolution des gènes pour lesquels nous avons davantage d'informations nous permettant de leur attribuer une fonction plus certaine dans le pouvoir pathogène. De plus, cette approche novatrice devrait nous permettre non pas de nous focaliser sur un dialogue précis entre une population et une plante hôte, mais de nous donner une vision plus globale des interactions entre les nématodes à kyste et leur plante. Cette approche semble être d'un intérêt particulier dans le contexte de gestion durable des sources de résistances au champ en permettant une prédiction du comportement des populations confrontées à certaines résistances ou plantes.

Dans la partie suivante, nous allons justement aborder les études que nous avons menées sur la variabilité de gènes du pouvoir pathogène au sein de la famille des Heteroderidae en choisissant non pas les candidats de cette première partie mais d'autres gènes décrits dans la bibliographie dont l'implication au niveau des interactions plante/nématode était plus certaine.

# Résultats

Variabilité de gènes du pouvoir  
pathogène des nématodes à kyste  
et évolution moléculaire de ces  
gènes

*Partie III*

## PRESENTATION DE LA PARTIE III

La troisième partie du manuscrit s'intègre dans une vision évolutive de l'interaction plante/nématode qui passe par une connaissance de la diversité des populations et espèces de nématodes d'un point de vue phylogénétique. L'étude de la variabilité des gènes du pouvoir pathogène et des pressions de sélection qu'ils ont subit au cours de l'évolution devrait permettre de retracer l'histoire évolutive de ces gènes. Les résultats de cette troisième partie ont fait l'objet d'une communication orale au congrès européen de nématologie en Bulgarie (2006) et d'une communication affichée accompagnée d'une courte présentation orale au congrès de biologie évolutive à Marseille (2006). Cette partie est divisée en deux chapitres :

- Le premier chapitre décrit la caractérisation de deux gènes du pouvoir pathogène (pectate lyase, cathepsine L) et d'un gène de ménage (facteur d'élongation 1 $\alpha$ ) en terme de structure et de polymorphisme de séquence pour 40 populations et espèces de la famille des Heteroderidae. Nous avons ainsi pu reconstruire les relations phylogénétiques des populations et espèces étudiées sur la base de ces séquences de gènes nucléaires liés ou non au parasitisme. Une analyse descriptive ainsi qu'une évaluation de l'utilisation des séquences de ces gènes en phylogénie est présentée sous forme d'article qui pourrait être soumis à *Molecular Biology and Evolution*. Enfin, nous développerons les différentes méthodes d'analyse qui nous ont permis de générer des données phylogénétiques fiables comme prérequis aux analyses des pressions de sélection auxquelles sont soumis ces gènes. En effet, l'évaluation de la sélection ne prend son sens et ne peut être interprétée que dans un cadre phylogénétique clairement établi.

- Le second chapitre présente les toutes dernières investigations menées sur l'évolution des séquences des gènes du parasitisme à différentes échelles (intra-spécifique, intra- et inter-genres) et le long des séquences en comparaison avec l'évolution du facteur d'élongation pour 40 populations de nématodes à kyste. Nous nous focaliserons sur les outils et les analyses utilisées pour identifier les types de pression de sélection qui régissent l'histoire évolutive des différents gènes du pouvoir pathogène en comparaison avec le gène de ménage.



## **PARTIE III :**

# **POLYMORPHISME ET EVOLUTION MOLECULAIRES DE GENES DU POUVOIR PATHOGENE**

### **INTRODUCTION**

La question centrale de cette troisième partie est la suivante :

#### **Comment ont évolué les gènes du pouvoir pathogène chez les nématodes à kyste ?**

Nous avons proposé d'apporter des éléments de réponse à cette question en améliorant les connaissances sur la variabilité et sur les pressions de sélection auxquelles ont été soumis les gènes qui gouvernent le parasitisme des plantes chez les nématodes. Pourquoi ? :

- 1- pour mieux comprendre les bases moléculaires de la diversité en terme de pathogénicité,
- 2- pour identifier le potentiel évolutif des gènes du pouvoir pathogène,
- 3- pour évaluer l'intérêt de développer de nouvelles résistances basées sur l'inhibition ou le blocage de protéines clés du parasitisme. Mais également pour mieux utiliser les résistances naturelles disponibles lorsque les protéines de nématodes et de plantes qui interviennent dans l'interaction seront connues.

Pour répondre à ces questions, il a fallu choisir les gènes les plus pertinents pour initier ce type d'approche, et également définir les genres, espèces et populations à inclure dans cette étude afin d'avoir la vision la plus large possible des modes évolutifs des gènes du pouvoir pathogène. Nous allons donc dans un premier temps détailler quelques éléments de bibliographie qui nous ont été nécessaires dans les choix des populations et des gènes étudiés.





# CHAPITRE 1

---

# CHAPITRE 1 :

## 1. ELEMENTS DE CONTEXTE : CHOIX DES GENES ET DES ESPECES

### 1.1. Choix des populations et des espèces au sein de la famille des Heteroderidae

Les populations ont été choisies de manière à représenter la variabilité existant au sein de la famille des Heteroderidae, tout en mettant l'accent sur le genre *Globodera* (28 populations) et encore plus particulièrement sur l'espèce *G. pallida* (15 populations) (Fig. 1). Le détail des populations étudiées est présenté dans l'article 4 (Tableau 1).

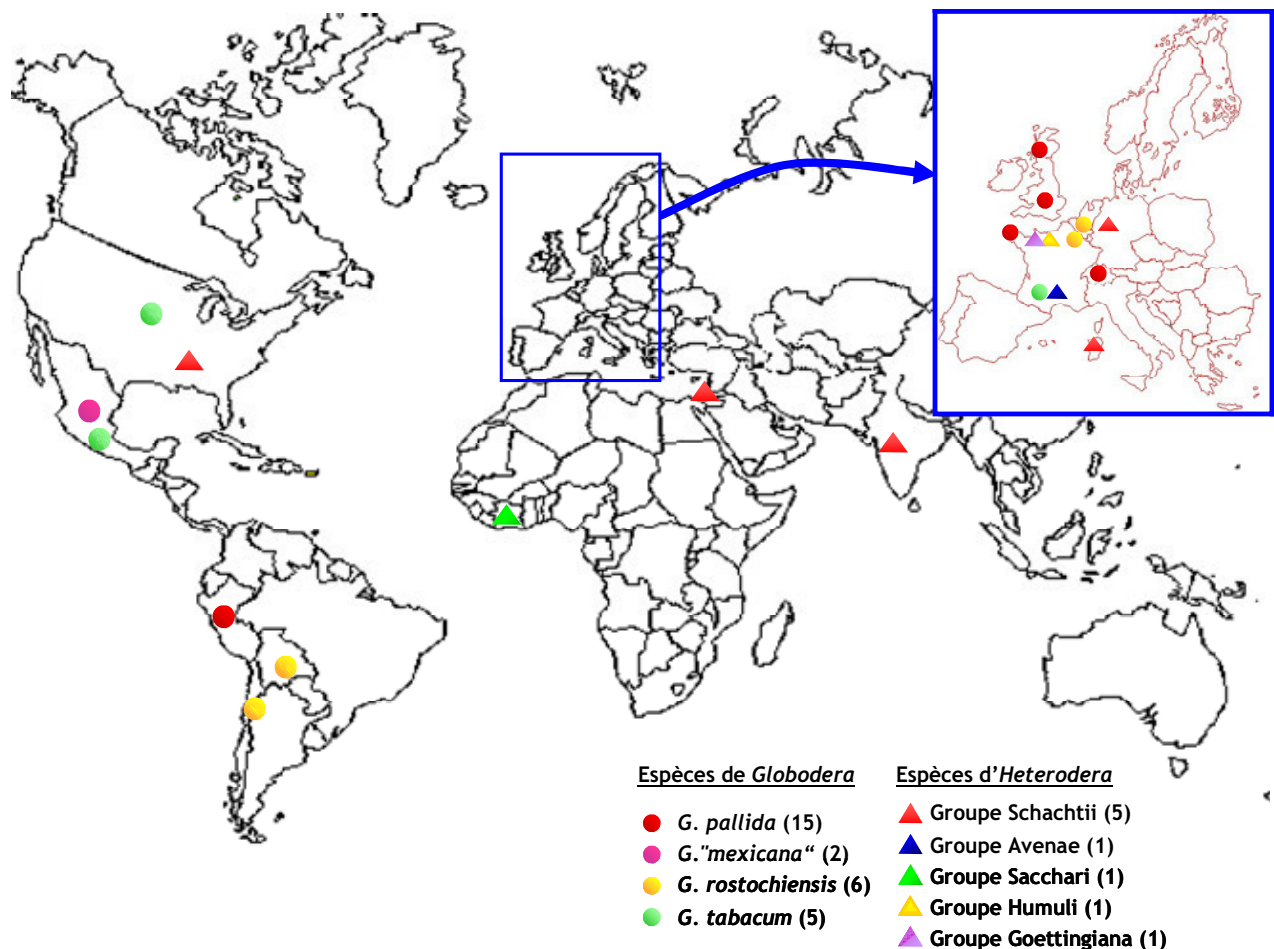


Figure 1 : Origine géographique des populations de nématodes des genres *Globodera* et *Heterodera* choisies pour l'étude du polymorphisme des gènes du pouvoir pathogène.

Nous avons utilisé plusieurs critères de choix des populations et espèces sur lesquelles nous voulions travailler :

- différences d'hôtes : nous avons sélectionné 4 espèces de *Globodera* présentant différentes gammes de plantes hôtes de la famille des Solanacées. Nous avons également ajouté des espèces d'*Heterodera* qui parasitent d'autres familles botaniques.
- diversité génétique : nous avons choisi un représentant de chaque groupe d'*Heterodera* définis par Subbotin *et al.* (2001). Nous avons également choisi des populations indigènes et importées de *Globodera* avec dans le cas de *G. pallida*, des représentants des 5 clades définis par Picard (2005) (Fig. 2). Pour l'espèce *G. tabacum*, nous avons choisis des représentants des différentes sous espèces définies par Marché *et al.*, (2001) comme relativement divergentes et également ayant des gammes d'hôtes différentes.
- différences de pouvoir pathogène : ceci se traduit chez les nématodes du genre *Globodera* par des différences de pathotypes (Kort *et al.*, 1977). Nous avons donc choisi pour l'espèce *G. pallida*, des représentants de tous les pathotypes et chez *G. rostochiensis*, les pathotypes dont nous disposons. Chez les *Heterodera*, nous avons à disposition des populations virulentes et avirulentes contre le gène de résistance *HS1<sup>pro1</sup>*.

Au total, nous disposons ainsi de 40 populations pour étudier la variabilité des gènes du pouvoir pathogène au sein de la famille des Heteroderidae.

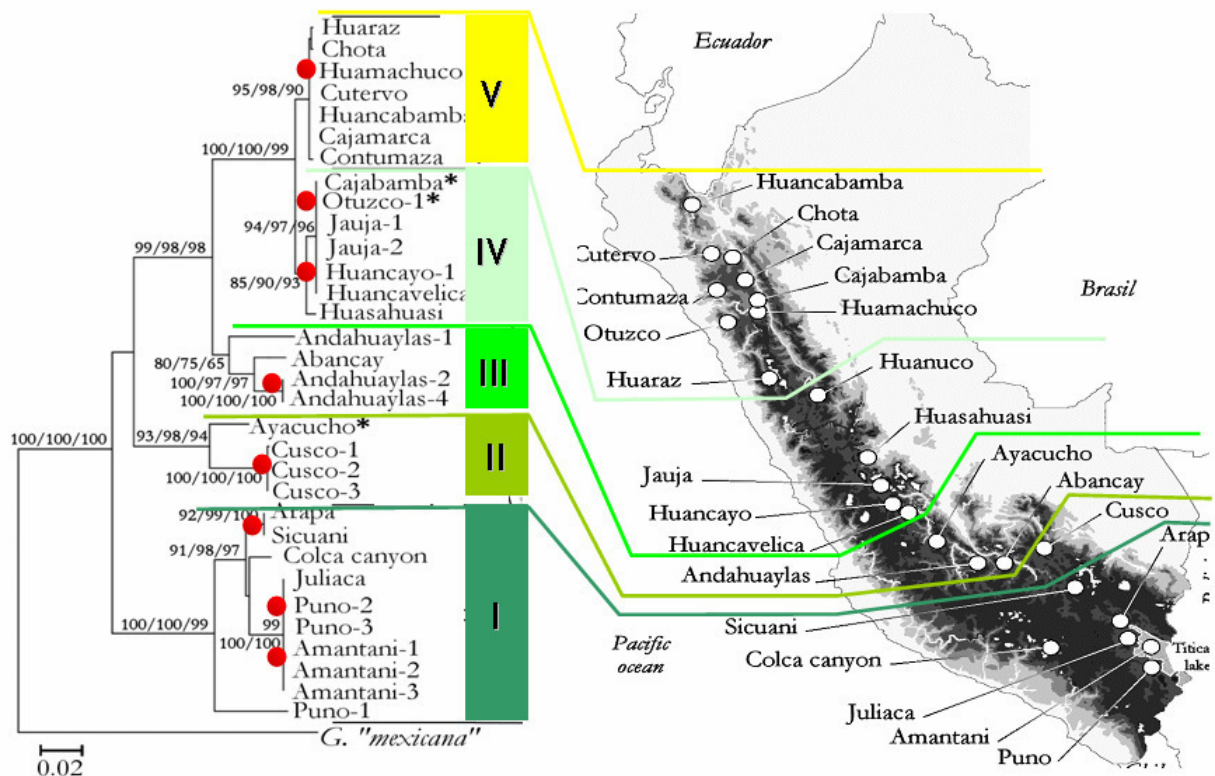


Figure 2 : Phylogéographie des populations de *G. pallida* originaires du Pérou, basée sur les séquences mitochondriales (cytochrome b) (Picard, 2005). Les clades I à V regroupent les populations d'une même origine géographique le long du transect Nord-Sud du Pérou. Les points rouges symbolisent les populations choisies pour les études du polymorphisme des gènes du pouvoir pathogène.

## 1.2. Choix des gènes du pouvoir pathogène

Comme mentionné dans l'introduction bibliographique, peu de gènes isolés chez les nématodes à kyste ont une fonction précise déjà démontrée. Nous nous sommes appuyés sur les données disponibles dans les bases de données et sur les connaissances que nous avons acquises sur les gènes *la7*, *IVg9* et *rbp-1* pour choisir les gènes les plus appropriés. Ces données nous ont permis d'évaluer la faisabilité de l'étude sur ces gènes : appartenance à une famille multigénique, niveau de variabilité entre les copies paralogues, amplification chez tous les Heteroderidae et amplification à partir d'un kyste ou d'une larve pour minimiser les besoins. Pour que les comparaisons des tempos d'évolution prennent leur sens, il fallait s'assurer de l'implication des gènes choisis dans le parasitisme. Ainsi, nous avons cherché à identifier des candidats impliqués à différentes phases du parasitisme avec le plus de certitudes possible. Un autre critère de choix était la taille de la famille de ces gènes. Idéalement, l'utilisation d'un gène mono-copie aurait été la solution la plus simple pour s'assurer de comparer des gènes orthologues, cependant, aucun gène de ce type impliqué dans le pouvoir pathogène n'a été isolé chez les nématodes phytoparasites. Pour s'assurer de comparer des copies orthologues, nous n'avons pas choisi de nous appuyer que sur le pourcentage d'homologie entre deux séquences, mais également sur la structure du gène en terme de position et de taille des introns le long de la séquence génomique. En effet, la famille des  $\beta$ -1,4-endoglucanases identifiée chez *H. glycines* (Gao *et al.*, 2004) présente des variations importantes du nombre et de la répartition des introns entre les copies paralogues. Nous nous sommes aussi basé sur l'identification du meilleur homologue dans les bases de données disponibles pour le nématode *C. elegans*. Une autre exigence des méthodes d'analyse des taux d'évolution est la taille de la séquence codante exploitable que nous avons fixée arbitrairement à un minimum de 300 pb.

Nous n'avons identifié aucun gène pour lequel nous possédions toutes les données ou qui remplissait tous les critères. Nous avons donc dû faire des compromis pour choisir les gènes qui s'en approchaient le plus.

Parmi les gènes impliqués dans les phases précoces d'invasion et de migration, étaient disponibles les données sur les cellulases ou les pectate lyase. Le choix s'est ensuite fait sur la base de difficultés techniques rencontrées lors des premiers tests. Les essais d'amplification que nous avons réalisés pour travailler sur les cellulases ont montré que nous amplifions plusieurs copies à la fois par la présence de plusieurs bandes en gel d'agarose chez les *Heterodera*. Aucune amplification n'a été générée chez les *Globodera*. De plus, nous ne sommes parvenus à générer ces amplifications que sur des quantités importantes d'ADN d'*Heterodera*. Or, l'objectif de ces premiers tests était 1- d'amplifier le gène d'intérêt à partir du contenu larvaire d'un seul kyste, qui contient des individus frères, et 2- d'amplifier le gène d'intérêt avec les mêmes couples d'amorces sur toutes les espèces choisies afin de favoriser l'amplification de copies orthologues du gène. Nous avons donc choisi la pectate lyase qui remplissait tous ces critères techniques, dont la famille semble être plus réduite que celle des cellulases. Nous aborderons un peu plus loin, les détails concernant ces enzymes de dégradation des parois des cellules végétales.

Nous avons également cherché à étudier des gènes du pouvoir pathogène impliqués à différentes phases parasitaires et/ou sécrétés par différents organes. C'est pourquoi nous nous sommes intéressés aux chorismates mutases, aux cathepsines L ainsi qu'aux RanBPM. Les données

bibliographiques mentionnant plusieurs copies paralogues de chorismate mutase et également l'existence de différents allèles d'une même copie (Bekal *et al.*, 2003), nous n'avons pas choisi ce gène du fait des difficultés que nous aurions certainement rencontrées pour identifier des copies orthologues. Quant aux RanBPM, nous n'avions au début de notre étude que peu d'informations sur ces gènes (*rbp-1*), nous nous sommes donc finalement focalisé sur la cathepsine L pour laquelle nous disposons de plus d'informations sur son rôle potentiel, sa localisation et sur les séquences chez *Meloidogyne incognita*. Ce gène est exprimé dans un compartiment cellulaire différent (intestin) de celui des pectate lyase (glandes salivaires ventrales). De plus ce gène a été identifié comme différentiellement exprimé entre des individus virulents et avirulents de *M. incognita*, contre le gène de résistance *Mi* de la tomate (Neveu *et al.*, 2003).

Enfin, en comparatif, nous avons besoin d'étudier un gène non soumis à des pressions de sélection liées au parasitisme. La globine et le facteur d'élongation  $1\alpha$  sont couramment décrits comme des gènes de ménage, et utilisés comme marqueurs en phylogénie (Suarez *et al.*, 2006 ; Goetze *et al.*, 2006 ; Shingles *et al.*, 2006 ; Rokas *et al.*, 2002). Nous avons retenu le facteur d'élongation qui s'est avérée amplifiable à partir d'une seule larve de nématode. Ce gène de ménage a été choisi comme référent tout au long de l'étude sur la variabilité et l'évolution des gènes.

En outre, les phylogénies réalisées chez les nématodes sont la plupart du temps basées sur des régions d'ADN ribosomique (Blaxter *et al.*, 1998 ; Subbotin *et al.*, 2001 ; Holterman *et al.*, 2006) ou mitochondrial (Picard, 2005), mais rarement sur des gènes de ménage ou des gènes du pouvoir pathogène. Nous avons donc évalué l'utilité du facteur d'élongation  $1\alpha$  pour résoudre les relations phylogénétiques en comparaison avec les gènes nucléaires liés au parasitisme (pectate lyase et cathepsine L).

### **A. Les pectates lyases**

Les parois des cellules végétales sont constituées de deux composants majeurs, la cellulose et la pectine, organisés de façon très structurée (Fig. 3). Elles constituent une première barrière physique contre toute intrusion d'organismes phytopathogènes. La pectine, qui est une forme hautement méthylée de la pectate, est dégradée par les enzymes de type pectinases. Les pathogènes de plantes, que ce soit des champignons, des bactéries, des virus ou des nématodes, ont mis en place des "outils" leur permettant de dégrader la paroi des cellules végétales afin de s'introduire dans la plante. De nombreuses enzymes de dégradation des parois existent chez les organismes phytoparasites, le plus souvent des cellulases ou des pectinases. Parmi les pectinases, différents types de protéines ont été identifiées : les pectine méthyl-estérases, les pectines acétyl estérases, les polygalacturonases, les pectines lyases ou encore les pectates lyases. Ces dernières ont la capacité de dépolymériser les composants pectiques de la paroi primaire et de la lamelle moyenne des cellules de la plante hôte par clivage des ponts glycosyliques internes de la pectate ou des régions de la pectine qui ne sont pas hautement méthylées (Fig. 3 et 4).

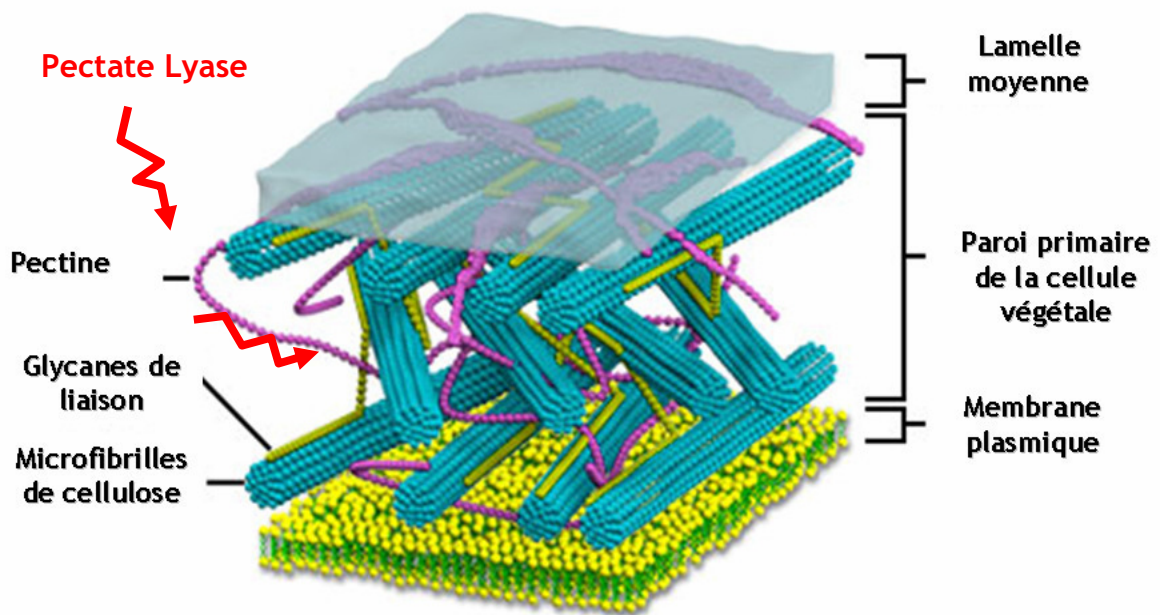


Figure 3 : Organisation structurale de la paroi des cellules végétales (<http://micro.magnet.fsu.edu/cells/plants/cellwall.html>). Les flèches rouges symbolisent l'action des pectates lyases sur la paroi végétale.

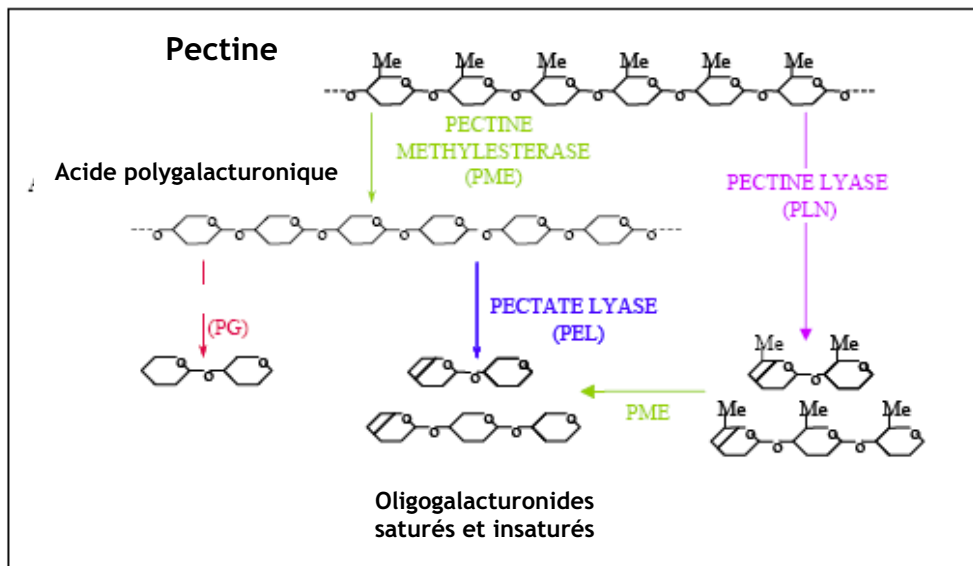


Figure 4 : Catabolisme des pectines

Certaines pectates lyases ont été identifiées comme importants facteurs de virulence chez plusieurs organismes phytopathogènes pour permettre la pénétration et la colonisation de l'hôte (Barras *et al.*, 1994, Rogers *et al.*, 2000, Huang *et al.*, 2005). Les pectates lyases sont réparties en cinq classes définies à partir des homologies identifiées avec des protéines déjà connues et parfois en fonction de leur structure tridimensionnelle :

- **Classe I** : correspond à la superfamille des pectates lyases regroupant les protéines bien décrites de bactéries (Pel A, B, C, D et E décrites chez *Erwinia chrysantemi*) et certaines protéines de plante.
- **Classe II** : rassemble les protéines de type pectate lyase périplasmique d'*Erwinia carotovora*.
- **Classe III** : regroupe certaines pectates lyases identifiées chez *Erwinia carotovora* (Pel B, Pel-3) ou chez le champignon *Fusarium solani*. Les pectates lyases de cette classe possèdent 4 régions conservées et de nombreux résidus cystéine caractéristiques (Shevchik *et al.*, 1997).
- **Classe IV et V** : chacune de ces classes regroupe des enzymes spécifiques identifiées chez *E. chrysanthemi*.

Les pectinases de classe III ont été mises en évidence pour la première fois chez des second stades larvaires (J2) de *G. rostochiensis* (Popeijus *et al.*, 2000). Le gène *pel-1* code une pectate lyase de 28 KD possédant un peptide signal indiquant la sécrétion probable de la protéine et est exprimé dans les glandes subventrales (localisation cellulaire des transcrits) des nématodes au stade infestant (J2). Ces enzymes ont été identifiées chez d'autres nématodes à kyste (*H. glycines*, De Boer *et al.*, 2002) mais aussi chez les nématodes à galle. Elles présentent une très grande diversité chez les nématodes phytoparasites. En effet, les trois copies isolées chez *Meloidogyne incognita* sont très différentes entre elles (Huang *et al.*, 2003 ; Huang *et al.*, 2005) mais également différentes des copies caractérisées chez les nématodes à kyste (Fig. 5). L'interrogation des bases de données a montré une homologie avec les pectate lyase de classe III d'origine fongique ou bactérienne suggérant un potentiel transfert horizontal de ce gène au cours de l'évolution des nématodes.

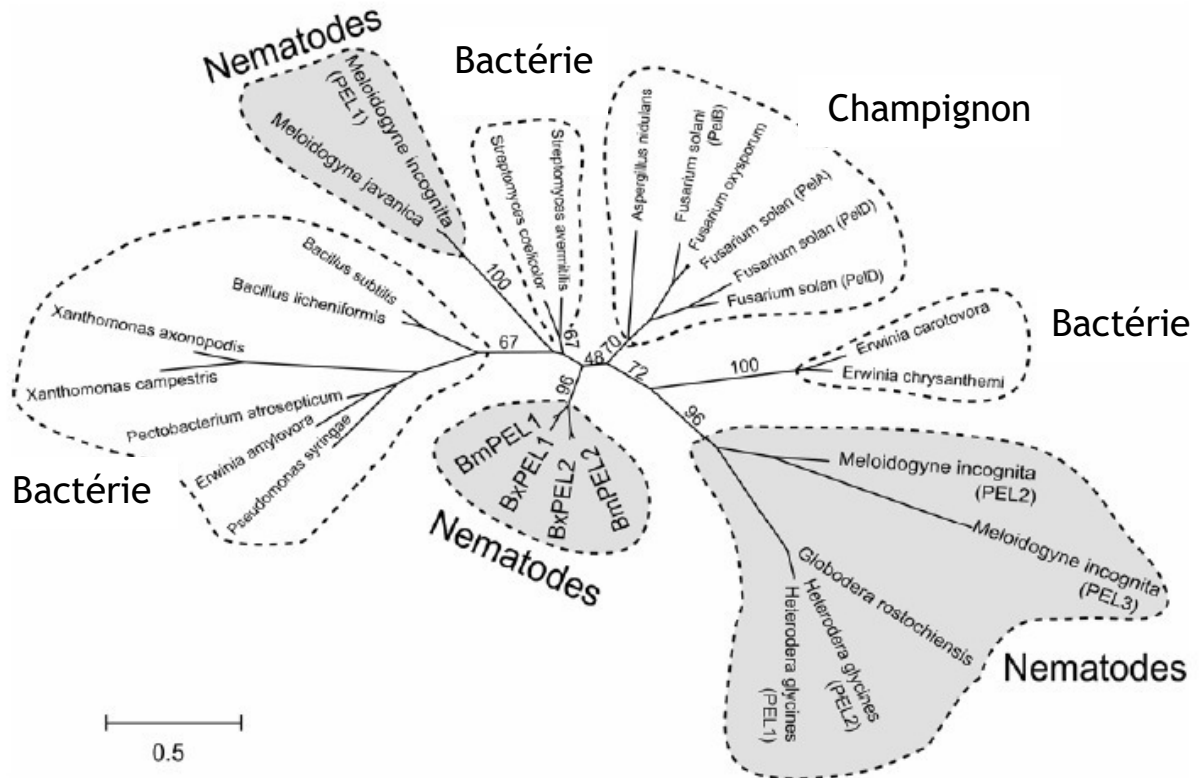


Figure 5 : Arbre phylogénétique non enraciné de protéines polysaccharides lyases de classe III généré en analyse de maximum de vraisemblance (Kikuchi *et al.*, 2006).

Quelles soient d'origine fongique, bactérienne ou de nématode, les pectates lyases de classe III possèdent une structure protéique commune comportant quatre régions signatures (Shevchik *et al.*, 1997) (Fig. 6) ainsi qu'un signal de sécrétion pour les pectates lyases de nématodes. Aucune fonction particulière n'a été attribuée à ces quatre régions. Une autre caractéristique des pectates lyases de classe III est la présence de nombreux résidus cystéine, de l'ordre de 10 à 14, dont la position est très conservée chez *Erwinia carotovora* (Shevchik *et al.*, 1997). Ces résidus cystéine semblent être impliqués dans la formation de ponts disulfures responsables de l'importante rigidité de ces protéines et sont retrouvés chez les pectate lyases de nématodes.

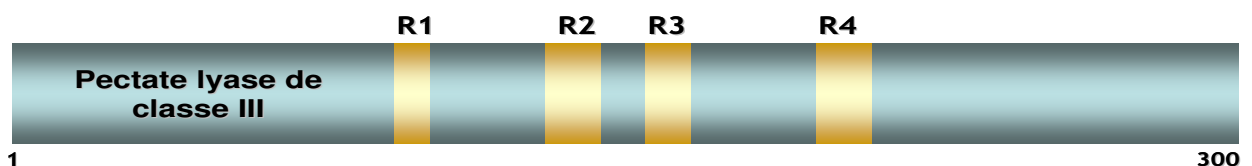


Figure 6 : Structure générale d'une pectate lyase de classe III, d'après Huang *et al.*, 2005. La taille de la séquence est indiquée en acides aminés. Les régions "signatures", R1 à R4, sont les régions définies par Schevchik *et al.*, 1997.



Les copies de pectate lyase identifiées chez les nématodes sont très divergentes en terme de :

- 1- nombre d'introns : celui-ci est variable entre les deux copies identifiées chez *M. incognita* présentant 2 petits introns de moins de 200 pb pour *Mi-pel-1* et un grand de 420 pb, contre deux grands pour *Mi-pel-2* de 577 et 572 pb,
- 2- taille des séquences codantes : 813 pb pour *Mi-pel-1* contre 963 pb pour *Mi-pel-2*,
- 3- d'homologies protéiques : seulement 49% de similarité entre MI-PEL-1 et MI-PEL-2 (Huang *et al.*, 2005). Ces deux gènes ne présentent également que peu de similarité avec les pectates lyases identifiées chez d'autres nématodes : 55% entre MI-PEL-1 et GR-PEL-1 (*G. rostochiensis*). Des analyses génomiques par Southern blot chez *M. javanica* et *M. incognita* ont montré que les pectates lyases identifiées chez les nématodes à galle appartiennent probablement à une petite famille multigénique (Doyle & Lambert, 2002 ; Huang *et al.*, 2005).

La présence de plusieurs copies très différentes du gène pectate lyase chez les nématodes suggère que malgré l'apparente fonction commune de toutes les copies de pectate lyase de classe III ainsi que leurs structures protéiques proches, il pourrait exister une spécificité des copies utilisées chez les deux familles de nématodes. Est-ce en fonction du mode de pénétration et de migration, en fonction des plantes hôtes ? Est-ce une arme supplémentaire pour le nématode qui pourrait exprimer une copie plutôt qu'une autre selon les conditions de parasitisme et/ou selon les plantes hôtes ? Même si nous n'avons pas de données complètes disponibles sur la taille des familles multigéniques de ces pectates lyases chez les nématodes, les études sur la variabilité des gènes connus codant les pectates lyases au sein des familles de nématodes à kyste et à galle devraient permettre d'apporter des éléments de réponse à ces questions.

### **B. Les cathepsines L**

Il existe différentes classes de protéinases, toutes présentes chez les nématodes parasites : les protéases à sérine, les aspartiques, les metalloprotéases et enfin les protéases à cystéine (ou cystéine protéases) sur lesquelles nous allons nous focaliser.

La super famille des cystéines protéases est divisée en quatre clans en fonction des homologies de séquence identifiées (Fig. 7). La majorité des cystéines protéases appartiennent à la famille des protéines papain-like (C1) du clan CA. Cette famille regroupe trois sous familles dont les cathepsine L qui sont comme les autres cathepsines (B, C, H, K, et S) localisées soit dans les vésicules cellulaires soit directement sécrétées (Sajid & McKerrow, 2002). Dans ce dernier cas, un peptide signal est identifié dans la séquence nucléotidique. C'est le cas de la cathepsine L identifiée chez *Meloidogyne incognita* qui présente pourtant une localisation intestinale des transcrits (Neveu *et al.*, 2003). Comment les protéines cathepsine L peuvent alors être sécrétées ou excrétées chez les nématodes phytoparasites ? Shingles *et al.* (2006) émettent l'hypothèse que le lumen serait perdu du pharynx à l'anus chez les femelles à maturité, ce qui expliquerait que les protéines localisées dans l'intestin ne seraient pas compartimentées mais libres dans la cavité du nématode qui serait alors en mesure de les excréter. De plus, cette protéine a été immunolocalisée dans les produits sur système E/S (Excretory/Secretory) des nématodes filaires (Guiliano *et al.*, 2004) suggérant une sécrétion des cathepsines L par ce biais.

Le clan CA des cystéines protéases est également caractérisé par la sensibilité des protéines qui le constitue aux inhibiteurs de protéase généraux, tel que le E64 (L-trans-epoxysuccinyl-leucyl-amido (4-guanidino) butane), faisant d'elles des cibles particulièrement intéressantes pour créer de nouvelles méthodes de lutte *via* ces inhibiteurs de protéases. Les cathepsines L ont des rôles très divers dans les interactions hôte/parasite (Sajid & Mckerrow, 2002). Chez certains trématodes tels que *Fasciola hepatica*, parasites d'animaux vertébrés, les cathepsines de type L sont sécrétées dans l'intestin et régurgitées pour dégrader les matrices protéiques constituant les cellules animales (Halton, 1997). Chez *Diplostomum pseudospathaceum*, il semblerait que ces cystéines protéases localisées dans les cercaires soient impliquées dans la pénétration du parasite au travers de la peau des oiseaux aquatiques (Moczon, 1994). Les cystéines protéases sont dans certains cas impliqués dans l'échappement aux défenses de l'hôte qu'il soit animal ou végétal. C'est le cas des *Trichomonas vaginalis* qui excrètent des cystéines protéases capables de cliver les immunoglobulines chez l'homme, ou des effecteurs du type AVRpt2 chez la bactérie phytoparasite *Pseudomonas syringae* qui sont impliqués dans les suppressions de réactions hypersensibles des plantes (Nomura *et al.*, 2005). Quel que soit le modèle étudié, les cathepsines L semblent jouer un rôle à différents niveaux de l'interaction, dans l'invasion des tissus hôtes, la nutrition, les mues larvaires ou encore l'échappement aux défenses mises en place par l'hôte. Cette protéase serait impliquée dans la digestion chez les nématode des genres *Meloidogyne* et *Globodera* (Shingles *et al.*, 2006 ; Lilley *et al.*, non publié). Chez les nématodes phytoparasites *Meloidogyne incognita*, la cathepsine L a été identifiée comme différentiellement exprimée chez les seconds stades larvaires de nématodes virulents et avirulents contre le gène de résistance *Mi* de la tomate (Neveu *et al.*, 2003).

Une très grande diversité de structure des gènes codant les cathepsines L est observée chez les organismes vivants, de 0 intron chez *Plasmodium falciparum* à 7 introns chez les mammifères (Hu *et al.*, 2006). Cependant la structure codante est quant à elle relativement bien conservée au sein du règne animal avec une structure en deux parties : une pro région et une deuxième zone correspondant à la protéine mature (Ménard *et al.*, 1998 ; Silva *et al.*, 2004) comme indiqué sur la figure 8A. D'une manière générale, ce type d'enzyme est synthétisé sous forme de pro enzyme et subit un clivage protéolytique puis un relargage du peptide pro domaine pour donner une protéine mature (Silva *et al.*, 2004). La structure tridimensionnelle de cette protéine a été déterminée (Fig. 8B) permettant d'identifier les zones de repliements de la protéine et de mieux évaluer l'importance des mutations observées entre les séquences de cathepsine L.

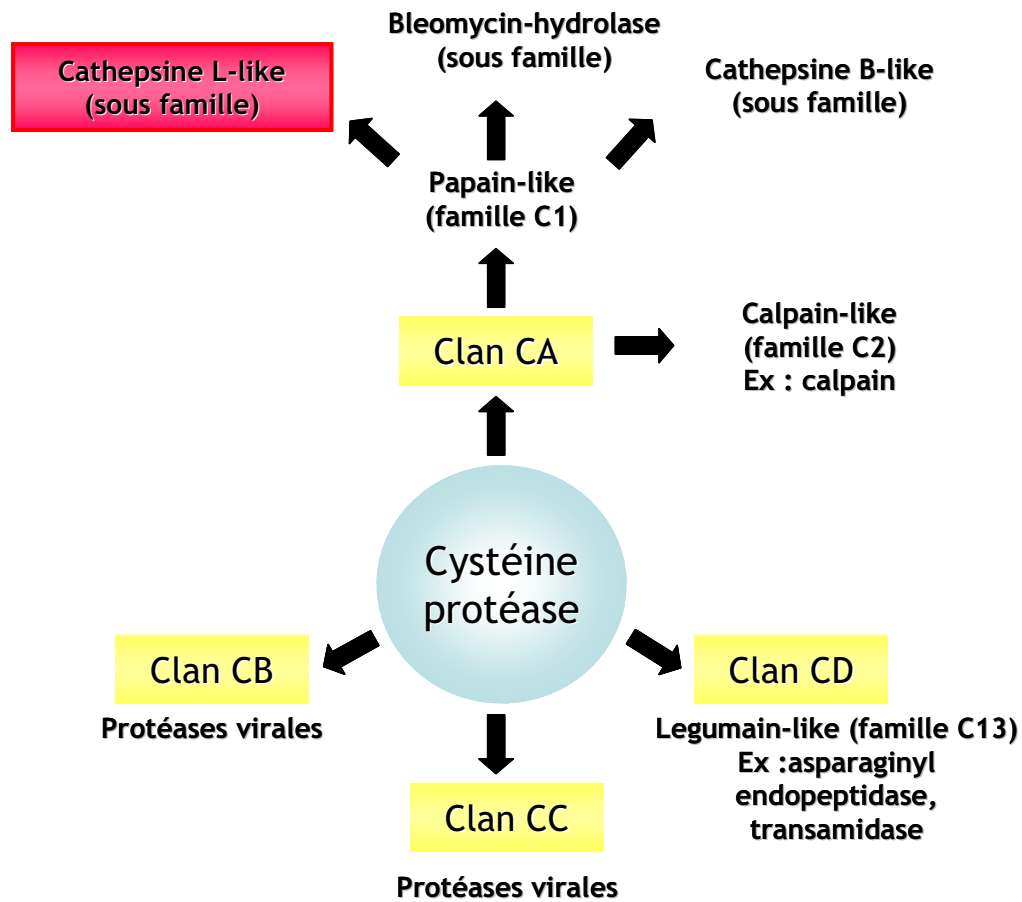


Figure 7 : Schéma de la superfamille des cystéines protéases. D'après Sajid & McKerrow, 2002. Les sous familles sont déterminées par homologie de séquence autour des résidus acides aminés du site actif.

Cette protéine a été identifiée chez les nématodes les plus problématiques d'un point de vue agronomique : *Heterodera glycines* (Lilley *et al.*, 1996), *Globodera pallida* (Koritas et Atkinson, 1994), *Meloidogyne hapla*, *M. javanica* et *M. incognita* (Michaud *et al.*, 1996, Neveu *et al.*, 2003). Le criblage d'une banque d'ADNc de J2 d'*H. glycines* n'a révélé qu'une copie de cathepsine L avec une ORF de 1122 pb codant une protéine de 374 acides aminés selon Silva *et al.* (2004). Mais deux types de transcrits ont été décrits par Urwin *et al.* (1997) dont les séquences protéiques déduites présentent 67% de similarité. Les analyses génomiques par Southern blot ont également montré qu'il existe probablement plus de 2 copies de cathepsine L chez les nématodes à kyste du soja (Urwin *et al.*, 1997). Comme dans le cas des pectates lyases, les cathepsines L de nématodes semblent appartenir à une petite famille multigénique.

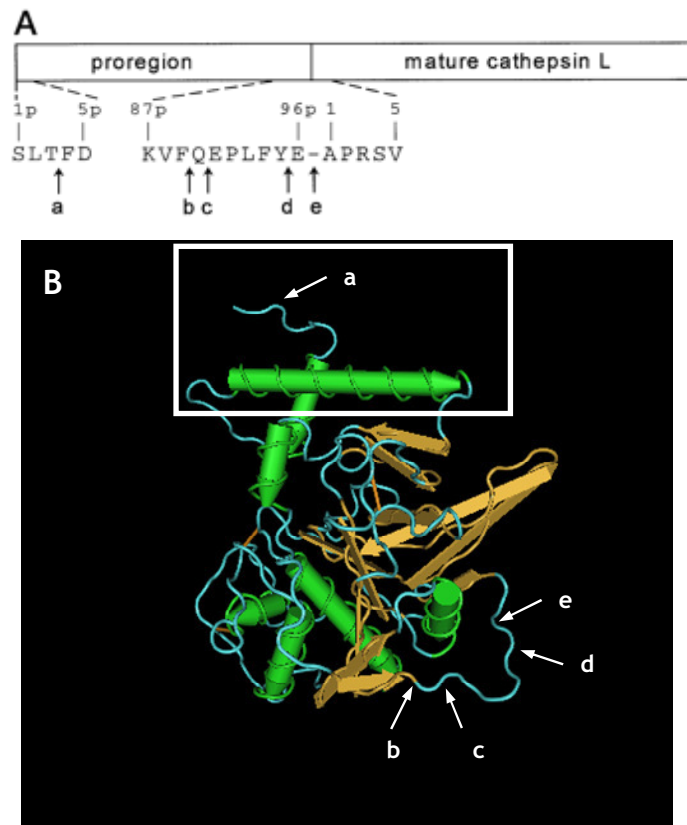


Figure 8 : Représentation schématique de la procathepsine L et des régions de clivage (flèches a à e), d'après Ménard *et al.* (1998). A : dessin de la structure primaire de la procathepsine L humaine. B : représentation tridimensionnelle de la cathepsine L humaine visualisée avec Cn3D 4.1 (N° accession 1CS8). En vert sont représentées les hélices  $\alpha$  et en marron les feuillets  $\beta$ . La région du pro-domaine est encadrée.

### C. Le facteur d'élongation $1\alpha$

Le facteur d'élongation  $1\alpha$  (EF- $1\alpha$ ) est une des protéines les plus abondantes exprimées dans les cellules eucaryotes (Slobin *et al.*, 1980). Le gène EF- $1\alpha$  est exprimé de façon constitutive et est la clé des mécanismes de traduction protéique chez les eucaryotes (Merrik *et al.*, 1992 ; Fig. 9). Il n'est donc, *a priori*, pas impliqué dans un quelconque processus parasitaire. Ce gène est très utilisé pour retracer la phylogénie au niveau spécifique (Goetze *et al.*, 2006 ; Rokas *et al.*, 2002) et semble être un marqueur relativement efficace même s'il présente au moins deux copies fonctionnelles chez la drosophile ou encore chez des copépodes marins pour lesquels les copies sont difficilement identifiables (Goetze *et al.*, 2006). Ceci rend l'identification de gènes orthologues plus délicate. Aucune information chez les Heteroderidae n'était accessible sur le gène complet, excepté chez *C. elegans* pour lequel le gène EF- $1\alpha$  complet fait 1880 pb. Cependant des ESTs de *Globodera* et d'*Heterodera* étaient disponibles dans les bases de données pour ce gène.

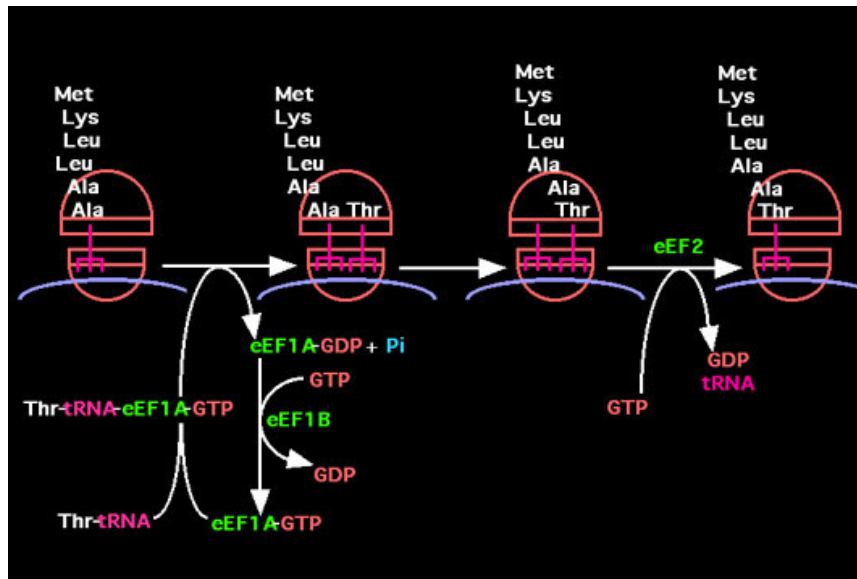


Figure 9 : Implication du facteur d'élongation 1 $\alpha$  (eEF1A) dans le fonctionnement cellulaire, plus particulièrement, lors des étapes de traduction.

L'article du chapitre 1 rend compte des résultats obtenus en terme de polymorphisme de structure et de séquence des gènes (amplification sur ADN génomique) pectate lyase, cathepsine L et facteur d'élongation pour les 40 populations et espèces de nématodes sélectionnées au sein de la famille de Heteroderidae. Lors de cette étude, nous nous sommes posés les questions suivantes :

- 1- Les séquences que nous avons obtenues sont-elles des copies orthologues des gènes amplifiés ? Nous avons pour cela comparé les séquences génomiques en terme de nombre et de répartition des introns.
- 2- Les données générées nous permettent-elles de reconstruire une phylogénie cohérente des espèces sélectionnées ? Nous avons tenté d'estimer la validité de l'utilisation des tels marqueurs en fonction de différentes échelles taxonomiques. Nous avons également évalué la pertinence de l'utilisation des introns pour définir les relations phylogénétiques entre les espèces et populations sélectionnées.

Le chapitre 2 regroupe les résultats des analyses de séquences qui nous ont permis d'apporter de nouvelles données pour évaluer l'évolution et le potentiel évolutif des gènes du pouvoir pathogène en comparaison avec le référent (facteur d'élongation). Toutes ces analyses reposent sur les relations phylogénétiques établies dans le chapitre 1 de manière à pouvoir interpréter les mécanismes de sélection observés en congruence avec l'évolution des populations et espèces. Ainsi, nous nous sommes posé les questions suivantes :

- 1- Est-ce que les gènes du parasitisme présentent des taux de mutations différents du gène de ménage ?
- 2- Comment sont réparties ces mutations le long du gène (en fonction des domaines fonctionnels) ? Et au sein des populations et espèces ? Existe-t-il des variations en fonction de l'échelle taxonomique à laquelle on se place ?
- 3- Existe-t-il des zones préférentiellement soumises à la sélection positive ? Ou des populations ou espèces chez lesquelles la sélection apparaîtrait différente du reste du jeu de données ?
- 4- Comment peut-on interpréter les phénomènes de sélection diversificatrice ou purificatrice dans le cadre des interactions plante/nématode ?

## **2. ARTICLE 4**

**ENCADRÉ N° 5 : ANALYSIS OF PARASITISM GENES IN CYST NEMATODES TO RECOUNT THE EVOLUTIONARY HISTORY OF THE HETERODERIDAE**

**POLYMORPHISME DIFFERENTIEL ENTRE LES GENES DU PARASITISME ET LE GENE DE MENAGE ?**

**1- Variabilité des gènes pectate lyase (PL), cathepsine L (CathL) et facteur d'élongation 1 $\alpha$  (EF)**

- Les taux de sites variables sont similaires : 35,9% pour PL, 28,7% pour CathL et 30,4% pour EF.
- Les distances génétiques sont proches à l'échelle intra genre ou intra espèce (~ 0,03) mais très différentes lorsque les comparaisons se font entre les genres *Globodera* et *Heterodera* : 0,149 pour EF, 0,186 pour PL et 0,270 pour CathL. Ces données suggèrent une accélération du tempo d'évolution des gènes du parasitisme entre les deux genres.

**2- Distribution de la variabilité dans les codons**

La fréquence des mutations en première et deuxième position de codon, correspondant pour la majorité à des substitutions non synonymes, est élevée (~ 45%) pour les gènes du parasitisme en comparaison avec le gène de ménage (< 20%). Ceci traduit un relâchement des pressions de sélection qui n'autorisent en règle générale que les mutations neutres.

**EVOLUTION DES GENES DU POUVOIR PATHOGENE**

**1- Relations phylogénétiques des genres et espèces d'Heteroderidae**

- Une topologie identique est observée sur la base des gènes CathL et EF avec regroupement des espèces *G. tabacum* et *G. rostochiensis* versus *G. pallida* et *G. "mexicana"* pour le genre *Globodera* qui est un groupe monophylétique.
- Le gène pectate lyase présente un regroupement des nématodes à kyste de la pomme de terre d'un côté (*G. pallida* et *G. rostochiensis*) et des nématodes du tabac de l'autre (*G. tabacum*). S'agit-il d'une convergence évolutive des gènes pectate lyase sous l'influence de la plante hôte ?

**2- Rôle informatif des introns en phylogénie**

- Les introns permettent d'obtenir une meilleure robustesse de l'arbre des cathepsine L.
- Pour les gènes PL et CathL, certains introns ont évolué par acquisition de séquences de 10 à 173 pb au cours des spéciations du genre *Globodera*.

## **ANALYSIS OF PARASITISM GENES IN CYST NEMATODES TO RECOUNT THE EVOLUTIONNARY HISTORY OF THE HETERODERIDAE**

Alexandra BLANCHARD, Didier FOUVILLE, Olivier PLANTARD, Eric GRENIER.

Institut National de la Recherche Agronomique (INRA) - Agrocampus, Unité Mixte de Recherche  
Biologie des Organismes et des Populations appliquée à la Protection des Plantes (UMR BiO3P).

Domaine de la Motte. BP 35327. 35653 LE RHEU CEDEX. FRANCE

Tel: +33-2-23.48.51.73 / Fax: +33-2-23.48.51.50

Corresponding authors. E-mail addresses: [alexandra.blanchard@rennes.inra.fr](mailto:alexandra.blanchard@rennes.inra.fr),  
[eric.grenier@rennes.inra.fr](mailto:eric.grenier@rennes.inra.fr)

*En préparation*



## **ABSTRACT**

The use of nuclear genes as phylogenetic markers is not usually performed in nematode family to recount the evolutionary history of these organisms. We amplified and characterised three nuclear genes - two parasitism genes (pectate lyase and cathepsin L) and a housekeeping gene (elongation factor) - among nearly 40 populations from 13 species belonging to the Heteroderidae family. We clearly showed that the intraspecific genetic distances are approximately the same whatever the gene studied with an average of 0.30. But focusing at the intergenera level, we observed higher distances using PL and CathL genes with an average of 0.27 and 0.18 respectively, compared to the EF that displayed only 0.15. We further investigate on the nature of the polymorphism observed by 1- looking at the position of the mutations inside the codons, 2- recounting the gene tree among the populations and species used. We first demonstrated that the mutations appeared along the parasitism genes with a high frequency on the two first positions within the codon. Differential evolutions between genes were supposed by the strong proportion of mutations observed on the first two positions of codon (~ 50% for both parasitism genes against ~ 80% for the EF gene, and differential indels in the CathL and PL introns between the species. The gene trees obtained were similar for CathL and EF, while the tree based on PL provides a different topology. We focused on the phylogenetic analyses to evaluate the relevance of the use of nuclear genes as phylogenetic markers. The methods to obtain the best trees are discussed.

## 1. INTRODUCTION

Sedentary plant parasitic nematodes are responsible for great damages on cultivated plant all around the world that represent 100 billion dollars per year. Cyst nematodes (Tylenchida: Heteroderidae) (Subbotin *et al.*, 2001) are parasites that have evolved strategies of parasitism involving complex molecular mechanisms to elaborate an extraordinary relationship with their host plant. Within the Heteroderidae family, the *Globodera* genus parasiting Solanaceous plants constitutes a particularly interesting group with species corresponding to several important pests of potatoes (*G. rostochiensis* and *G. pallida*) or tobacco (*G. tabacum*).

Cyst nematodes are able to penetrate the roots, to move towards the central cylinder in order to connect themselves to the vascular system and feed using the modified plant cells they have induced, called syncytium. They principally used salivary secretions to establish the molecular dialog with the plant that allowed their development and reproduction (Davis *et al.*, 2004).

We chose 40 populations among the *Globodera* and *Heterodera* genera representing the broadest diversity in Heteroderidae in term of host plant, genetic variability, pathotypes and geographic origin, to study their molecular evolution. We studied two parasitism genes, the pectate lyase (PL) and the cathepsin L (CathL) genes and we compared the data obtained to that of a housekeeping gene, the elongation factor 1 $\alpha$  that was considered as not implicated in parasitism. The pectate lyase is a cell wall degrading enzyme that induces the relaxation of the plant cell wall to facilitate the migration of the nematodes towards the central cylinder of the root (Popeijus *et al.*, 2000; De Boer *et al.*, 2002). Cathepsin L gene seemed to be involved in host defence evasion (Nomura *et al.*, 2005), the invasion of the host (Sajid and McKerrow, 2002), or the feeding (Shingles *et al.*, 2006). Recently, it has been shown as differentially expressed between virulent and avirulent population of the root-knot nematode *Meloidogyne incognita* (Neveu *et al.*, 2003a, Neveu *et al.*, 2003b). The elongation factor 1 $\alpha$  gene is currently used in phylogeny (Knutsen *et al.*, 2004; Rokas *et al.*, 2002; Suarez *et al.*, 2006) as a gene that can be efficient to resolve the interspecific and intergenera relationships (Goetze *et al.*, 2006).

None of the gene chosen has already been used to recount the phylogenetic history of the Heteroderidae family. Indeed, the phylogenetic data available in nematodes were principally based on ITS (Subbotin *et al.*, 2001), or at the intra-specific level on AFLP markers (Marché *et al.*, 2001) or on bidimensional gel electrophoresis of proteins (Bossis and Mugniéry, 1993). Such markers are sometimes of limited value for phylogenetic investigations due to homoplasy when comparing species that are too divergent or due to incomplete gene conversion in the case of ITS multicopy genes.

In this study, we first characterised three nuclear genes, EF, CathL, and PL among nearly 40 populations of nematodes belonging to 13 species of the Heteroderidae subfamily describing their intron-exon structure and the presence of functional domains within the sequence. We used all these data to recount evolutionary history of these genes among the cyst nematodes. The relevant use of these genes as markers to study the phylogenetic relationships of the populations and species was assessed by comparing methods to obtain well supported trees.

## 2. MATERIAL & METHODS

### 2.1. Biological material

A set of 40 populations representing 13 species of cyst nematodes were selected (table 1). This set of populations and species was chosen in order to represent the widest diversity in term of genetic variability, geographic origin, pathotypes or host ranges. The cysts of all those populations were rehydrated at least one night in distilled water before DNA extractions.

Table 1: Populations and species used to study the polymorphism and the evolutionary history of the pectate lyase, cathepsin L and elongation factor 1 $\alpha$  genes.

	Population name	Geographic origin	Code
<b>Genus <i>Globodera</i></b>			
<b><i>G. pallida</i> from Europe</b>	Luffness (Pa2/3)	Scotland	GPE1
	Ouessant (Pa2/3)	France	GPE2
	Chavornay (Pa2/3)	Switzerland	GPE3
	Pukekoe (Pa2)	New Zealand	GPE4
	Duddingston (Pa1)	England	GPE5
<b><i>G. pallida</i> from South America</b>	Arapa/Scicuani	South Peru	GPS2
	Amantani	South Peru	GPS3
	Puno	South Peru	GPS4
	Cutzco	Central Peru	GPS5
	Andahuaylas	Central Peru	GPS6
	Huancayo	North Peru	GPS7
	Chocon (P4A)	Peru	GPS8
	Otuzco (P5A)	Peru	GPS9
	Huamacucho (P6A)	Peru	GPS10
	<b><i>G. rostochiensis</i> from Europe</b>	RO1	France
RO4		The Netherlands	GRE2
RO5		The Netherlands	GRE3
<b><i>G. rostochiensis</i> from South America</b>	B1	Bolivia	GRS1
	B4	Bolivia	GRS2
	C1	Chile	GRS3
<b><i>G. mexicana</i></b>	Popocatepetl	Mexico	GM1
	Santa Ana	Mexico	GM2
<b><i>G. tabacum</i></b>			
<b><i>G. tabacum tabacum</i></b>	Miller Connecticut	USA	Gtt1
	Agen	France	Gtt2
<b><i>G. tabacum virginiae</i></b>	GV2	USA	Gtv
<b><i>G. tabacum solanacearum</i></b>	GS2	USA	Gts
<b><i>G. tabacum azteca</i></b>	75181	Mexico	Gta
<b>Genus <i>Heterodera</i></b>			
<b>Schachtii group</b>			
<b><i>H. cajani</i></b>	India	India	Hcaj
<b><i>H. schachtii</i></b>	HS41	Germany	HS41
	HS41 vir	Germany	HS41vir
<b><i>H. glycines</i></b>	VL1, TN16, TN19	USA	HG-VL1, HG-TN16, HG-TN19
<b><i>H. trifoli</i></b>	113	France (Corsia)	Htri
<b><i>H. ciceri</i></b>	Baayat	Syria	Hcic
<b>Avenae group - <i>H. avenae</i></b>	Ha41 Villasavary	France	Ha
<b>Sacchari group - <i>H. sacchari</i></b>	Bouaké	Ivory coast	Hsac
<b>Humuli group - <i>H. humuli</i></b>	Dutzenheim	France	Hh
<b>Goettingiana group - <i>H. carotae</i></b>	5001	France	Hca

## 2.2. DNA extraction, PCR amplification and sequencing

DNA was extracted from 1 cyst of each population or from one juvenile to amplify the elongation factor gene. The eggs of the cyst or the juvenile were crushed and the broyat was recovered in proteinase K buffer (10 mM Tris pH 8.8, 1 mM EDTA, 1% Monidet P40-Tergitol and 100µg/mL proteinase K). The samples were placed at -80°C during 15 min and then incubated 1h at 60°C and finally 15 min at 95°C. The tubes were centrifuged and 5 µL of the supernatant was used to amplify the three genes. We designed specific primers to amplify each gene in *Globodera* and *Heterodera* genera.

The cathepsin L gene was amplified in a final volume of 25 µL containing 1X buffer, 3 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.4 µM of each primer (CathepFwd: 5'-CGCGAGCGTGCAATTGA-3' for the *Heterodera* spp.; CathepFwd5: 5'-CGTGATCGTGCCATAGA-3' for the *Globodera* spp.; CathepRev5: 5'-CAACTGTTCTTACAATCCA-3' for both genera) and 1.25 U of Taq polymerase (Goflexi, Promega). The PCR program was 1 min at 95°C, 35 cycles with 30 s at 95°C, 30 s at 54°C, 3 min at 72°C, and a final elongation step at 72°C during 5 min.

The pectate lyase gene was amplified in a final volume of 25µL containing 1X buffer, 2 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.2 µM of each primer (PLFwd.2: 5'-GGCGGTTCCGGCATT-3' for both genera; PLRev: 5'-CCCAGAAAGTGACATTTTGT-3' for the *Globodera* spp; PLRev.2: 5'-CATTTTGTACCGTAGTTGGA-3' for the *Heterodera* spp.) and 1.2 U of Taq polymerase (Qbiogen). The PCR program was 1 min at 96°C, 35 cycles with 20 s at 96°C, 20 s at 55°C, 1.5 min at 72°C, and a final elongation step at 72°C during 5 min. Identical PCR conditions were used to amplify the pectate lyase cDNA using the primers PLFwd4 (5'-ATGCTTTTTGTTATCATTTCAAT-3') and PL Rev3 (5'-TAGTTGACAATTTTAAYWGCCGA-3') for the *Globodera* genus and PLFwd5 (5'-ATGCTTTTCATTCTACTGGTCAT-3') and PLRev4 (5'-TTAGCTCGCAATTTTAACTGCC-3') for the *Heterodera* genus.

The Elongation factor 1 α gene was amplified in a final volume of 25µL containing 1X buffer with MgCl<sub>2</sub>, 1µM of each primers (EF1Fwd: 5'-GCTTGGGTGYTKGACAAR-3', and EF1Rev: 5'-CATCYTRTCAGACGGCTCCA-3'), 1.125U of Taq Polymerase (QBiogen 15U/µl) and 0.1mM of each dNTP. The PCR program was 94°C for 4 min, 40 cycles of 30 s at 94°C, 30 s at 58°C and 1 min at 72°C, and a final elongation step of 5 min at 72°C.

All the PCR products were purified on G50 sephadex superfine (Amersham Biosciences) and sent for sequencing to Macrogen ([www.dna.macrogen.com](http://www.dna.macrogen.com)).

## 2.3. Sequence analysis and phylogenetic relationships establishment

For each gene, the dataset was partitioned into intron and exon sets. Each exon set was first separately analysed and concatenated in some cases. Alignments of the DNA sequences were generated using Multalin (Corpet, 1988; <http://prodes.toulouse.inra.fr/multalin/multalin.html>) with DNA 5-0 in alignment parameters and other default options. The alignments were manually corrected when necessary. After alignment of all the nucleotidic sequences gene, the dataset was reduced in length in order to conserve only the common portion to a maximum of populations.

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.1 (Kumar, Tamura & Nei 2004) to generate trees with the Neighbour Joining model. Phylogenetic analyses were inferred with the Maximum Likelihood (ML) or Maximum Parsimony (MP) models using optimally criteria defined by Modeltest 3.7 (Posada and Crandall, 1998) in PAUP 4.0 (Swofford, 1998). ML and MP trees were generated through heuristic searches with 100 stepwise additions. Gaps were encoding as missing data or specifically encoded using Gapcoder (Young & Healy, 2003) and implemented in PAUP.

### **3. RESULTS**

#### **3.1. Nucleotidic polymorphism of the elongation factor, cathepsin L & pectate lyase genes**

We amplified the three genes on the greatest number of populations and species among those selected. Some amplification difficulties for the PL and CathL genes were encountered with some species of *Heterodera* genus distantly related to *H. schachtii* or *H. glycines*, two species for which many EST can be found in databases. The amplified products were about 800bp for the pectate lyase gene (exons: 470 bp, introns: 300 bp) and nearly 2000 bp for the cathepsin L gene (exons: 860 bp, introns: 1200 bp). These amplification products correspond to about 60% of the total genomic sequence for the pectate lyase (PL) gene and 85% of that of the cathepsin L (CathL). A band of about 650 bp was obtained for the elongation factor 1 $\alpha$  (EF) containing 450 bp of exons and 200bp of introns. All the PCR products were directly sequenced and the sequences were deposited into genbank under the accession numbers XX.

##### ***A. Analysis of the pectate lyase gene sequences***

A portion of the PL gene was characterised among 28 populations of cyst nematodes at the genomic level and in 17 populations at the cDNA level. The PCR products amplified were aligned with the transcript sequences to identify the precise splicing site of the introns that were bordered by canonical cis-splicing sequence "GT-AG" (Blumenthal and Steward, 1997). The portion of the pectate lyase gene amplified contains 3 introns and 4 exons (2 complete and two partial). The average base composition of the exonic zone was 52.1% A+T rich (Table 2A). The polymorphism analyses of the dataset showed that out of the 123 variable sites observed, 60 are informative in parsimony (Table 2). In the coding region, the variability calculated - corresponding to the number of position presenting mutations among the dataset - was 35.9% (n= 123/342) (Table 3). If we looked at the localisation of the mutations within the codon, we observed that 27.5% were on the first base, 13.7% on the second base and 58.7% on the third base of the codons (Fig. 1). Similar results were obtained using the cDNA dataset presenting a longer coding sequence (Table 2 & Fig. 1).

Table 2: Nucleotides and amino acids analysis of the three genes among all the populations of the dataset.

A				
Pectate lyase	Size of the coding sequence used	Number of variable sites	Number of parsimony informative characters	%A+T
Nucleotidic gene sequences	342 bp	123	60	52.1
Amino acid sequences	114 aa	43	11	-
Nucleotidic cDNA sequences	690 bp	219	191	53.4
Amino acid sequences	230 aa	72	67	-
B				
Cathepsin L	Size of the coding sequence used	Number of variable sites	Number of parsimony informative characters	%A+T
Nucleotidic gene sequences	861	299	190	47.4
Amino acid sequences	287	111	58	-
C				
Elongation factor 1 $\alpha$	Size of the coding sequence used	Number of variable sites	Number of parsimony informative characters	%A+T
Nucleotidic gene sequences	371	132	101	47.1
Amino acid sequences	123	43	16	-

Table 3: Percentage of sites that are identified as variable (that displayed one or more mutations among the entire dataset). The percentage is calculated using the same range of populations.

	Pectate lyase	Cathepsin L	Elongation factor
<i>G. pallida</i>	9.9 %	10.9 %	12.7 %
<i>Globodera</i> genus	17.8 %	16.5 %	22 %
Heteroderidae family	35.9 %	28.7 %	30.4 %

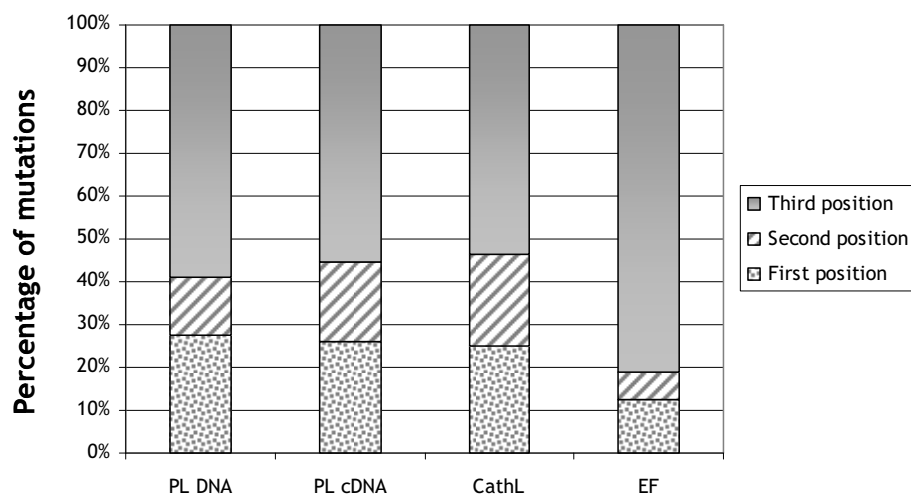


Figure 1: Distribution of the mutations within the codons for each gene and cDNA studied.

In the partial amino acids sequences deduced from our dataset, the conserved regions characteristic of the class III pectate lyase were all identified (Fig. 2A). The three conserved characteristic regions ranged from aa 35 to 42, 66 to 73 and 86 to 98 as described in other nematode PL (Doyle and Lambert, 2002 ; Popeijus *et al.*; 2000). The class III PL seemed to be cysteine rich as described by Schevchik *et al.* (1997), with seven residues identified in the 28 populations tested. An insertion of 9 nucleotides, thus without modification of the open reading frame, was specifically observed in the PL coding sequences, of *Heterodera* spp. encoding a "NSS" motif for the *H. glycines* population (TN16) tested and a "NSR" motif for only one population of *H. schachtii* (HS41) between the amino acids 32 and 33, where the intron number 3 is also inserted. These motifs were not found in other *Heterodera* species for which we obtained cDNA sequences such as two populations of *H. schachtii* (HS42, HS41vir) and the one population *H. trifoli* (see Table 1). The variability of the *Globodera* PL introns ranged from 46% for the intron number 2 to 62.9 % for the fourth (average intron variability: 53%). The average intron variability reached 70% if we include the *Heterodera* spp. sequences.

The introns number 2 and 3 displayed specific insertions in the *Heterodera* sequences while intron number 4 displayed specific insertions in *G. rostochiensis* and *G. tabacum* sequences (Figure 2A). The *G. pallida* and *G. "mexicana"* species shared the same intron/exon structure. Despite the relatively high rate of variability described within the PL introns, some highly conserved regions can also be identified in the PL intron sequences among all the populations including the *Heterodera* sequences (data not shown). Considering the intron/exon boundaries and the identification of several highly conserved regions in both introns and exons, we considered that we compared orthologs of the PL gene among our dataset.

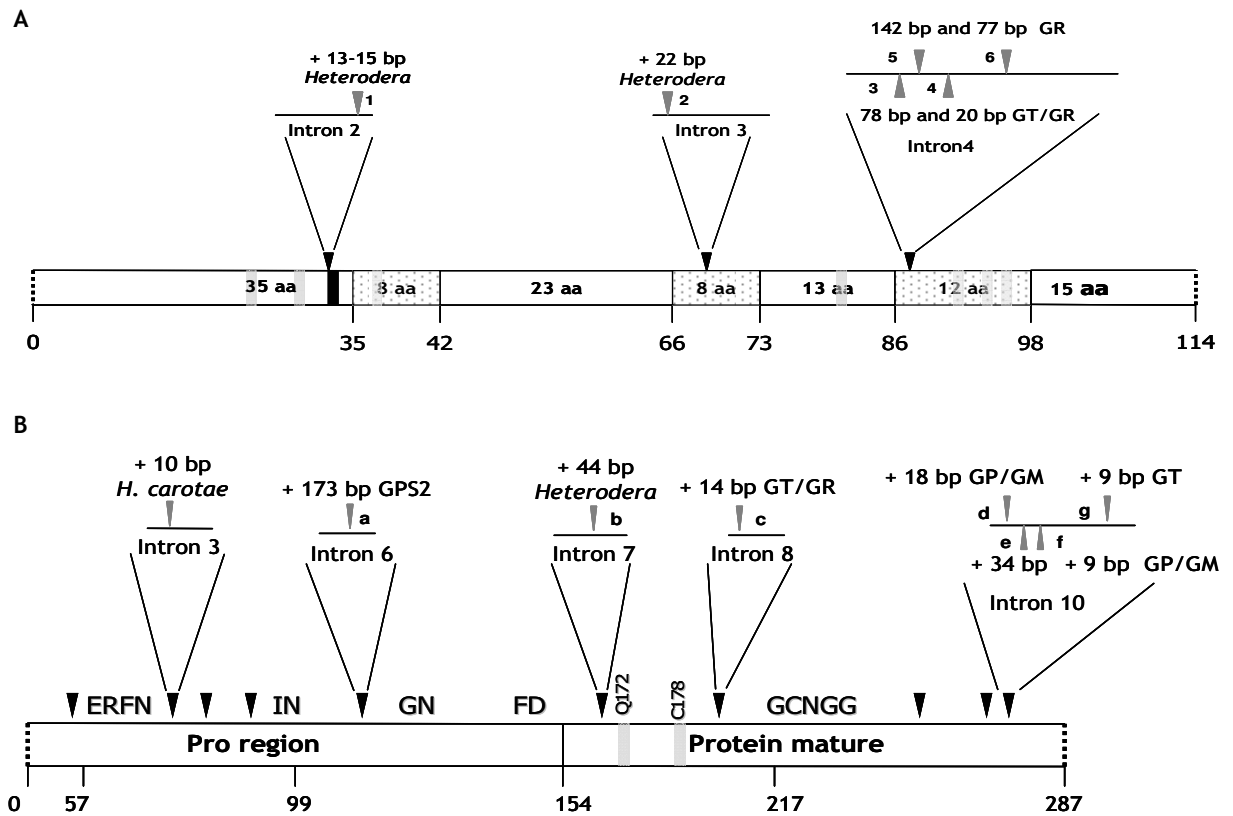


Figure 2: Schematic representation of the predicted proteins encoded by the partial pectate lyase and cathepsin L genes. The black arrows indicated the intron position - the introns were numbered from 2 because of the use of a partial pectate lyase and cathepsin L gene - and the grey arrows indicated intron positions differentially distributed among the Heteroderidae family (reported in the figure 6). A: Representation of the partial pectate lyase protein encoded by the PCR product amplified. The pectate lyase representation showed the class III pectate lyase characteristics. In grey boxes are represented the cystein residues, in dot boxes, the conserved regions and in black box, the specific insertion in *Heterodera* species sequences. B: Representation of the cathepsin L partial protein encoded by the PCR product obtained. The predicted protein representation showed the conserved characteristic regions of the cathepsin L protein that are both in the pro and mature regions.

### B. Analysis of the cathepsin L gene sequences

The PCR products were obtained on 32 populations from the content of only one cyst and were directly sequenced in 3 overlapping portions to obtain the full sequence. The positions of the intron/exon boundaries were predicted by comparing CathL gene sequences to cDNA sequences. All the introns were bordered by the canonical rule of "GT-AG", except for the intron number 4 in all the *Globodera* sequences that begin with "GC". The Cath L gene portion amplified was 861bp long



and contained 9 introns and 10 exons (two partial at the extremities). The average base composition of the exonic portion was 47.4% A+T rich as indicated in table 2. In the entire dataset, 299 variable sites were identified with 190 informative in parsimony which correspond to 34.7 % (299/861) of variable sites within the entire dataset. If we compared to PL data, using the same dataset, this average variability drops to 28.7% of variable sites (Table 3). An unequal repartition of the substitutions was showed within the codons, but relatively similar with what was observed for the PL gene: half of the mutations were observed on the third codon position and half on the first and second positions (Fig. 1). The deduced amino acid sequences contained the pro region domain from 1 to 125 and the mature region of the protein from 126 to the end (Fig. 2B). The amino acid sequences showed characteristic regions of the cathepsin L genes, as the ERFNIN, GNFD and GCNGG motifs described by Barret *et al.* (1998), and the crucial amino acids that are implicated in the active site formation: Q172 and C178 (Fig. 2B).

The cathepsin L gene contained numerous introns. Strong differences in term of variability were observed among the CathL introns due to indels and substitutions. Considering for example only the *Globodera* species, introns variability ranged from 9.2% to 73% for intron number 2 to 10 with an average of 39.4%. Including the *Heterodera* species, the variability ranges from 43% to 79% with an average of 68.4%. Specific insertions were observed at the species or genera level. The introns number 3 and 7 showed *Heterodera* specific zones, whereas introns number 6, 8 and 10 showed insertions specific of *Globodera* species (Fig. 2B). Insertions shared by two *Globodera* species were observed, it was between *G. pallida* and *G. "mexicana"*, or between *G. tabacum* and *G. rostochiensis*. All the intron/exon boundaries were conserved between all the populations tested, and then ensured that orthologs were compared.

### **C. Analysis of the elongation factor 1 $\alpha$ gene sequences**

The EF PCR products were obtained from one juvenile in 40 populations. The EF gene portion was 371bp and showed 128 variable sites with 190 informative in parsimony (Table 2C). The average base composition was 47% A+T rich (Table 2C). If we considered the same data set such as that for the parasitism genes, we identified 30.4% of variability (Table 3). The repartition of the substitutions within the codons was different than the one observed for the parasitism genes with 81.1% of the mutations observed on the third codon position (Fig. 1). EF introns showed a quite high variability. Considering only the *Globodera* species, introns number 1 and 2 showed respectively 51 and 82% of variability. If we add *Heterodera* sequences, the variability increased until 94% for the first intron and 86% for the second. Even if the variability observed was high, the intron/exon boundaries were always conserved in this gene.

### **3.2. Evaluation of saturation of the phylogenetic signal in the dataset**

When p-uncorrected distances for overall taxa were plotted against Kimura 2 parameters distances for each codon position, a linear relationship was observed for the first and second positions for all the three genes (Fig. 3). However, the saturation for the third codon positions was suggested by the nonlinear curve observed for the three genes (Fig. 3). We noted that this saturation was observed only when we compared the most divergent species, at the inter genera

level. At the genus level, no saturation of the dataset was detected whatever the gene analysed. As *Heterodera* spp. were mostly used as an out-group in this study, we considered that the saturation observed in the third position of the codon will not affect the topology of the trees in the ingroup (*Globodera* spp.).

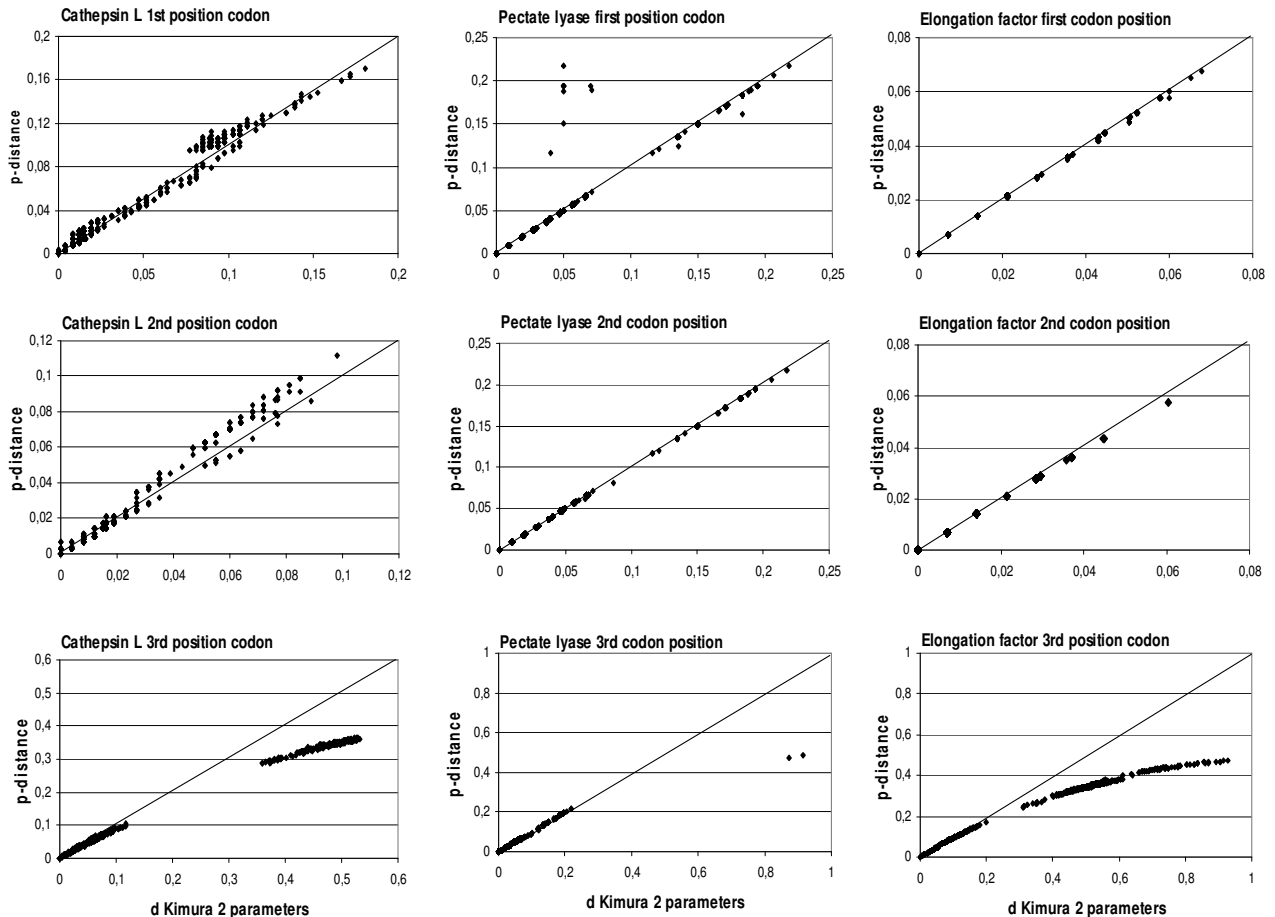


Figure 3: Evaluation of the saturation of the data sets in each codon position for the pectate lyase, cathepsin L and elongation factor genes. d kimura 2 parameters: genetic distances calculated using Mega with the Kimura 2 parameters model.

### 3.3. Distribution of the genetic distances among *Globodera* and *Heterodera* genera

The genetic distances (pairwise comparison) were estimated using the Kimura 2 parameters model that takes into account transition *versus* transversion bias (higher rates of transitions than transversions). A very clear bimodal distribution was observed for the three genes studied corresponding to two separate normal distributions (Fig. 4). The first mode corresponds to the intra specific or intra genus comparisons. For the three genes, the means of this first mode were very similar (0.032 for the EF, 0.029 for the CathL and 0.030 for the PL). The second mode corresponds to comparisons at the inter genera level (*Globodera* *versus* *Heterodera*). In contrary to the first

mode, the mean value of this second mode differs according to the gene studied. The smallest distance is observed for the EF gene (mean = 0.149), while higher genetic distances were observed for the CathL and PL genes (mean = 0.186 and 0.27 respectively). Although those comparisons were not based exactly on the same populations and species according to the gene studied, these conclusions hold true when considering reduced comparisons at the same species level.

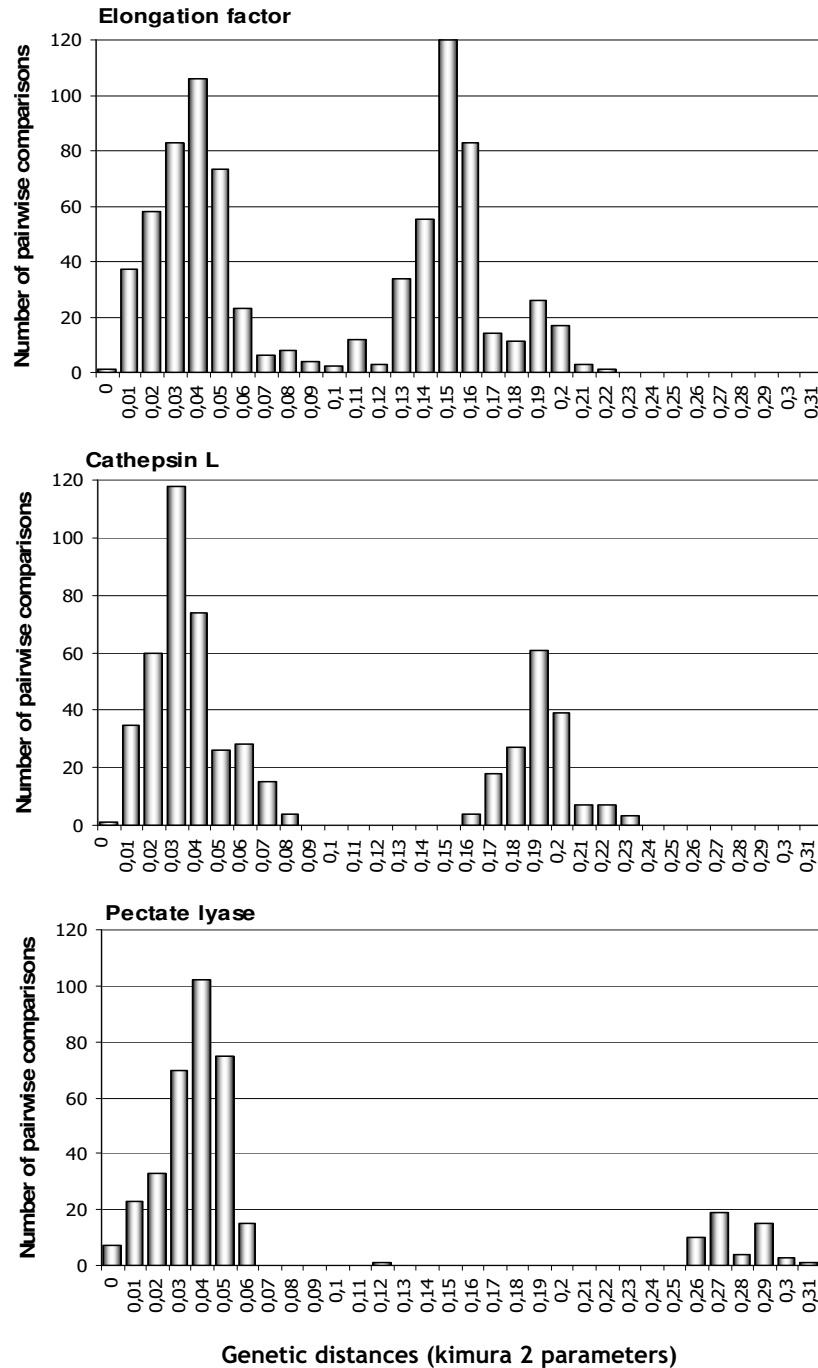


Figure 4: Distribution of the pairwise comparisons among classes of genetic distances that were calculated using MEGA with the kimura 2 parameters model.

### **3.4. Phylogenetic relationships of the species and genera of Heteroderidae**

After separate sequence alignment of the PL and CathL genes, each dataset was analysed using Mega software for the NJ model, or PAUP for the ML or MP models. In all cases, the monophyly of the *Globodera* and *Heterodera* genera was confirmed (Fig. 5 A & B).

For the *Globodera* genus, the different sequences representative of a given species (*G. rostochiensis*, *G. tabacum*, *G. pallida* or *G. "mexicana"*) always form a monophyletic group, usually with a high bootstrap support of 100% for each gene (Fig. 5 A & B). The *G. "mexicana"* sequences were found within the *G. pallida* cluster with a relatively supported node (99% for the PL tree, Fig. 5A, and 75% for the CathL tree, Fig. 5B). A sistergroup relationship between *G. rostochiensis* and *G. tabacum* was observed in the Cathepsin tree (bootstrap support ranging from 75 to 92 %) as for the EF tree (data not shown). This topology was different in the PL tree where *G. tabacum* appeared as the sistergroup of a *G. rostochiensis*/*G. pallida* group (this last group being supported by bootstrap value of 92%) (Fig. 5A).

### **3.5. Did the intron data improve the resolution of the phylogenetic relationships?**

The CathL introns were individually aligned and concatenated with the exon alignment. Then a NJ tree was generated in pairwise deletion. We observed a considerable improvement of the tree resolution, more particularly for the internal nodes (Fig. 5C). This tree strongly supports, in particular, the clustering of the *G. rostochiensis* and *G. tabacum* species and the clustering of the *G. "mexicana"* and *G. pallida* species (bootstrap value of 100%).

In contrary to this gene, the concatenation of the intron/exon data for the PL gene did not improve significantly tree resolution. We used the Gapcoder software to encode the characteristics indels observed in the PL introns and use them as phylogenetic informations. Despite some of these indels clearly allow the distinction of a GR/GT group from GP, the addition of these informations in the PL dataset did not improve the tree obtained in NJ with the exons (data not shown).

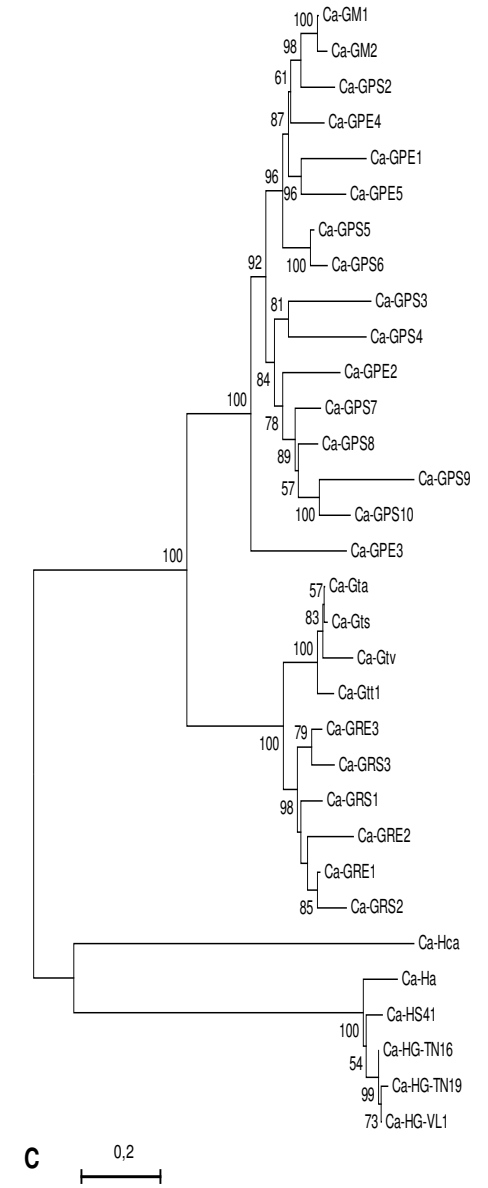
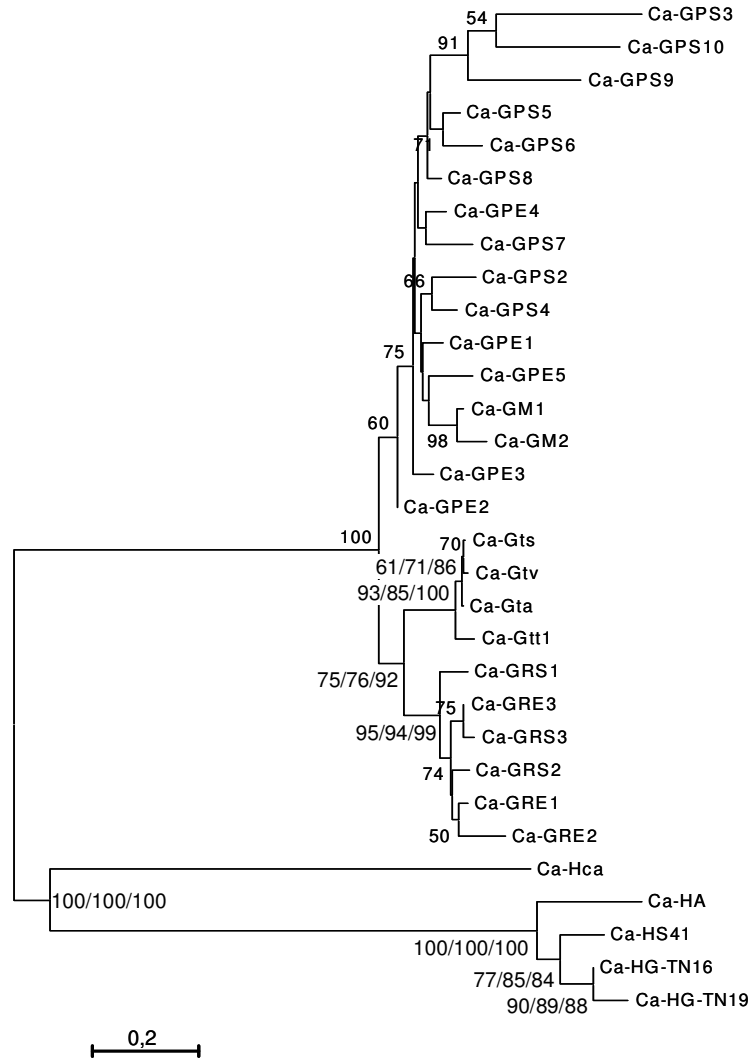
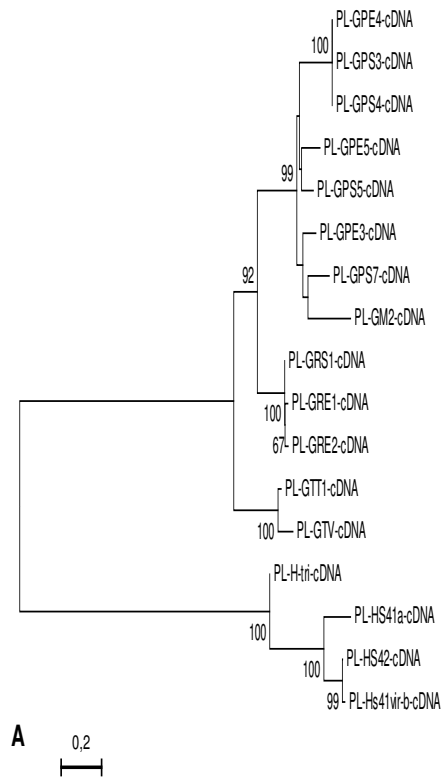


Figure 5: Tree inferred from separate phylogenetic analyses of the pectate lyase transcripts and cathepsin L genes. A: Phylogeny of pectate lyase sequences obtained from cDNA amplifications, based on Neighbour Joining analysis using Mega software. B: Phylogeny of cathepsin L gene obtained from coding region amplifications, based on Maximum Parsimony (MP) and Maximum Likelihood (ML) analyses using PAUP 4.0 software and with Neighbour Joining (NJ) model using MEGA software. Tree statistics are represented in MP/ML/NJ for the basal nodes that are the sustain nodes in MP and ML. The internal nodes are only supported using NJ model. C: Phylogeny of cathepsin L gene obtained from the entire gene amplifications (coding and non coding) based on Neighbour Joining analysis.

### **3.6. Evolution of the introns within the Heteroderidae family**

The concatenation of the exons of each gene in a single dataset is not allowed to recount the evolutionary history of the genes because of the conflicted topology observed between the PL tree and the CathL and EF trees. But this well supported tree could be used to understand the intron evolution in CathL and PL genes. We previously mentioned, in the part 3.1, specific indels in PL and CathL introns. These specificities were reported on the concatenated species tree (Fig. 6) and showed that for both genes, indels identification correlates to tree topology. For the CathL gene, indels were scattered among the different species branches but were mainly appeared after the *G. pallida* speciation.

The PL gene seemed to have various intron evolutions because no indels appeared in the *G. pallida* group. But 2 were observed before the *G. tabacum*/*G. rostochiensis* separation and 2 other after the *G. rostochiensis* speciation.

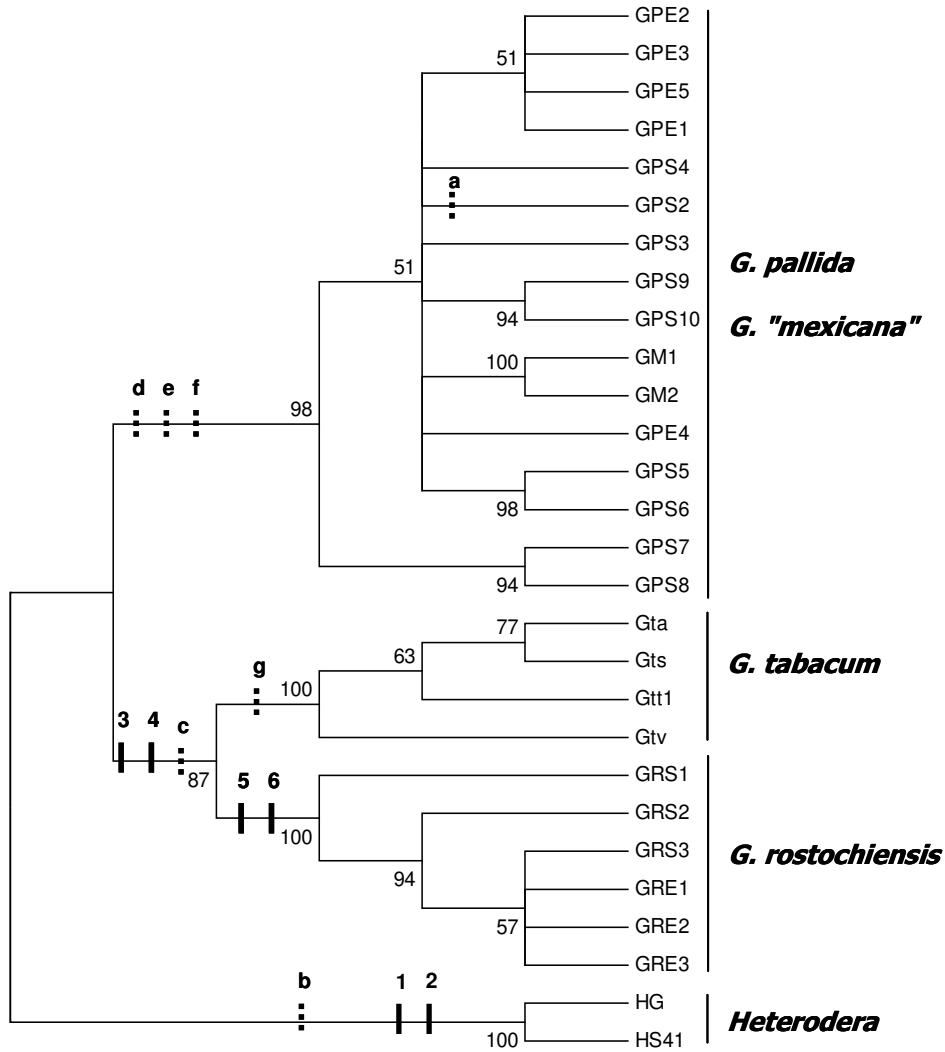


Figure 6: Condensed tree based on concatenated data of pectate lyase, cathepsin L and elongation factor exons using NJ model. The specific indels mentioned in figure 2 were reported onto the tree. The black bars indicated the indels of the PL gene and the dot bars are the indels of the CathL gene.

## 4. DISCUSSION

In this study, we examined the phylogenetic relationships of 40 population of nematodes among 13 species that belong to the Heteroderidae family, based on the three nuclear genes EF1 $\alpha$ , PL and CathL. In order to obtain the best topology, we evaluate the relevance of the use of the use of 1- parasitism or housekeeping genes as phylogenetic markers, 2- introns and/or exons data. We also evaluated the various uses of the nucleotide data set to obtain the tree with the best resolution (concatenated data, sequence size).

The pectate lyase and cathepsin L genes were previously identified in few species of nematodes (Popeijus *et al.*, 2000; De Boer *et al.*, 2002; Neveu *et al.*, 2003a; Shingles *et al.*, 2006) as genes involved in parasitism. In the present study, we described the characterisation of both genes in a large data set (near 40) of populations belonging to the Heteroderidae family. This approach required to compare orthologous copies of the genes studied within the populations. For both the PL and CathL genes, the sequences generated shared common structures of proteins such as conserved domains, a same number of introns associated with the same intron/exon boundaries. These features suggest that we compared orthologues copies of the PL and CathL genes since great variations in term of intron position were observed in paralogous copies of nematode pathogenicity genes such as cellulase genes (Gao *et al.*, 2004). Moreover numerous studies sustained the high conservation of the intron/exon structure among clades (COGs) (Rokas *et al.*, 1999; Wada *et al.*, 2002) even if some studies contradicted this rule (Goetze *et al.*, 2006). Our identification of orthologous relationships was reinforced by the numerous indels shared by different species in some intronic regions since it seemed unlikely that all these insertion-deletion events occurred at the same location independently in different lineages.

The average base composition observed for the three genes (46.6 to 52.9% of A+T) is congruent with the data described by Mitreva *et al.* (2006) that mentioned rates around 51% for the *Globodera* and *Heterodera* species, based on a huge of ESTs. They showed similar patterns of codon usage among species within the Heteroderidae family. However, compared to other nematode genera, the cyst nematodes showed particular pattern of G-C contents that probably could not only change base composition in third position codon but also alter first and second positions and even amino acids sequences (Mitreva *et al.*, 2006). It probably should be taking into account in molecular evaluation of the evolution of the parasitism gene in Heteroderinae subfamily.

It was quite surprising to note that the number of sites that displayed mutations at least in one of the population tested - called variable site - is equal whatever the gene we look at. But the difference between housekeeping gene and parasitism genes was revealed with the study of the repartition of these variables sites within the codons. Approximately 80% of the variable sites were identified in the third codon position for the EF gene while for the parasitism genes, approximately 45% of the variable sites were on first and second codon positions. This means that the PL and CathL genes possess more variable sites that lead to amino acids mutations than EF. These data suggest that diversifying selection could acts on these genes. This was also supported by the genetic distances that were greater in parasitism genes comparing *Globodera* and *Heterodera* genera than



in housekeeping gene, suggesting an acceleration of the evolutionary tempo of these genes. Diversifying selection could be an indication of an adaptation of the nematode to external constraints as it was shown for genes implicated in interactions (Weber *et al.*, 2005; Sawires *et al.*, 2006; Tsai *et al.*, 2006) or genes involved in the adaptation to an habitat or a host (Paulsen *et al.*, 2002; Chen *et al.*, 2006).

The phylogenetic relationships observed among *Globodera* populations using the nuclear genes studied were congruent with those generated using the 2D technique or the ITS-RFLP (Bossis and Mugniéry, 1993; Grenier *et al.*, 2001), except for the PL gene that clustered *G. pallida* and *G. rostochiensis* as a sister group of *G. tabacum*. Given that both nematodes are potato cyst nematodes, we can hypothesis that the mutations observed in the PL gene were shaped in some instance by the plant host cell wall components.

The intron informations greatly improved the internal nodes of the CathL tree and therefore appeared to constitute an important source of phylogenetic information and more particularly the indels of these non coding regions as it was described by Gonzalez *et al.* (2006). Previous studies conducted on non nematode genes already showed that introns can be considered as neutral markers that contain useful information for molecular phylogenetics (Friesen *et al.*, 1997; Creer *et al.*, 2006). But to date, no clear information is available to know how to encode the indel informations particularly in non coding sequences (Creer *et al.*, 2006). Classically, the indels are considered as missing data (Kumar *et al.*, 2004; Swofford *et al.*, 1998), but regarding to our dataset, such a treatment clearly discard a valuable source of information that certainly not appeared randomly in the sequences. Some studies try to evaluate the optimal use of these informations to study molecular phylogenetics (Gonzalez *et al.*, 2006) but it seemed difficult to transpose the conclusions from one gene to another and even more from a model to another one. Then to obtain an optimal use of these data and since the Gapcoder analysis did not improve the resolution of the tree, further studies are now required. As shown in figure 6, a way to take into account the indel informations to improve artificially the tree resolution is to map the indels onto the phylogeny (Rokas *et al.*, 2000). By this way, we showed that even if the *G. tabacum* and *G. rostochiensis* species appeared separately grouped in the PL tree, they shared two specific indels of 78 and 20 bp. It seemed more probable that these indels appeared in a common ancestor of the two species rather than two times independently during the *Globodera* evolution (Rokas *et al.*, 2000). But if this hold true then the PL tree topology is doubtful and the PL gene may represent an inappropriate marker to recount the phylogenetic relationships of the Heteroderidae. This is not the case of the CathL and EF genes that could be considered as new markers to recount the evolutionary history of this nematode family. The parasitism gene CathL was a more suitable marker at the intra specific level than the EF gene that did not allowed a well supported identification of the relationships of the populations in the *G. pallida* species. Further investigations are underway to evaluate the relevant of these markers at larger scale.

## REFERENCES

- Barret AJ, Rawling ND, Woessner JF. 1998. Hand-book of proteolytic enzymes. Eds, Academic press, London.
- Bossis M, Mugniéry D. 1993. Specific status of six *Globodera* parasites of Solanaceous plants studied by means of two-dimensional gel electrophoresis with a comparison of gel patterns by a computed system. *Fundam Appl Nematol.* 16:47-56.
- Blumenthal T, Steward K. 1997. RNA processing and gene structure. In: Riddle, D.L., Blumenthal, T., Meyer, B.J., Priess, J.R. (eds), *C. elegans II*. Cold Spring Harbor Laboratory Press, pp. 117-145.
- Chen S L, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D, Bieri T, Meyer RR, Ozersky P, Armstrong JR, Fulton RS, Latreille JP, Spieth J, Hooton TM, Mardis ER, Hultgren SJ, Gordon JI. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. *Proc Natl Acad Sci.* 103:5977-5982.
- Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16: 10881-10890.
- Creer S, Pook CE, Malhotra A, Thorpe RS. 2006. Optimal intron analyses in the *Trimeresurus* Radiation of Asian Pitvipers. *Syst Biol.* 55:57-72.
- Davis EL, Hussey RS, Baum TJ. 2004. Getting to the roots of parasitism by nematodes. *Trends Parasitol.* 20:134-141.
- De Boer JM, McDermott JP, Davis EL, Hussey RS, Popeijus H, Smant G, Baum TJ. 2002. Cloning of a putative pectate lyase gene expressed in the subventral esophageal glands of *Heterodera glycines*. *J Nematol.* 34:9-11.
- Doyle EA, Lambert KN. 2002. Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. *Mol Plant-Microbe Interact.* 15:549-556.
- Friesen VL, Congdon BC, Walsh HE, Birt TP. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphism. *Mol Ecol.* 6:1047-1058.
- Gao BL, Allen KD, Davis EL, Baum TJ, Hussey RS. 2004. Developmental expression and biochemical properties of a beta-1,4-endoglucanase family in the soybean cyst nematode, *Heterodera glycines*. *Mol Plant Pathol.* 5:93-104.
- Goetze E. 2006. Elongation factor 1-[alpha] in marine copepods (Calanoida: Eucalanidae): Phylogenetic utility and unique intron structure. *Mol Phylogenet Evol.* 40:880-886.

**Gonzalez D, Cubeta MA, Vilgalys R.** 2006. Phylogenetic utility of indels within ribosomal DNA and [beta]-tubulin sequences from fungi in the *Rhizoctonia solani* species complex. *Mol Phylogenet and Evol.* 40:459-470.

**Grenier E, Bossis M, Fouville D, Renault L, Mugniéry D.** 2001. Molecular approaches to the taxonomic position of Peruvian potato cyst nematodes and gene pool similarities in indigenous and imported populations of *Globodera*. *Heredity.* 86:277-290.

**Knutsen AK, Torp M, Holst-Jensen A.** 2004. Phylogenetic analyses of the *Fusarium poae*, *Fusarium sporotrichioides* and *Fusarium langsethiae* species complex based on partial sequences of the translation elongation factor-1 alpha gene. *Int J Food Microbiol.* 95:287-295.

**Kumar S, Tamura K, Nei M.** 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform.* 5:150-163.

**Marché L, Valette S, Grenier E, Mugniéry D.** 2001. Intra-species DNA polymorphism in the tobacco cyst-nematode complex (*Globodera tabacum*) using AFLP. *Genome.* 44:941-946.

**Mitreva M, Wendl M, Martin J, Wylie T, Yin Y, Larson A, Parkinson J, Waterston R, McCarter J.** 2006. Codon usage patterns in Nematoda: analysis based on over 25 million codons in thirty-two species. *Genome Biol.* 7:R75.

**Neveu C, Abad P, Castagnone-Sereno P.** 2003a. Molecular cloning and characterization of an intestinal cathepsin L protease from the plant-parasitic nematode *Meloidogyne incognita*. *Physiol Mol Plant Pathol.* 63:159-165.

**Neveu C, Jaubert S, Abad P, Castagnone-Sereno P.** 2003b. A set of genes differentially expressed between avirulent and virulent *Meloidogyne incognita* Near-isogenic lines encodes secreted proteins. *Mol Plant-Microbe Interact.* 16:1077-1084.

**Nomura K, Melotto M, He SY.** 2005. Suppression of host defense in compatible plant-*Pseudomonas syringae* interactions. *Curr Opin Plant Biol.* 8:361-368.

**Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, Read TD, Dodson RJ, Umayam L, Brinkac LM, Beanan MJ, Daugherty SC, Deboy RT, Durkin AS, Kolonay JF, Madupu R, Nelson WC, Ayodeji B, Kraul M, Shetty J, Malek J, Van Aken SE, Riedmuller S, Tettelin H, Gill SR, White O, Salzberg SL, Hoover DL, Lindler LE, Halling SM, Boyle SM, Fraser CM.** 2002. The *Brucellus* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc Natl Acad Sci.* 99:13148-13153.

**Popeijus H, Overmars HA, Jones J, Blok VC, Goverse A, Helder J, Schots A, Bakker J, Smant G.** 2000. Degradation of plant cell walls by a nematode. *Nature.* 406:36-37.

- Posada D, Crandall KA.** 1998. Modeltest: testing the model of DNA substitution. *Bioinform.* 14:817-818.
- Rokas A, Kathirithamby J, Holland PW.** 1999. Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. *Insect Mol Biol.* 8:527-530.
- Rokas A, Holland PWH.** 2000. Rare genomic changes as a tool for phylogenetics. *Trends Ecol Evol.* 15:454-459.
- Rokas A, Nylander JAA, Ronquist F, Stone GN.** 2002. A Maximum-Likelihood Analysis of Eight Phylogenetic Markers in Gallwasps (Hymenoptera: Cynipidae): Implications for Insect Phylogenetic Studies. *Mol Phylogenet Evol.* 22:206-219.
- Sajid M, McKerrow JH.** 2002. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol.* 120:1-21.
- Sawires YS, Songer JG.** 2006. *Clostridium perfringens*: Insight into virulence evolution and population structure. *Anaerobe.* 12:23-43.
- Shevchik VE, Robert-Baudouy J, Hugouvieux-Cotte-Pattat N.** 1997. Pectate lyase Pell of *Erwinia chrysanthemi* 3937 belongs to a new family. *J Bacteriol.* 179:7321-7330.
- Shingles J, Lilley CJ, Atkinson HJ, Urwin PE.** 2006. *Meloidogyne incognita*: Molecular and biochemical characterisation of a cathepsin L cysteine proteinase and the effect on parasitism following RNAi. *Exp Parasitol* In Press, Corrected Proof.
- Suarez CE, Norimine J, Lacy P, McElwain TF.** 2006. Characterization and gene expression of *Babesia bovis* elongation factor-1[alpha]. *Int J Parasitol.* 36:965-973.
- Subbotin SA, Vierstraete A, De Ley P, Rowe J, Waeyenberge L, Moens M, Vanfleteren JR.** 2001. Phylogenetic Relationships within the Cyst-Forming Nematodes (Nematoda, Heteroderidae) Based on Analysis of Sequences from the ITS Regions of Ribosomal DNA. *Mol Phylogenet Evol.* 21:1-16.
- Swofford DL.** 1998. PAUP: phylogenetic analysis using parsimony. Version 4. Sinauer Associates, Sunderland, Mass. X.
- Wada H, Kobayashi M, Sato R, Satoh N, Miyasaka H, Shirayama Y.** 2002. Dynamic insertion-deletion of introns in deuterostome EF-1 alpha genes. *J Mol Evol.* 54:118-128.
- Weber E, Koebnik R.** 2006. Positive selection of the Hrp pilin HrpE of the plant pathogen *Xanthomonas*. *Journal of bacteriology.* 188:1405-1410.
- Young ND, Healy J.** 2003. GapCoder automates the use of indels characters in phylogenetic analysis. *BMC Bioinformatics* 4(6).

**Tsai YH, Orsi RH, Nightingale KK, Wiedmann M. 2006. *Listeria monocytogenes* internalins are highly diverse and evolved by recombination and positive selection. Infect Genet and Evol. 6:378-389.**

### 3. RESULTATS COMPLEMENTAIRES : CARACTERISATION DES GENES DU PARASITISME DANS LES 40 POPULATIONS DE *GLOBODERA* ET D'*HETERODERA*

#### 3.1. Amplification et caractérisation du gène cathepsine L

Une portion du gène cathepsine L (Fig. 10) a été amplifiée sur 33 des 40 populations initialement sélectionnées (tableau 1 de article 4 ; voir conditions d'amplification en annexe 2). Les amplifications ont été réalisées sur 1 kyste. Certaines amplifications, en particulier chez les *Heterodera*, se sont avérées impossible sur quelques populations, ceci étant probablement lié à la qualité des kystes et/ou à la quantité de matériel faible que représente 1 kyste (Gtt2, Hcaj, HS41vir, Htri, Hcic, Hsa, Hh). Une bande entre 2000 et 2500 pb a été amplifiée dans toutes les populations de *Globodera* excepté pour Gtt2.

Tous les produits PCR ont été purifiés sur Sephadex G50 et séquencés (Macrogen) en plusieurs fois. Les séquences obtenues étant de bonne qualité, elles ont été corrigées et alignées pour toutes les populations puis analysées. L'alignement des séquences obtenues avec la séquence du transcrit correspondant chez *G. pallida* (N° d'accèsion : AY999065) disponible dans les bases de données a permis d'identifier les introns (Figure 11). La description du polymorphisme des séquences est détaillée dans l'article 4. Si la répartition des introns tout au long de la séquence est la même chez les populations *Globodera* et *Heterodera*, la taille des introns est variable (Fig. 11). L'analyse des séquences nous a permis de déceler une erreur d'identification pour une des populations. La population GPS1 est en fait une population de *G. rostochiensis*. Cette première analyse a ensuite été confirmée par une analyse ITS (qui permet de discriminer *G. pallida* et *G. rostochiensis*). Ces données montrent la haute spécificité des séquences du gène cathepsine L à l'échelle de l'espèce.

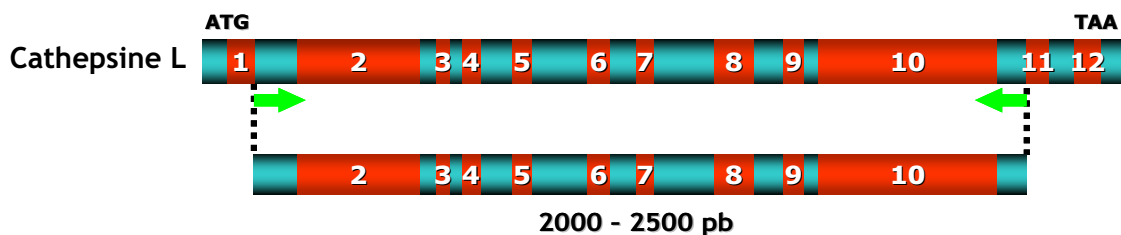


Figure 10 : Structure du gène cathepsine L et taille du produit PCR amplifié. Les flèches vertes symbolisent les amorces utilisées (voir annexes 1 & 2). Les exons sont en bleu et les introns en rouge.

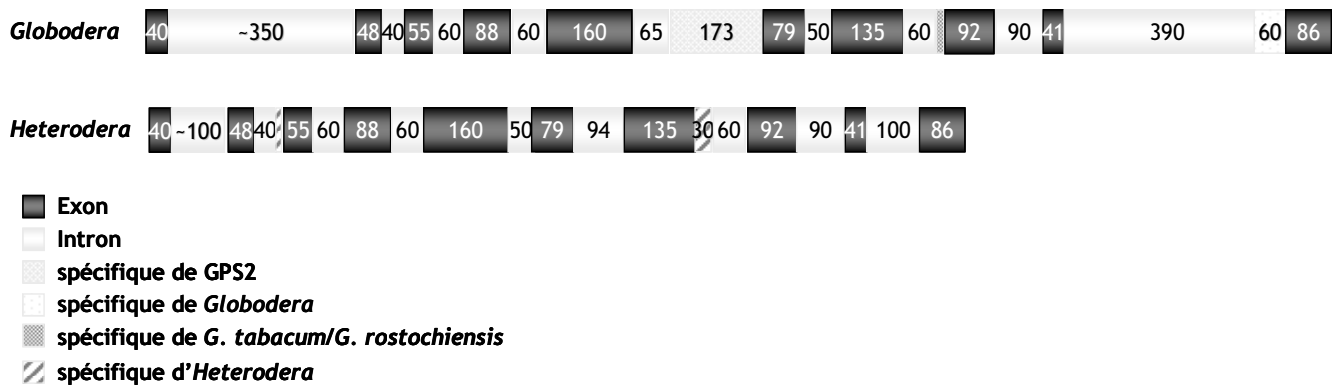


Figure 11 : Structure du gène cathepsine L dans les populations de *Globodera* et d'*Heterodera* sélectionnées. Les séquences des introns montrent des motifs spécifiques de certaines espèces. La taille des introns et des exons est indiquée en nucléotides.

L'alignement des séquences codantes montre que la structure de la protéine prédite est conservée au sein des populations et espèces de nématodes (Fig. 12) avec la zone dite "pro région" et la protéine mature, comme décrit par Silva *et al.* (2004). Deux indels apparaissent spécifiques des séquences d'*Heterodera* (Fig. 12) dans la zone pro région. Une délétion de 3 acides aminés est observée chez tous les *Heterodera* excepté pour la séquence *H. carotae* et une délétion ("R") est spécifique de tous les *Heterodera*. Une zone particulièrement variable a été identifiée dans la région mature de la protéine, spécifiquement chez les populations GPS3, GPS10 et GPS9 (Fig. 12). Les connaissances en matière de repliement de la protéine peut permettre d'accorder une importance variable aux mutations observées en fonction de leur position sur la protéine *in vivo*. Trois motifs conservés caractéristiques des cathepsines L (Barret *et al.*, 1998) ont été identifiés : "ERFNIN", "GNFD", "GCNGG" (Fig. 13). Certains acides aminés définis comme cruciaux pour la formation des sites actifs ont également été identifiés et sont conservés dans toutes les populations et espèces testées : Q172 et C178 (Barret *et al.*, 1998).

Résultats – Variabilité et évolution moléculaires des gènes du pouvoir pathogène



Figure 12 : Alignement protéique des séquences cathepsine L obtenues chez 33 populations de nématodes à kyste (Multaïn). Le domaine de la "pro région" est souligné, le reste de la protéine correspondant à la zone mature. Les indels spécifiques des séquences obtenues chez les *Heterodera* sont surlignées en vert et la zone de la protéine mature qui est plus variable chez les populations GPS3, GPS9 et GPS10 sont surlignées en jaune. Les positions des introns sont indiquées par des flèches noires.



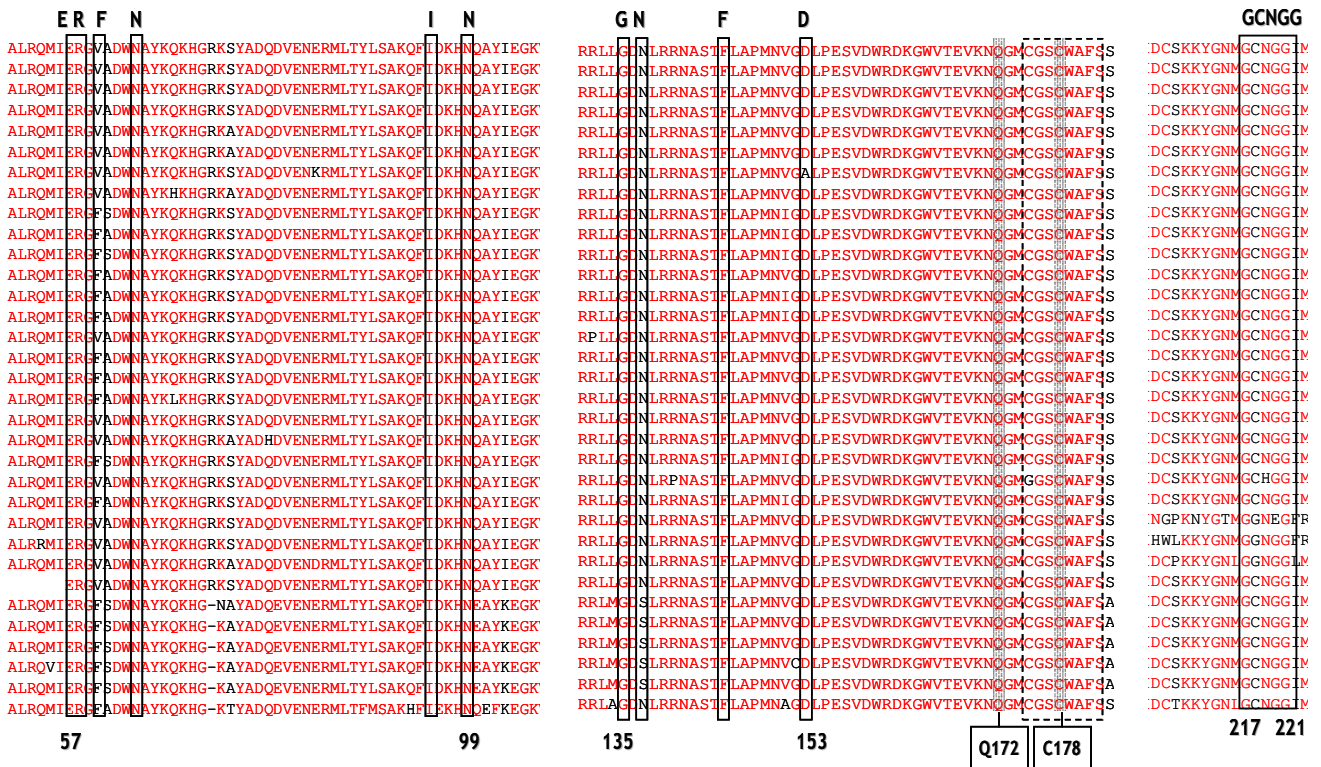


Figure 13 : Alignements des séquences protéiques partielles prédites à partir des données génomiques des gènes cathepsine L. Les motifs conservés sont indiqués en encadrés noirs. Deux acides aminés essentiels sont conservés dans les séquences dont nous disposons (Q172 et C178 ; Barret *et al.*, 1998). Le motif encadré en pointillé est très conservé dans toutes nos séquences, comme décrit par Sajid *et al.* (2002).

### 3.2. Amplification et caractérisation du gène pectate lyase

Une portion du gène pectate lyase (Fig. 14) a été amplifiée sur 28 populations des 40 populations initialement sélectionnées (tableau 1 de article 4 ; voir conditions d'amplification en annexe). Les amplifications chez les populations d'*Heterodera* se sont avérées relativement difficiles. Ceci étant probablement lié aux divergences de séquences entre les nématodes à kyste des genres *Globodera* et *Heterodera* et à la mauvaise qualité de certains kystes qui ne nous a pas permis d'amplifier ce gène sur de faibles quantités d'ADN. Ainsi, nous ne sommes parvenu à obtenir des séquences du gène pectate lyase que pour les populations : *H. glycines* et *H. schachtii* (HS41). Le produit amplifié fait environ 800 à 1100 pb selon les populations. Par la suite, afin d'obtenir des séquences plus longues mais au niveau transcriptomique, nous avons amplifié les transcrits pectate lyase sur 20 populations par RT-PCR. Les séquences amplifiées sont alors plus longues d'environ 310 pb par rapport à la zone codante obtenue lors des amplifications sur ADN génomique. Cette technique a également permis d'obtenir davantage de séquences chez les *Heterodera* (5).

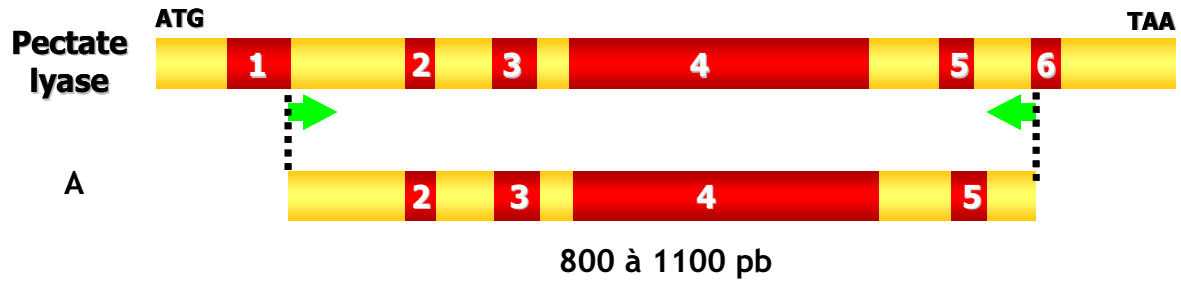


Figure 14 : Structure du gène pectate lyase et taille du produit PCR amplifié sur ADN génomique. Les flèches vertes symbolisent les amorces utilisées (voir annexes 1 & 2).

Le même nombre d'introns a été identifié chez toutes les populations testées et la même répartition de ces introns a été observée. Cependant, des variations de taille importantes ont pu être identifiées en particulier pour l'intron numéro 3 qui présente des séquences caractéristiques des différentes espèces utilisées (Fig. 15).

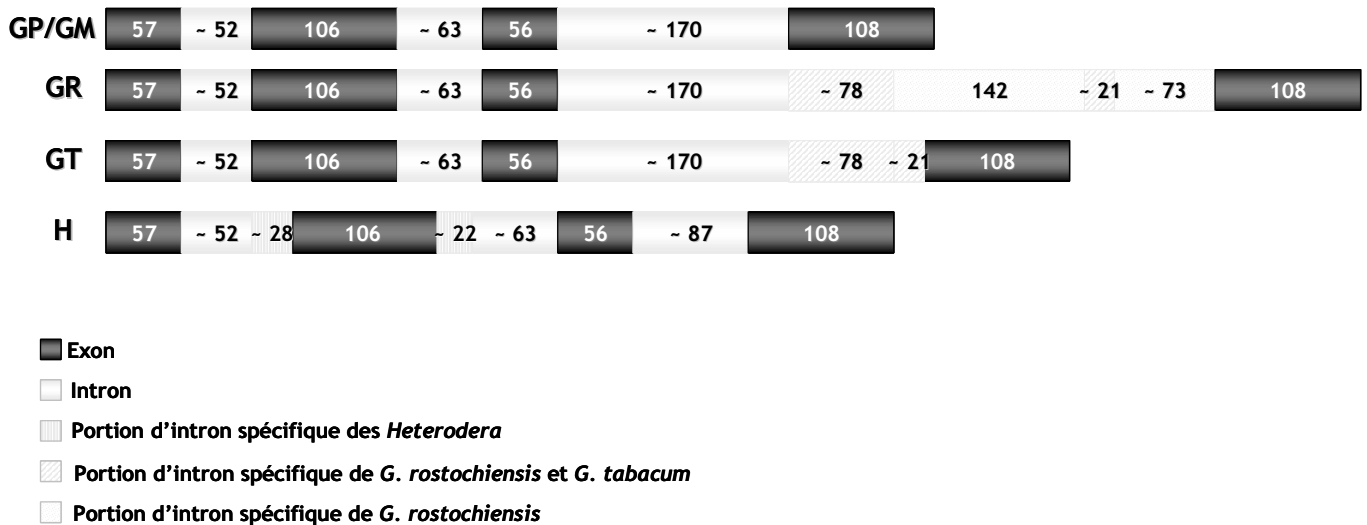


Figure 15 : Structure du gène pectate lyase dans les populations de *Globodera* et d'*Heterodera* sélectionnées. GP : *G. pallida* ; GM : *G. "mexicana"* ; GR : *G. rostochiensis* ; GT : *G. tabacum* ; H : *Heterodera*. La taille des introns et des exons est indiquée en nucléotides.

Les séquences du gène pectate lyase amplifiées présentent toutes les caractéristiques des pectate lyase de classe III : 4 régions conservées et de nombreux résidus cystéines, tel que décrit par Shevichik *et al.*, 1997 (Fig. 16). Ces régions ne semblent pourtant pas plus conservées que le reste de la séquence si l'on compare les séquences en acide aminés déduites des séquences de transcrits obtenues chez les nématodes des genres *Globodera* et *Heterodera* (Fig. 16). Ce jeu de

données nous a permis d'identifier une insertion d'un acide aminé dans les séquences spécifiques des populations du genre *Heterodera*. Nous avons également identifié une insertion de 3 acides aminés, "NSS", dans la séquence pectate lyase de la population HS41 (*H. schachtii* qui est avirulente contre *HS1<sup>pro1</sup>*) alors que cette insertion est absente de la population virulente correspondante. Cependant, cette insertion est également absente de toutes les autres populations d'*Heterodera* testées.

Concernant la pectate lyase, aucune structure tridimensionnelle n'est actuellement disponible dans les bases de données. Nous ne sommes donc pas en mesure d'estimer l'impact de ces insertions dans la séquence codante sur le repliement ou l'activité de la protéine.

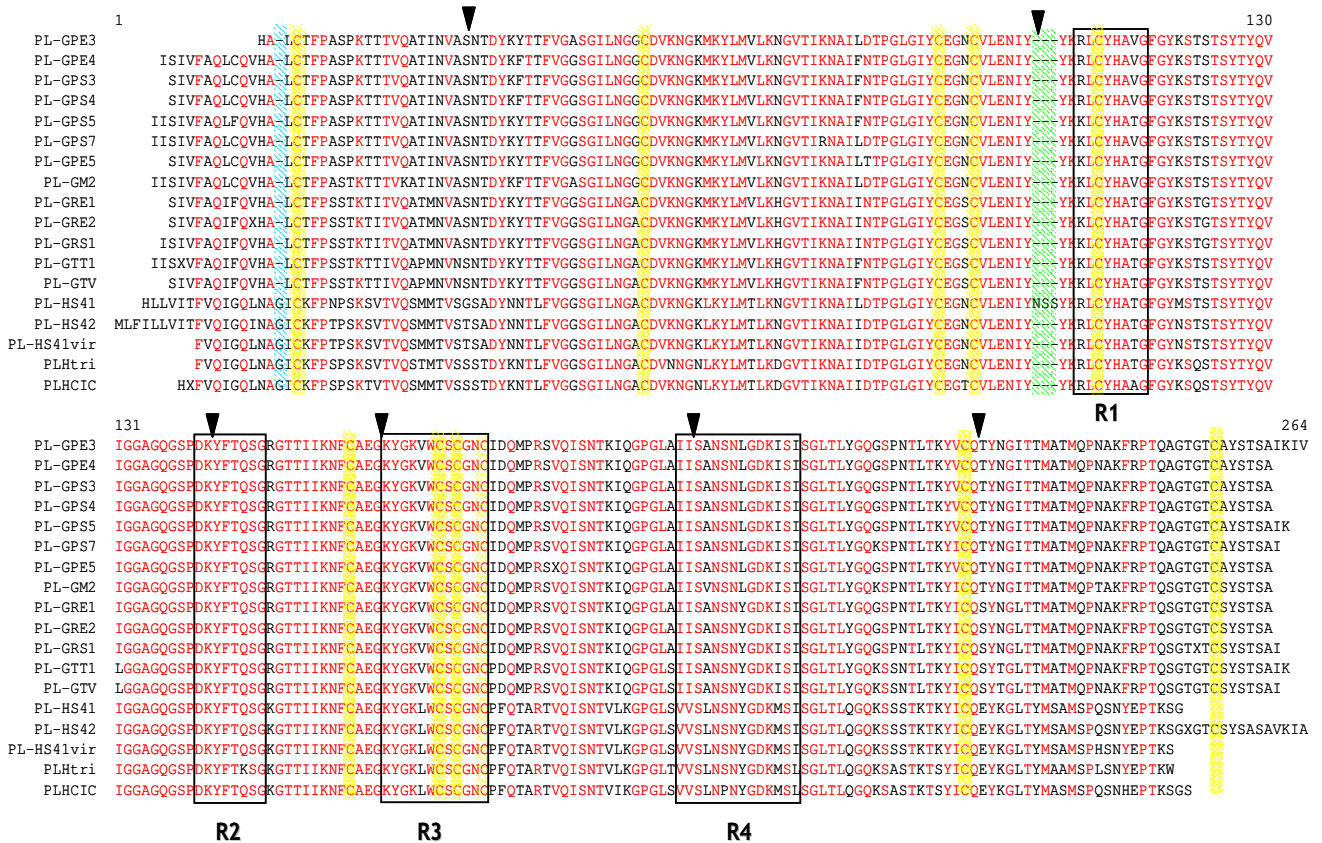


Figure 16 : Alignement en acide aminés des séquences de pectate lyase amplifiées sur ADNc de populations de *Globodera* et d'*Heterodera* représentant une portion du transcrit pleine longueur. En jaune sont surlignés les résidus cystéine (11), caractéristiques des pectates lyases de classe III. Les encadrés délimitent les 4 régions conservées (R1 à R4) également caractéristiques des pectate lyase de classe III. En vert est surlignée l'insertion spécifique de la population HS41 avirulente contre le gène de résistance *HS1<sup>pro1</sup>* de la betterave et en bleu l'insertion spécifique des *Heterodera*. La position des introns est spécifiée par une flèche noire.

# CHAPITRE 2

---

## CHAPITRE 2 :

Nous allons détailler dans ce dernier chapitre les toutes dernières analyses réalisées sur les trois jeux de données (EF, CathL et PL) afin d'aller plus loin dans l'interprétation des observations de l'article 4 et de tenter de définir les pressions de sélections qui ont agi sur ces gènes au cours de l'évolution des Heteroderidae.

### 1. ELEMENTS DE CONTEXTE : CONTRAINTES SELECTIVES DE LA RELATION HOTE/PATHOGENE

Les interactions entre un pathogène et son hôte sont souvent très étroite et très élaborées. Chez les deux partenaires, des gènes sont spécifiquement impliqués dans ces relations et sont soumis perpétuellement à des pressions évolutives (Fig. 17). Les plantes ont évolué en interaction avec les microbes épiphytes, symbiotiques et pathogènes et possèdent ainsi des récepteurs de surface qui reconnaissent des molécules produites par les pathogènes, les PAMPS (Pathogen-Associated Molecular Patterns). Cette reconnaissance très spécifique permet à la plante de mettre en place des mécanismes de défense contre les attaques pathogènes. Les parasites quant à eux ont évolué pour mettre en place des stratégies de survie en supprimant notamment les défenses de l'hôte grâce à des mécanismes d'interférence dans la reconnaissance ou encore la sécrétion de molécules bloquant la transduction du signal de défense (Chisholm *et al.*, 2006) (Fig. 18).

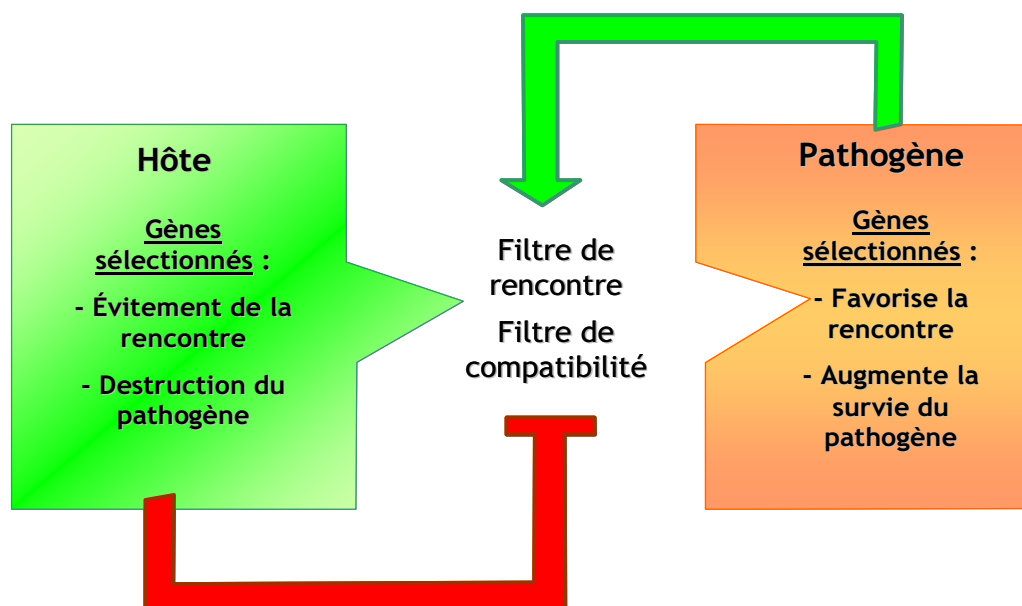


Figure 17 : Les interactions hôte-parasites : mécanismes évolutifs (d'après Combes, 2000).

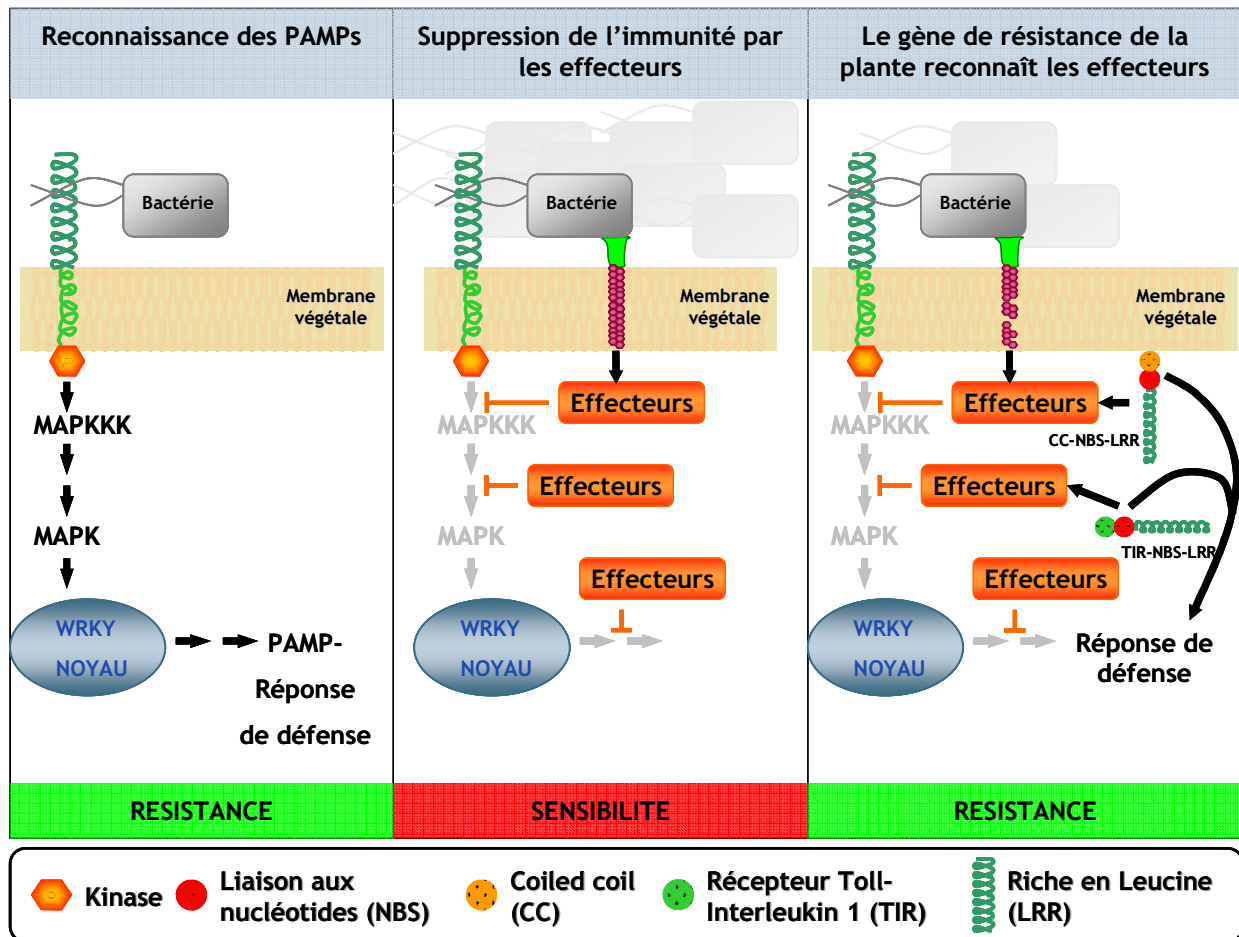


Figure 18 : Modèle d'évolution de la résistance dans le cas des bactéries (d'après Chisholm *et al.*, 2006).

### 1.1. Evolution des gènes du pouvoir pathogène

Les pathogènes utilisent différentes stratégies pour contourner les défenses constitutives et induites des plantes, incluant la dégradation des composés anti-microbiens et la production de molécules supprimant l'induction des mécanismes de défense des plantes. C'est le cas d'enzymes de détoxification produite par un champignon capable d'hydrolyser tant les saponines anti-microbiennes que les phytoalexines produites par la plante pour se défendre (Bouarab *et al.*, 2002 ; George *et al.*, 2001). Dans le premier article de la partie 2, nous avons évoqué une éventuelle adaptation moléculaire des parasites en réponse aux défenses mises en place par la plante hôte. En effet, la relation souvent très étroite qui s'est établie entre la plante et son bio-agresseur fait que chacun des partenaires exerce une pression sélective sur l'hôte, c'est la coévolution évoquée par Van Valen en 1973 sous le nom d'hypothèse de la reine rouge. L'hôte et son parasite évoluent probablement de façon séquentielle chacun leur tour de manière à répondre aux adaptations de l'autre sans qu'aucun des deux ne prenne l'avantage ou alors seulement de façon transitoire. Dans les cas d'adaptation du parasite à l'hôte, il a été montré à plusieurs reprises des signes forts de sélection positive sur les gènes directement impliqués dans l'interaction (Weber *et al.*, 2006 ; Chen

*et al.*, 2006 ; Sawires *et al.*, 2006 ; Tsai *et al.*, 2006). Ainsi les gènes du pouvoir pathogène seraient soumis à des pressions de sélection diversificatrices. De nombreuses méthodes sont disponibles pour étudier la sélection au niveau moléculaire mais elles présentent certaines limites du fait que la sélection n'agit sans doute pas de la même façon sur tous les sites d'un gène ni sur le gène complet, en continu dans toutes les lignées. Ces différents paramètres rendent l'évaluation des pressions de sélection délicate. C'est pourquoi il apparaît aussi nécessaire de caractériser les gènes sur lesquels il est intéressant de rechercher de la sélection positive au niveau de leur séquence (variabilité nucléotidique) et au niveau de leur structure (identification des domaines fonctionnels, de la structure tridimensionnelle, de sites actifs).

## **1.2. Pressions de sélection exercées par la plante hôte**

Si les gènes du pouvoir pathogène des parasites évoluent plus rapidement que des gènes non liés au parasitisme, tels que les gènes de ménage, c'est parce qu'ils sont soumis à des pressions particulières. La plante, lorsqu'elle reconnaît les molécules libérées par le parasite, exerce une pression de sélection sur celui-ci. Deux solutions s'offrent alors au pathogène : soit il évolue pour s'adapter en contournant les défenses de l'hôte, soit il meurt. Devant les multiples attaques de pathogènes, les plantes ont créé pour se défendre, un système de défense composé d'un éventail important de gènes de défense et de résistance - près de 600 identifiés chez *Arabidopsis thaliana* (Jones & Takemoto, 2004) - capables de reconnaître un large spectre de protéines exogènes. En effet, la majorité des gènes de résistance des plantes ont une structure de type NBS-LRR (Fig. 19) dont les domaines LRR sont très variables et confèrent la spécificité de reconnaissance du pathogène (Baker *et al.*, 1997). On peut supposer que face à la grande diversité de ce type de gènes chez les plantes, les parasites doivent eux aussi avoir une possibilité d'échapper à cette reconnaissance et ceci pourrait passer par le biais de mutations, même ponctuelles. Un seul changement en acide aminé peut être suffisant, dans un domaine d'interaction pour inhiber la reconnaissance entre les gènes de pathogénicité et les gènes de reconnaissance de l'hôte (Pouttu *et al.*, 1999). Ces gènes sont donc soumis à des contraintes sélectives qui font que s'ils n'évoluent pas, les organismes qui les portent perdent leur capacité à se développer dans leur hôte. De tels gènes sont donc logiquement soumis à pression de sélection diversificatrice ou positive. Cependant, les pathogènes possédant plusieurs copies fonctionnelles d'un même gène impliqué dans le parasitisme peuvent n'avoir évolué que sur une copie de ce gène alors que les autres copies ne seraient pas contraintes. On parle alors de réservoir de gènes du pouvoir pathogène (Lambert *et al.*, 2005) qui permet au parasite de s'adapter rapidement aux plantes.



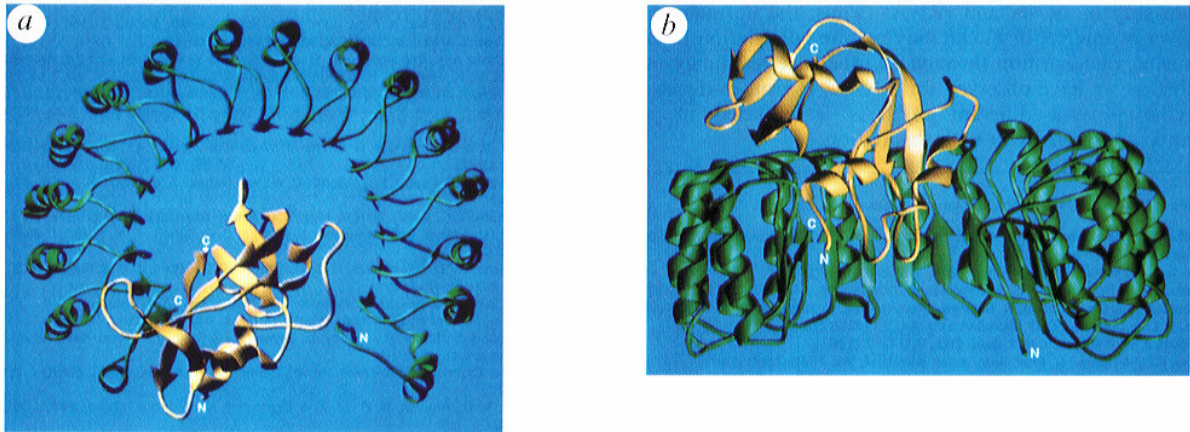


Figure 19 : Structure d'un gène de résistances de type NBS-LRR, d'après Baker *et al.* (1997). En jaune est représenté le domaine variable LRR (Leucin Rich Repeat) et en vert, le domaine plus conservé de type NBS (Nucleotid Binding Site).

## 2. ARTICLE 5 : ANALYSES DES PRESSIONS DE SELECTION

Les résultats présentés dans cette partie sont les dernières données acquises qui feront prochainement l'objet d'un nouvel article.

### INTRODUCTION

Selon la théorie neutraliste de l'évolution moléculaire, les mutations qui apparaissent sont majoritairement neutres, c'est-à-dire qu'elles n'ont pas d'effet sur la fonction de la protéine ou le phénotype en général (Kimura, 1968; Kimura, 1983). Les gènes en règle générale sont pour autant soumis à deux grands types de pressions de sélection : 1- diversificatrices (sélection positive), qui tendent à générer de la diversité ou 2- purificatrices qui tendent à contraindre l'évolution moléculaire vers la conservation des acides aminés. Nous pouvons considérer que le niveau évolutif de base chez les nématodes est représenté par le gène de ménage qui ne subit aucune pression sélective directement liée aux besoins parasitaires des nématodes à kyste. Lorsque nous comparons ce niveau avec celui observé pour les gènes liés au pouvoir pathogène, nous constatons que ces gènes semblent être soumis à pressions diversificatrices par rapport au gène de ménage du fait que près de la moitié des mutations se situent en première et seconde position dans les codons. Ces mutations correspondent majoritairement à des substitutions non synonymes qui vont engendrer des modifications protéiques. Nous pouvons alors penser que ces gènes ont subi une accélération du tempo d'évolution par accumulation de mutations au sein de la famille des Heteroderidae. Ceci laisse supposer que les gènes pectate lyase et cathepsine L sont soumis à sélection positive sur certaines régions du gène ou dans certaines lignées.



Les distances génétiques distribuées de façon bimodale au sein de la famille des Heteroderidae ainsi que la présence d'un "gap" dans la distribution, présenté dans l'article 4 sont également une indication d'accélération du tempo d'évolution. En effet, lorsque nous comparons les distances génétiques entre *Globodera* et *Heterodera*, elles deviennent très éloignées de celles que l'on peut trouver en comparant les espèces à l'intérieur des genres, sans qu'il y ait de valeurs intermédiaires comme cela a été observé pour le gène de ménage.

Ces deux types de données indiquent une évolution moléculaire différente entre les gènes liés au parasitisme et le gène de ménage et permet donc de faire l'hypothèse de pressions de sélection différentielles pour ces deux types de gènes. Cependant, il semble peu envisageable que la sélection ait agi de la même façon sur tous les sites d'un gène, en continu dans toutes les lignées (Chen *et al.*, 2006). Il est donc nécessaire d'identifier des zones et/ou des lignées qui seraient préférentiellement soumises à sélection diversificatrice ou purificatrice. En fonction des données de structure des protéines prédites dont nous disposons, nous avons tout d'abord étudié la répartition des mutations en fonction de la position dans les codons, le long de la séquence nucléotidique pour les trois gènes, puis recherché des positions protéiques ainsi que des lignées soumises à la sélection positive.

## RESULTATS

### 1. Répartition des mutations pour chaque gène, le long de la séquence amplifiée en fonction de la position dans le codon

#### 1.1. Le gène *cathepsine L*

La répartition des sites variables (position nucléotidique présentant au moins une mutation au sein du jeu de données) est relativement homogène entre les positions et également le long de la séquence nucléotidique (Fig. 20). Même si la troisième position de codon présente le plus de mutations, les premières et deuxième positions sont aussi des zones qui subissent de nombreuses mutations tout le long de la séquence.

Une zone, encadrée en rouge sur la figure, semble présenter davantage de mutations si on cumule les données des trois positions par rapport au reste de la séquence. Cette zone se trouve dans la région mature de la protéine. Si les deux premières positions de codon ne sont que très rarement mutées au maximum du possible dans les fenêtres de neuf nucléotides, ceci est très fréquent sur la troisième position de codon qui atteint trois sites mutés sur trois à de nombreuses reprises. Aucun type de mutation ne semble spécifiquement être lié aux "points chauds" que constituent les positions d'introns.

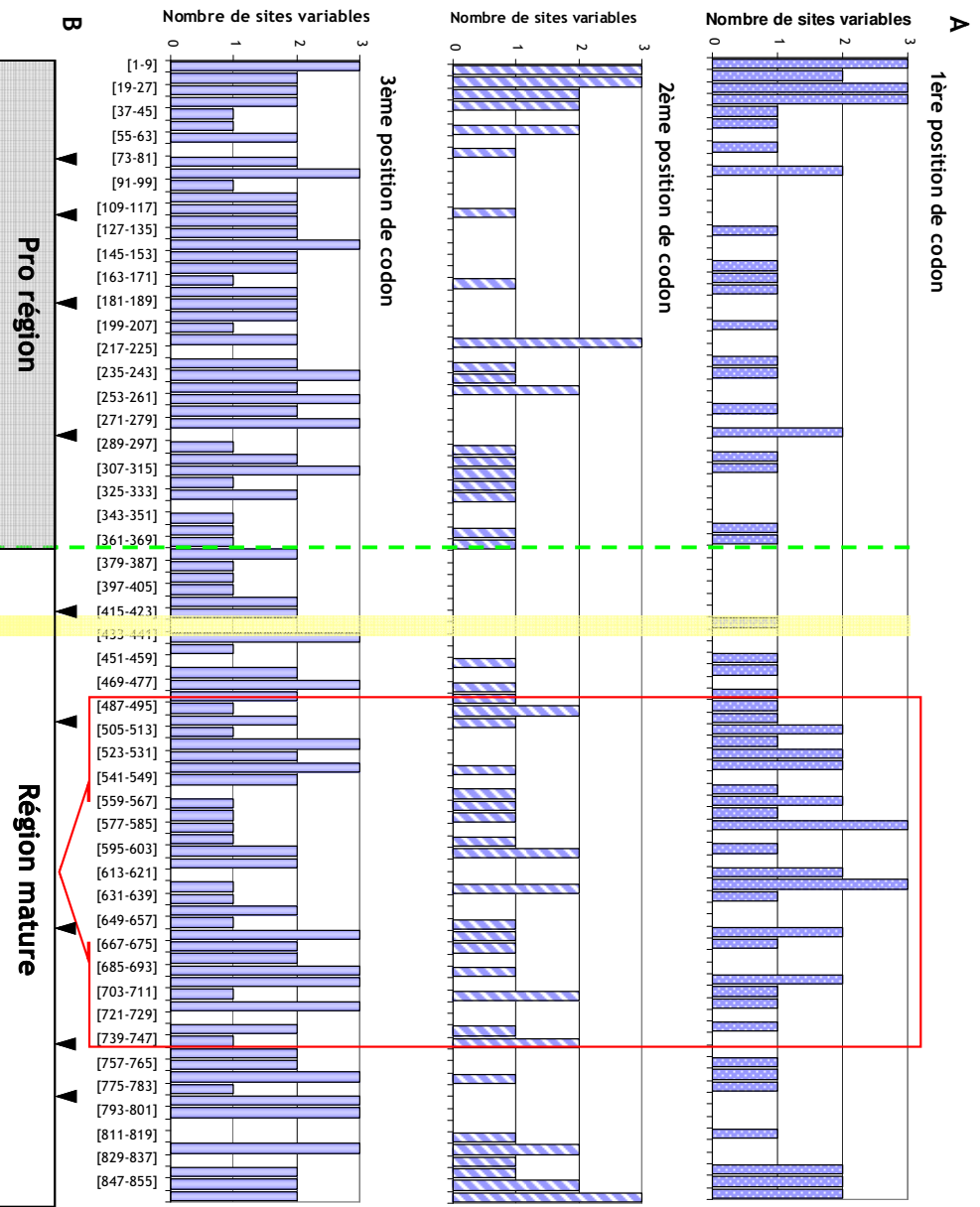


Figure 20 : A : Répartition des mutations en 1<sup>ère</sup>, 2<sup>ème</sup> ou 3<sup>ème</sup> position de codon le long de la zone codante du fragment du gène cathepsine L amplifié. En abscisse est indiquée la taille de la séquence en classe de 9 nucléotides soit 3 codons. Ainsi, le nombre maximal de mutation pour chaque position, dans une classe est de 3. B : Schéma du gène cathepsine L présentant les zones spécifiques des cathepsines de type L. En jaune est représentée la zone très conservée "CGSCWAFS" et l'encadré rouge délimite au contraire une zone particulièrement variable. Les flèches indiquent la position des introns. Les pointillés verts délimitent la pro région.

### 1.2. Le gène *pectate lyase*

Pour le gène *pectate lyase*, les mutations les plus fréquentes sont clairement localisées en troisième position de codon (Fig. 21). Cependant, un taux élevé de mutations est observé en première et deuxième base, avec une fréquence plus importante en première position. Les régions spécifiques des *pectates lyases* de classe III, R1 à R4 ne présentent pas moins de sites variables que le reste de la séquence contrairement à ce qui était attendu. Ce ne sont donc pas réellement des régions conservées, mais probablement plus des régions présentant des motifs/domaines communs à toutes les *pectates lyases* de cette classe. Trois zones de trois acides aminés ont été identifiées totalement conservées entre les genres et espèces (Flèches violettes sur la figure 21) et deux

supplémentaires ne présentent de mutations qu'en première position de codon, indiquées par des flèches bleues (Fig. 21). Ce sont des régions du gène pectate lyase particulièrement conservées chez les nématodes à kyste. De même que pour la cathesine L, les mutations en troisièmes sites sont plus fréquentes que sur les autres positions de codon, avec un maximum de mutations par fenêtre de 9 nucléotides atteint plusieurs fois le long de la séquence, indiquant une saturation du jeu de données sur cette position.

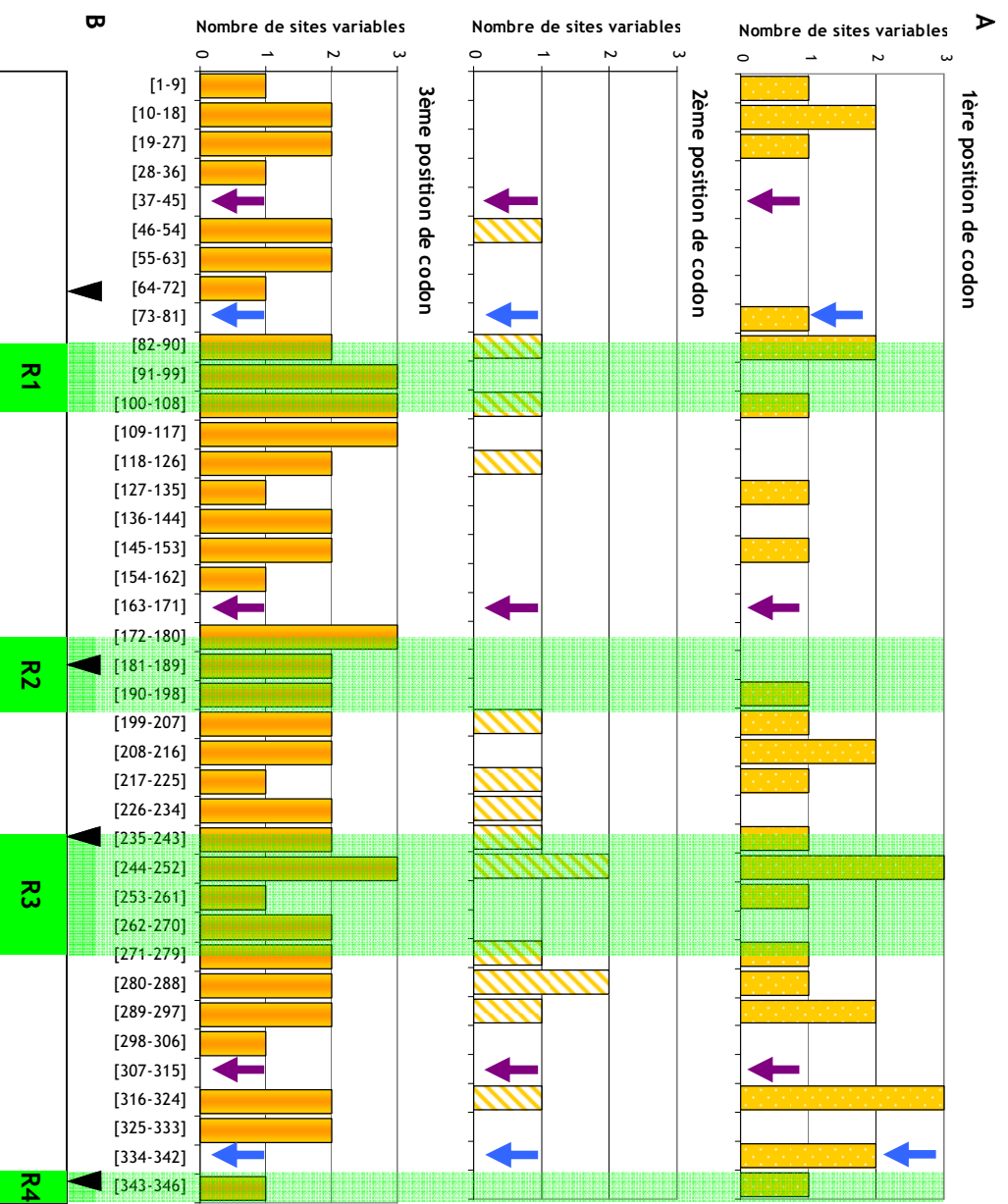


Figure 21 : A : Répartition des mutations en 1<sup>ère</sup>, 2<sup>ème</sup> ou 3<sup>ème</sup> position de codon le long de la zone codante du fragment du gène pectate lyase amplifié. En abscisse est indiquée la taille de la séquence nucléotidique en classe de 9 nucléotides soit 3 codons. B : Schéma du gène pectate lyase présentant les zones spécifiques des pectate lyase de classe III. Sont représentées en vert les régions R1 à R4 spécifiques des pectates lyases de classe III (Shevchik *et al.*, 1997) (le domaine R4 est incomplet). Les flèches indiquent la position des introns. Les flèches violettes indiquent les fenêtres de 9 nucléotides pour lesquelles aucune mutation n'est observée, et les flèches bleues les fenêtres ne présentant qu'une ou deux mutations en première position de codon. Ces zones sont considérées comme des régions particulièrement conservées.

### 1.3. Le gène facteur d'élongation 1 $\alpha$

Très clairement, pour le gène de ménage, les sites où les mutations sont accumulées sont les troisièmes positions de codon (Fig. 22). Sur 430pb, seuls 5 sites en deuxième position et 13 sites en première position ont été identifiés comme site variable avec le plus souvent 1 seul site variable par fenêtre de 9 nucléotides. Ce gène semble très nettement n'avoir que très peu de sites sur lesquelles les mutations sont possibles indiquant l'action de pressions de sélections de purificatrice sur ce gène de ménage, c'est à dire qui contre sélectionnent les individus ayant accumulé des mutations sur ces positions là. Ceci suggère également que ce gène a un rôle tel que toute mutation dans la séquence nucléotidique entraînant un changement d'acide aminé peut être défavorable au nématode.

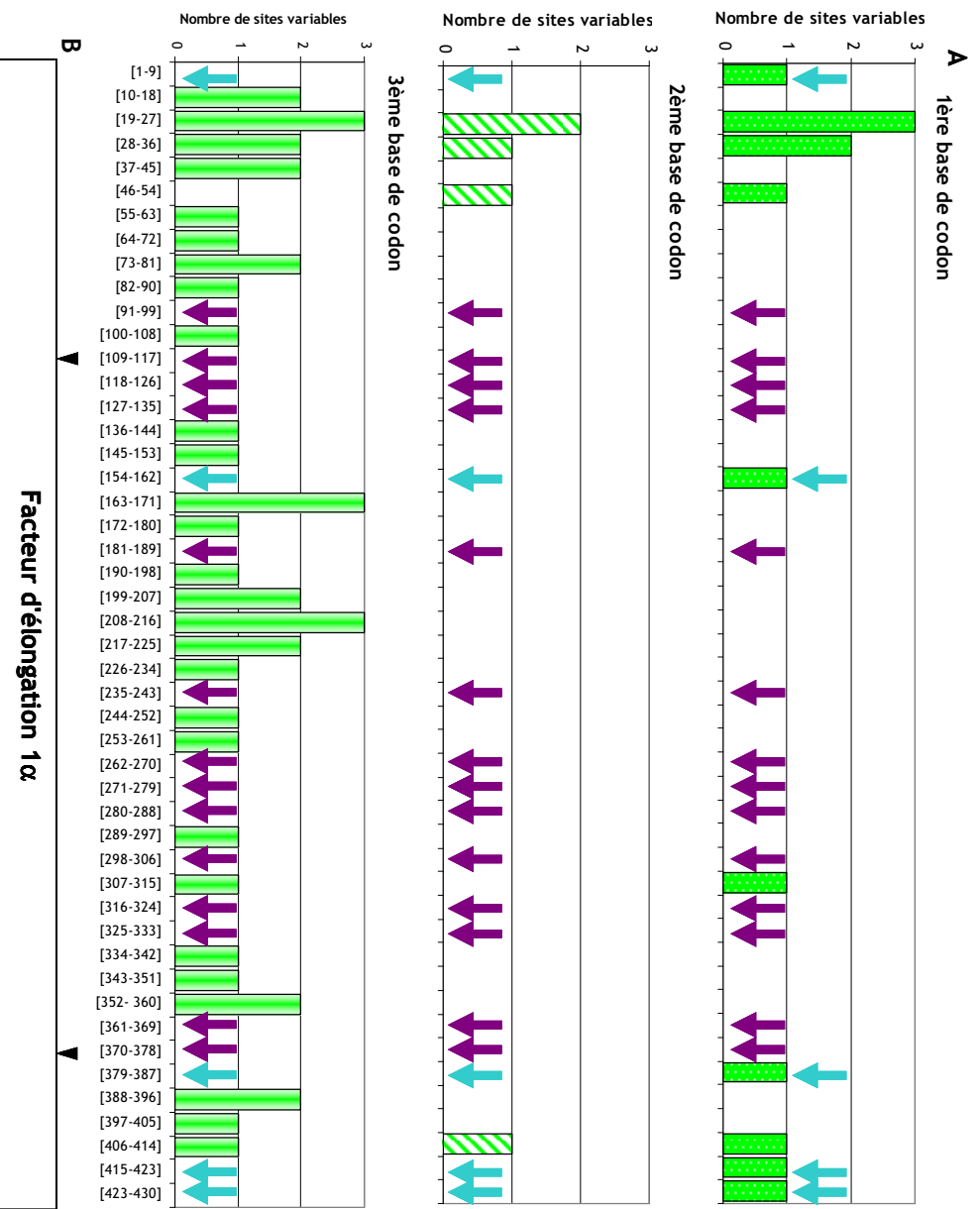


Figure 22 : A : Répartition des mutations en 1<sup>ère</sup>, 2<sup>ème</sup> ou 3<sup>ème</sup> position de codon le long de la zone codante du fragment du gène facteur d'élongation 1 $\alpha$  amplifié. En abscisse est indiquée la taille de la séquence en classe de 9 nucléotides soit 3 codons. B : Schéma du gène facteur d'élongation 1 $\alpha$ . Les flèches représentent la position des introns. Les flèches violettes indiquent les fenêtres de 9 nucléotides pour lesquelles aucune mutation n'est observée, et les flèches bleues les fenêtres ne présentant qu'une ou deux mutations en première position de codon.

## **2. Identification de la sélection agissant sur les gènes du pouvoir pathogène et sur le gène de ménage**

Pourquoi chercher à identifier les pressions de sélection des gènes du pouvoir pathogène ? Les gènes impliqués dans l'adaptation à un nouvel habitat, une nouvelle niche écologique ou dans notre cas à une nouvelle plante est une cause de l'évolution des gènes notamment par le biais de la sélection positive qui tend à générer davantage de diversité et donc de possibilités d'adaptation. En effet, un taux de mutation élevé confère un avantage de fitness : les mutations persistent plus longtemps et les individus mutés semblent être plus virulents (Chen *et al.* 2006). Ainsi les gènes associés à l'adaptabilité sont souvent soumis à sélection positive (Sawires *et al.* 2006). Weber *et al.* (2006) ont montré que certaines protéines extracellulaires de pathogènes responsables de l'adaptation évolutive sont sous sélection positive pour échapper au système de surveillance des plantes et éviter la mise en place de leur système de défense. Ainsi, la détection de sélection positive sur les gènes du pouvoir pathogènes serait :

1- un argument supplémentaire pour appuyer l'implication de ces gènes dans la relation hôte/parasite, ce qui n'a pas encore été clairement montré,

2- un moyen de détecter des zones de la protéine correspondante qu'il faudrait éviter pour générer des résistances artificielles du fait de leur fort potentiel à muter pour permettre l'adaptation du parasite à diverses conditions,

3- un moyen de différencier des évolutions de séquences dans les différentes populations et espèces afin de prédire les capacités d'adaptation d'une population, d'une espèce ou d'un genre aux conditions extérieures.

Ainsi, dans un second temps, nous avons initié des études de recherche de sélection positive dans deux dimensions, le long des séquences et dans les différentes populations et espèces, afin de préciser la nature de la sélection agissant sur les gènes du pouvoir pathogène ainsi que les zones protéiques d'action de ces pressions.

Les types de sélections sont définis en fonction du ratio  $Ka/KS (=dN/dS)$ ,  $Ka$  et  $dN$  étant le taux de substitutions non synonymes et  $KS$  et  $dS$  le taux de substitutions synonymes. Lorsque ce ratio est inférieur à 1, la sélection est purificatrice, lorsqu'il est égal un 1, la sélection est neutre et lorsqu'il est supérieur à 1, la sélection est diversifiante ou positive (Yang *et al.*, 1997). Nous avons alors cherché à définir les différents types de pressions sélection agissant le long des séquences en faisant l'hypothèse, d'après les résultats précédents, qu'elles ne sont pas identiques sur tout le gène, comme ce qui a été montré à plusieurs reprises chez d'autres pathogènes (Chen *et al.*, 2006).

### **2.1. Analyse des ratios des taux de substitution non synonyme vs synonyme par fenêtre glissante ou "sliding window"**

Nous avons recherché pour chaque gène à localiser les régions qui pourraient être soumises à sélection positive. Pour cela, nous avons utilisé le logiciel DnaSP 4.0 (Rozas *et al.*, 2003) pour analyser en "fenêtre glissante" (sliding window) la répartition des ratios Ka/Ks (substitutions non synonymes/synonymes) le long des séquences nucléotidiques. Les résultats de l'analyse sont présentés sur la figure 23. Les ratios sont nettement inférieurs pour le gène de ménage ( $< 0,25$ ) que pour les gènes du parasitisme ( $< 3,5$ ). Les valeurs globales calculées pour le facteur d'élongation étant très inférieur à 1, valeur à laquelle on définit la sélection comme neutre, le gène est alors considéré comme soumis à sélection purifiante, d'autant que la majorité des ratios se situent au-dessus de 0,05. Même si des augmentations brutales sont observées (Fig. 23C), elles restent inférieures à 0,25. Les ratios calculés le long des séquences des gènes du pouvoir pathogène sont très différents de ceux calculés pour le gène de ménage. Par contre les profils entre les gènes pectate lyase et cathepsine L sont similaires en terme de valeur. En effet, la majorité des ratios se situent entre 0,5 et 1 indiquant une sélection plutôt neutre sur une grande partie des deux séquences. Certains sites semblent être soumis à pression diversificatrice (pics  $> 1$ , Fig. 23 A & B). Deux pics ont été identifiés pour le gène cathepsine et un pour le gène pectate lyase avec des valeurs de Ka/Ks compris entre 1,25 et 1,75 pour les deux pics cathepsine et avoisinant les 3 pour la pectate lyase. Concernant la cathepsine, si le nombre de sites variables ne semblait pas défini par les régions du gène (pro région vs région mature), les ratios montrent que les substitutions non synonymes sont plus nombreuses dans la région mature que dans la pro région, à l'image de ce qui avait été décrit par Silva *et al.* (2004), même si les ratios restent faibles. Les deux pics indiquant des zones sous sélection positive sont situés dans la partie codant la protéine mature de la cathepsine L, là où avait été précédemment localisée la zone la plus variable. Pour la pectate lyase, aucun ratio spécifique n'a été mis en relation avec la présence des régions conservées de la classe III. Le seul pic indiquant de la sélection positive sur le gène pectate lyase se trouve dans la région R3 (Fig. 23B).

Ces analyses confirment que certaines régions des protéines codées par les gènes cathepsine L et pectate lyase sont soumises à sélection positive. Nous avons donc cherché à déterminer plus précisément l'action de cette sélection au niveau des différentes populations et espèces, et également à identifier les acides aminés concernés par cette sélection positive.

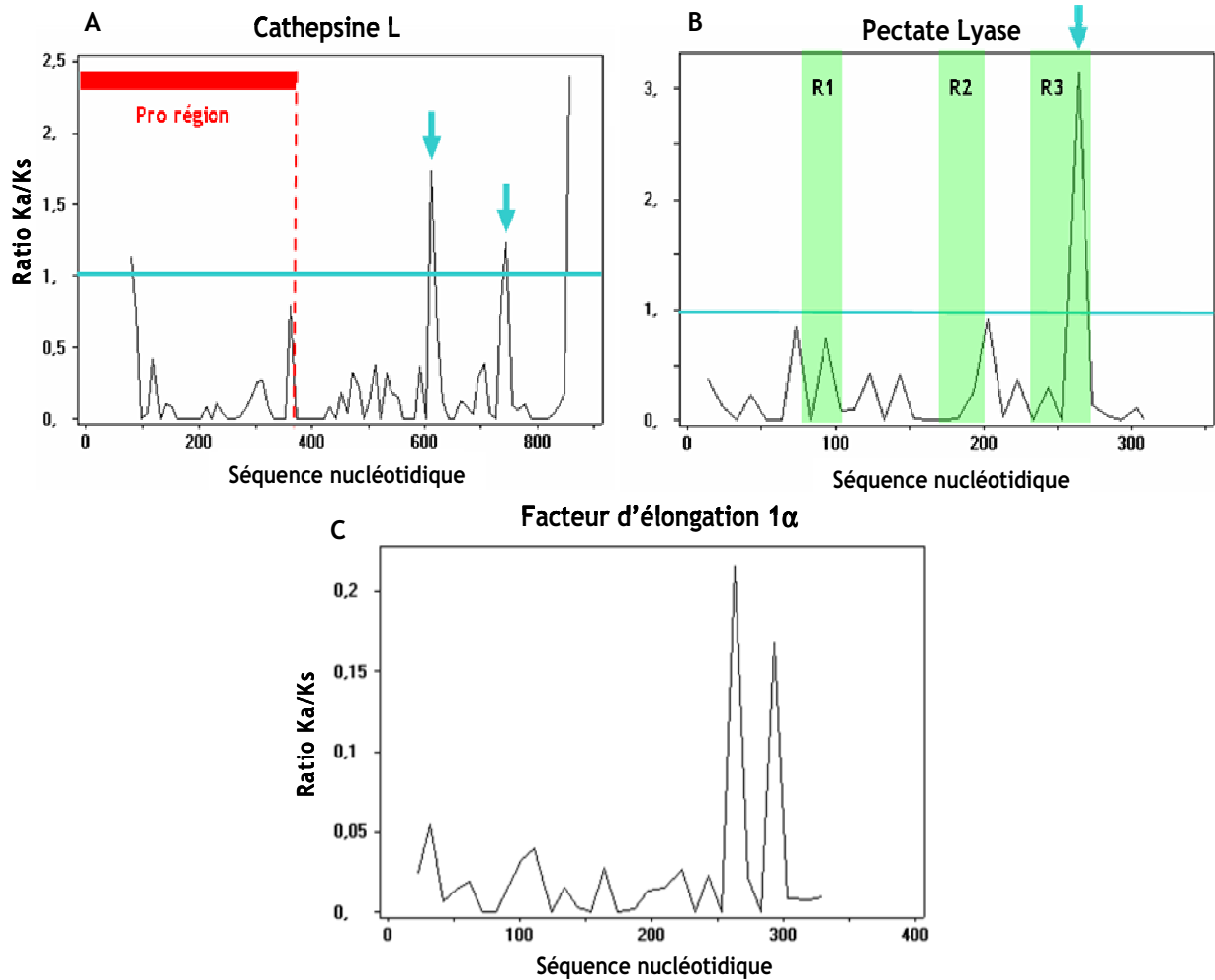


Figure 23 : Analyse en fenêtre glissante de l'alignement nucléotidique du gène cathepsine L (A), pectate lyase (B) et facteur d'élongation (C) chez toutes les populations testées par utilisation du logiciel DnaSP 4.0 (Rozas *et al.*, 2003). Le graphique représente en ordonnée les ratios Ka/KS et en abscisse les positions nucléotidiques. Les flèches bleues indiquent des sites *a priori* soumis à sélection positive ( $Ka/Ks > 1$ ). Les régions caractéristiques des pectates lyases de classe III sont représentées en vert.

## 2.2. Analyses des taux de substitutions différentiels entre les sites et les espèces pour les gènes du pouvoir pathogène : recherche de sélection positive

Différentes méthodes d'analyse de la sélection sont disponibles mais peu prennent en compte l'hétérogénéité des pressions sur les différents sites d'un gène ou au sein des différentes lignées étudiées. La méthode la plus utilisée et qui semble être la plus puissante est l'analyse de la sélection par maximum de vraisemblance (ML). Pour cela, le logiciel PAML (Phylogenetic Analysis by Maximum Likelihood) développé par Ziheng Yang (1997) a été utilisé. Les résultats des analyses préliminaires réalisées avec le module *codeml* sont présentés ci-dessous.

### A. Détection de la sélection positive pour le gène cathepsine L

L'utilisation du module *codeml* nous a permis de rechercher les sites préférentiellement soumis à sélection positive dans tout le jeu de données. Ce module permet de réaliser une analyse en maximum de vraisemblance des séquences nucléotidiques codantes en utilisant le modèle de substitution des codons de Goldman et Yang (1994) qui considère le codon comme une unité d'évolution. Ce module permet à la fois de calculer l'usage des codons dans la matrice mais également d'en estimer le niveau de substitutions synonymes et non synonymes. Ce module est celui qui est utilisé pour identifier les sites ou les lignées sous sélection positive (Yang, 1997). Les indications de la documentation du logiciel ont été suivies pour réaliser ces analyses.

#### A-1. Recherche des sites sous sélection positive

Ce modèle permet de faire varier le ratio de substitutions synonymes/non synonymes pour chaque codon (chaque site) (Nielsen et Yang, 1998; Yang *et al.*, 1998). Plusieurs modèles sont implémentés au module *codeml* et ont été testés sur des jeux de données réels.

Les paramètres utilisés ont été les suivants : *runmode* = 0, permettant de prendre en compte l'arbre de gène que nous entrons dans le modèle; *model* = 0, modèle de substitution de codons recommandé par Yang (1997) pour la recherche de sites sous sélection positive. Ensuite les différents modèles de variation sur les sites sont testés (*NSsites* = 1 à 8).

Les différents modèles proposés par le module ont été testés et le modèle M8 a détecté trois sites sous sélection positive - Q1, S43 et H287 - sur l'ensemble des séquences avec un score de maximum de vraisemblance  $\text{LogL} = -3434,91$ . Les valeurs  $dN/dS$  reportées le long de la séquence protéique montrent que certains sites présentant des valeurs supérieures à 1 ne sont pourtant pas détectés sous sélection positive par le modèle (Fig. 24). Ces résultats indiquent que sur l'ensemble des populations et espèces testées, ces trois acides aminés varient d'avantage que le reste de la séquence et devraient être plus variables également si nous utilisons un jeu de données plus importants, en restant dans la même famille des Heteroderidae. Deux de ces sites sont localisés dans la pro région de la cathepsine L et le troisième dans la région mature.



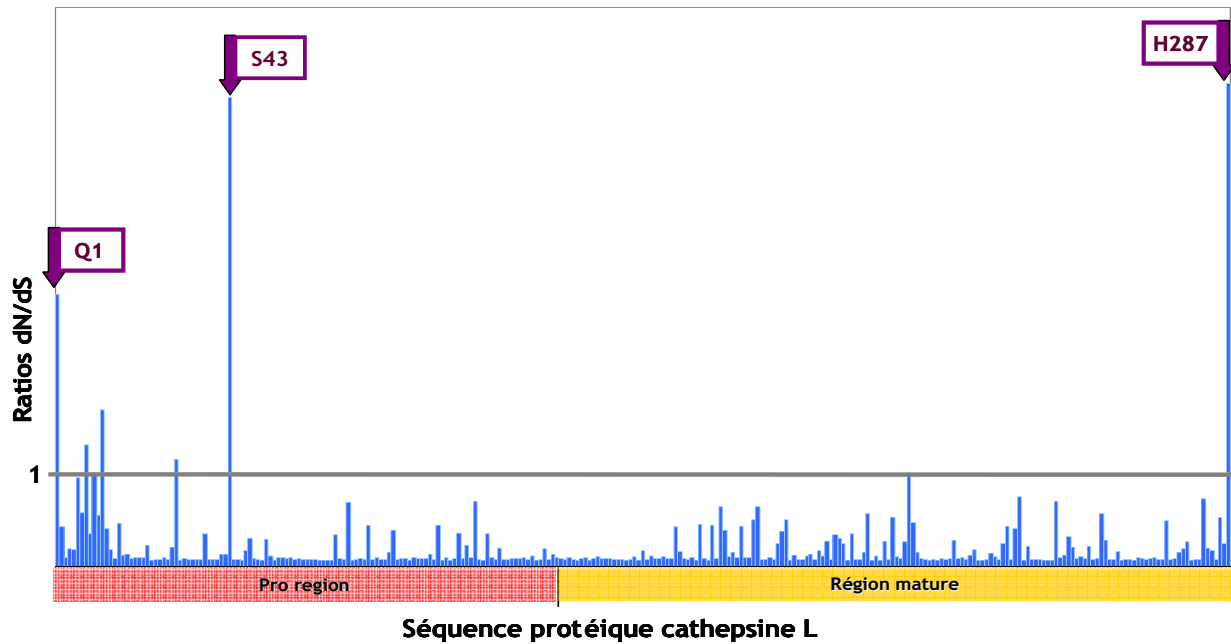


Figure 24 : Distribution des valeurs de dN/dS estimées par PAML le long de la séquence cathepsine L chez les 33 populations et espèces d'Heteroderidae. Les sites sous sélection positive sont indiqués par une flèche sur laquelle sont apposées la nature et la position de l'acide aminé concerné. La valeur 1 indiquée correspond la valeur du ratio dN/dS pour lequel la sélection est considérée comme neutre.

#### A-2. Recherche de sélection positive variable sur les différentes branches de l'arbre cathepsine

Nous avons précisé précédemment que les pressions de sélections ne sont pas homogènes sur un gène entier. Ces pressions de sélection ne sont probablement pas non plus équivalentes sur l'ensemble des populations et espèces d'une même famille. Nous avons ainsi recherché des variations du ratio dN/dS le long des branches de l'arbre cathepsine. Le module *codeml* permet de faire varier le ratio de substitution synonymes/non synonymes sur les différentes branches de l'arbre de gène entré dans la matrice (Yang, 1998; Yang et Nielsen, 1998). De même que pour l'analyse précédente, les recommandations de Yang sont suivies (Yang, 1997). Les paramètres utilisés ont été les suivants : runmode = 0, permettant de prendre en compte l'arbre de gène que nous entrons dans le modèle; model = 1, qui autorise un ratio dN/dS libre sur chaque branche; NSsites = 0.

La figure 25 présente les résultats obtenus en utilisant l'arbre généré en NJ, décrit dans l'article 4. Des accélérations de l'évolution des séquences sont observées sur différentes branches internes ou distales. Les populations indiquées en vert, correspondant à deux groupes de l'espèce *G. pallida*, présentent une évolution conjointe puisque l'augmentation des taux de mutations semble être apparue avant la séparation des différentes populations. La valeur du ratio dN/dS est relativement important (2,048) sur la branche qui mène aux populations GPS3, GPS9 et GPS10 pour lesquelles nous avons précédemment identifié une zone présentant un nombre important de mutations au niveau protéique (Fig. 12). En revanche, pour l'espèce *G. tabacum*, indiquée en rouge, une augmentation des taux de mutations semble suivre l'apparition des sous-espèces. La branche qui

mène aux populations de *G. rostochiensis* introduites en Europe, GRE1 et GRE2, présente également un ratio élevé (1,78) Ces deux populations sont les deux seules qui sont contrôlées par le gène de résistance H1. De même la branche qui mène à la population Santa Ana de *G. "mexicana"* présente une forte accélération de l'accumulation de substitutions non synonymes (ratio = 1,50).

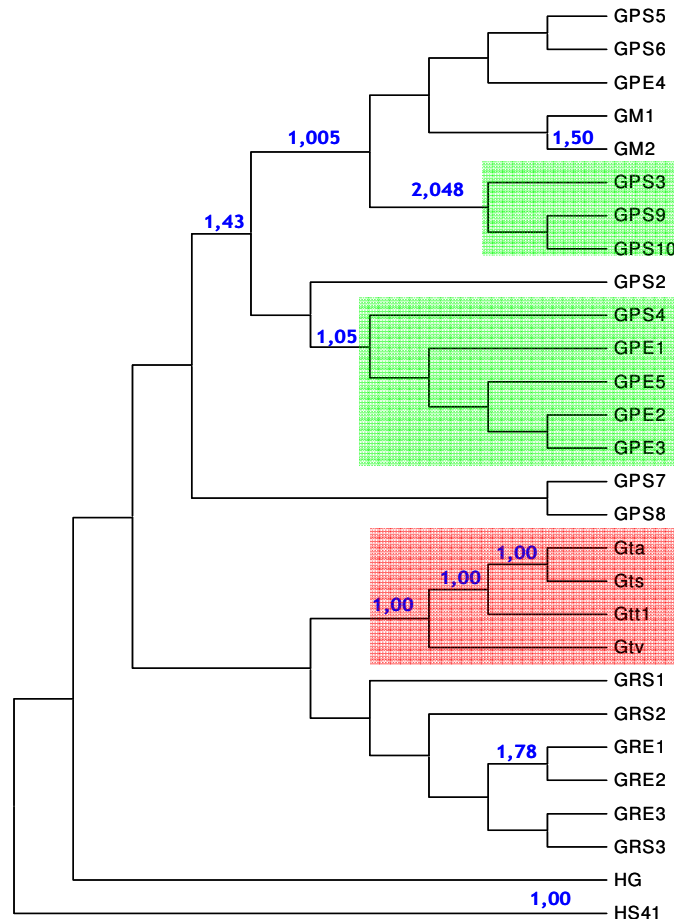


Figure 25 : Arbre phylogénétique du gène cathepsine L amplifié chez 33 populations de la famille des Heteroderidae. Les valeurs dN/dS estimées par le logiciel PAML dont la valeur est supérieure ou égale à 1 sont indiquées sur les branches.

### A-3. Recherche conjointe des zones préférentiellement soumises à sélection positives dans les lignées et sur les sites de la protéine

Le module *codeml* permet de détecter la sélection positive à la fois dans les différentes lignées et sur les différents sites (Yang et Nielsen, 2002). Les paramètres recommandés pour cette analyse sont les suivants : runmode = 0; model = 2 et NSsites = 2.

Cette approche permet de donner une vision globale de l'évolution du gène à travers les différentes populations et espèces et ainsi de déterminer le potentiel évolutif du gène. Au total, 28 sites sous sélection positive ont été mis en évidence dont 22 se situent dans la zone hypervariable

identifiée chez les populations GPS3, 9 et 10. Cette zone, indiquée en vert sur la figure 26, peut ainsi être globalement considérée sous sélection diversificatrice. Si nous reportons cette zone sur la structure protéique tridimensionnelle de la procathepsine L chez l'homme (Fig. 27), cette zone est localisée à l'extérieur de la protéine. C'est donc potentiellement un domaine qui peut interagir avec les protéines de plantes. Par ces analyses, nous pouvons ainsi relier la variabilité et les pressions de sélections subies par le gène cathepsine L avec une potentielle implication dans la relation hôte/parasite. Ces analyses permettent ainsi également d'apporter de nouvelles informations d'un point de vue fonctionnel sur des protéines dont l'activité est encore inconnue.

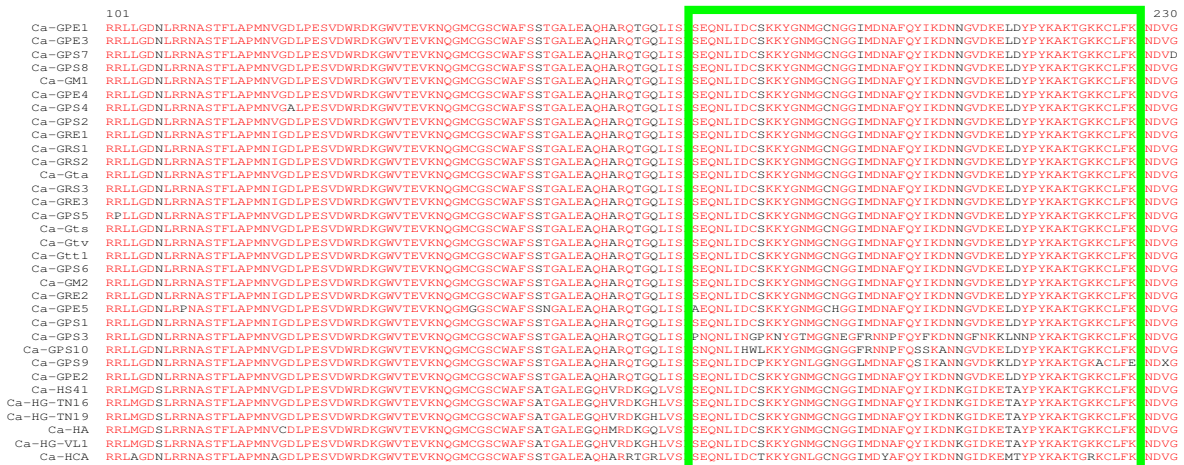


Figure 26 : Aligment protéique partiel de la cathepsine L montrant la zone comprenant les 22 sites sous sélection positive identifiée *via* le logiciel PAML dans la zone protéique qui avait préalablement été définie comme hypervariable.

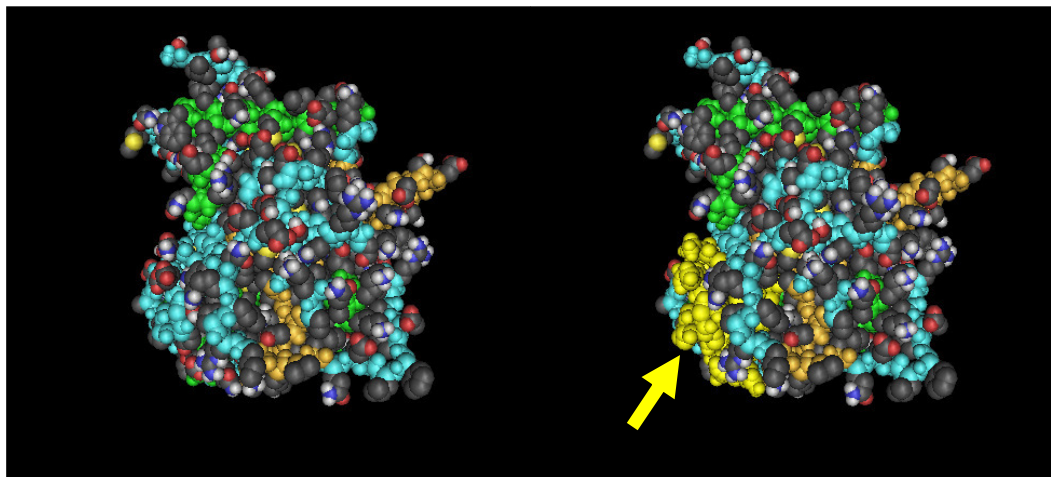


Figure 27 : Schéma de la structure tridimensionnelle de la protéine procathepsine L chez l'homme (N° accession 1CS8) visualisée avec le logiciel Cn3D 4.1. Sur la structure de droite, est représentée en jaune la zone de la protéine qui semble être plus variable et qui est sous sélection positive.

### B-. Détection de la sélection positive pour le gène pectate lyase

Les analyses séparées ont également été menées sur le gène pectate lyase afin d'identifier les sites, les branches de l'arbre et finalement les zones globales soumis à sélection diversificatrice. Les mêmes paramètres que pour le gène cathepsine L ont été utilisés pour les différents tests.

#### B-1. Recherche des sites sous sélection positive

De la même façon que pour le gène cathepsine L, le modèle M8 du module *codeml* a permis d'identifier trois sites - I3, K28 et A112 - sous sélection positive avec une score de maximum de vraisemblance  $\text{LogL} = -966,27$ . Lorsque nous reportons les valeurs de ratio dN/dS le long de la séquence en acides aminés (Fig. 28), nous constatons également que certains sites présentant des valeurs de  $\text{dN/dS} > 1$  ne sont pour autant définis comme sites sous sélection positive. Le site K28 est le premier acide aminé de la région R1 spécifique des pectate lyase de classe III (Fig. 16) et le site A112 est en position -1 par rapport à la région R4.

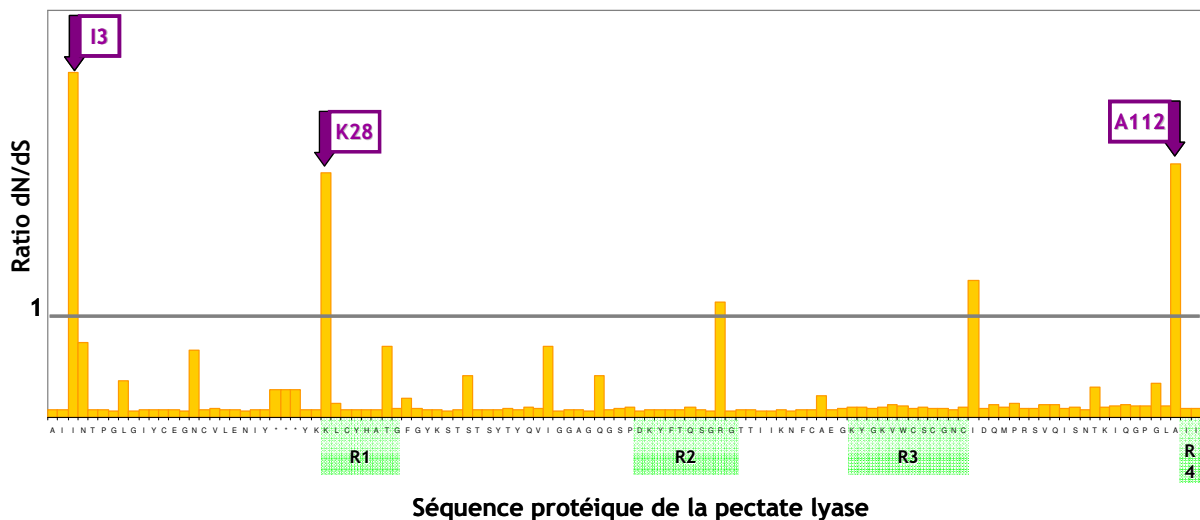


Figure 28 : Distribution des valeurs de dN/dS estimées par PAML le long de la séquence pectate lyase chez les 28 populations et espèces d'Heteroderidae. Les sites sous sélection positive sont indiqués par une flèche sur laquelle sont apposées la nature et la position de l'acide aminé concerné. La valeur 1 indiquée correspond la valeur du ratio dN/dS pour lequel la sélection est considérée comme neutre. Les régions spécifiques des pectate lyase de classe III sont indiquées en vert.

#### B-2. Recherche de sélection positive variable sur les différentes branches de l'arbre pectate lyase

Une importante hétérogénéité des ratios de taux de substitutions est observée sur l'arbre phylogénétique des pectate lyase, mais contrairement au gène cathepsine L, ces accélérations sont détectées seulement sur les branches distales de l'arbre (Fig. 29). En effet, les branches conduisant à GPS7 et GPS8 montrent une accélération importante de l'évolution du gène pectate lyase avec des ratios dN/dS respectifs de 2,023 et 1,038. Une forte accumulation de substitutions non synonymes

semble également être apparue sur la branche qui conduit à GPE4 (3,94) et également sur la branche GPE1 avec un taux beaucoup plus faible (1,04). De plus, les populations introduites en Europe de *G. rostochiensis*, GRE2 et GRE3, présentent conjointement une accélération de l'évolution du gène pectate lyase. Enfin, la branche qui mène à la sous espèces *G. tabacum solanacearum* semble également avoir subit une accumulation plus rapide de substitutions au cours de l'évolution de l'espèce *G. tabacum*.

Les données conjointes d'estimation des taux de substitution sur les différents sites dans les différentes populations n'ont pas pu être générées pour des raisons techniques qui restent à élucider.

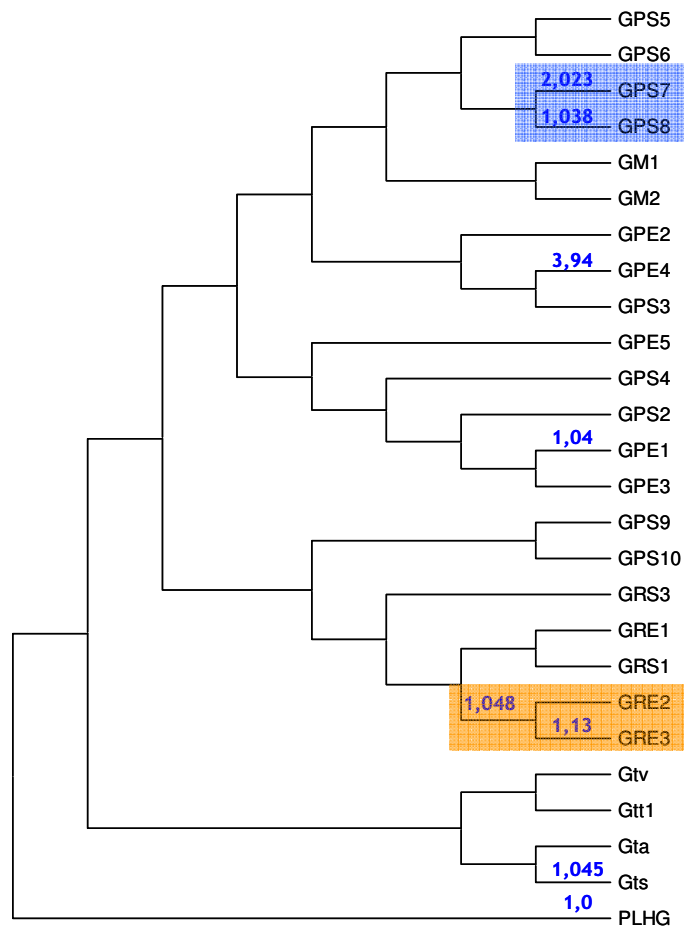


Figure 29 : Arbre phylogénétique du gène pectate lyase amplifié chez 28 populations de la famille des Heteroderidae. Les valeurs dN/dS estimées par le logiciel PAML dont la valeur est supérieure ou égale à 1 sont indiquées le long des branches de l'arbre.

## DISCUSSION

Cette étude préliminaire des taux de substitution synonyme *versus* non synonyme montre que les gènes du parasitisme sont effectivement soumis à des pressions de sélection diversificatrices. Nous avons montré par des analyses détaillées que ces pressions sont inégalement réparties le long des séquences des gènes et également dans les populations et espèces par utilisation du logiciel d'analyse PAML qui donne des données beaucoup plus précises que les données de DnaSP qui ne prend par ailleurs pas en compte la vraisemblance des ratio calculés.

### 1. Quel rôle des sites sous sélection positive ?

Que ce soit pour le gène pectate lyase ou pour le gène cathepsine L, les trois sites sous sélection positive au sein des populations et espèces d'Heteroderidae ne correspondent pas à des zones connues des protéines correspondantes, excepté pour un acide aminé de la pectate lyase qui est présent dans une région définie comme conservée, mais pour laquelle aucune fonction n'a été identifiée. Il serait très intéressant de mieux connaître les différentes zones d'activités de ces protéines afin d'identifier l'impact en terme d'activité protéique des mutations sur ces différents sites. Ces sites sous sélection positives sont rares (3 par protéine) et jouent probablement un rôle fondamental dans l'interaction plante/nématode. L'hypothèse la plus probable est une implication de ces sites dans la reconnaissance des protéines avec lesquels ils interagissent, en l'occurrence, les protéines végétales. Alors qu'au contraire, les régions sous sélection purificatrices seraient les zones d'activité des protéines de nématode.

### 2. Comment expliquer une accélération de l'évolution des gènes du parasitisme dans certaines populations et/ou espèces ?

Par l'étude de la variation des taux de substitutions dans les différentes populations et espèces, nous avons montré deux évolutions différentes des gènes du pouvoir pathogène.

#### A. Le gène *cathepsine L*

Il est intéressant de voir que toutes les populations européennes de *G. pallida* excepté GPE4 ont subi une évolution conjointement avec l'espèce GPS4 qui est une population provenant du sud du Pérou, plus rapide que les autres populations sud-américaines analysées. Ces données iraient dans le sens des résultats de Picard (2005) qui montrent une origine sud américaine des populations de *G. pallida* introduites en Europe. Nous pouvons alors émettre l'hypothèse d'un patron d'évolution commun entre les populations d'origine et les populations introduites puisque de fortes capacités adaptatives pourraient être nécessaires à l'implantation de ces populations en Europe, à l'image de ce qu'à décrit Anthoine (2006) pour les populations d'un autre nématode phytoparasite, *Nacobus aberrans*. En effet, les critères d'implantation d'une espèce dans un nouvel environnement reposent sur l'opportunité de pouvoir se développer, et dans notre cas de parasiter, mais aussi sur les

capacités à s'installer ce qui passe par une adaptation de la population nouvellement introduite au milieu. Si l'hypothèse d'une évolution plus rapide d'une population introduite est souvent reliée à un effet fondateur, dans notre cas, nous ne sommes pas en mesure de faire cette association sur la base du marqueur cathepsine puisque la population d'origine est également soumise à sélection diversifiante.

Nous avons également montré que les populations de *G. rostochiensis*, GRE1 et GRE2 qui sont les deux populations contrôlées par le gène de résistance H1 sont soumises à sélection positive. Cette accélération des tempos d'évolution pourrait s'expliquer par le fait que la pression qu'exerce la plante *via* le gène H1 sur ces populations est extrêmement forte et contraindrait ces nématodes à évoluer plus rapidement que la normale pour pouvoir à terme contourner le gène de résistance.

Une accélération de l'évolution du gène cathepsine L a également été détectée dans l'espèce *G. tabacum*. Cette évolution différentielle sur chaque branche précédant la différenciation des sous espèces pourrait être le reflet d'une adaptation des sous espèces à différents hôtes (Marché *et al.*, 2001), ce qui serait congruent avec le rôle putatif de la cathepsine L dans l'alimentation du nématode ou dans l'échappement aux défenses de l'hôte.

Par le biais de ces différentes observations, nous pouvons identifier quelques tendances évolutives, bien que celles-ci n'expliquent pas tout. Aucun patron réellement clair n'a été identifié par rapport à ce que nous connaissons des populations choisies (origine géographique, diversité génétique, pathotypes...). Nous observons également un ratio élevé à l'embranchement des populations GPS3, GPS9 et GPS10. Ces populations ne semblent pas partager de caractéristiques communes si ce n'est que les populations GPS9 et GPS10 sont respectivement des pathotypes PA5 et 6. De même, la population de *G. "mexicana"* originaire de Santa Ana évolue plus rapidement que la population Popocatepetl sans que nous puissions l'expliquer.

### **B. Le gène pectate lyase**

Contrairement à la cathepsine L qui présentent des évolutions différentielles à l'échelle des espèces entières, la pectate lyase ne présente des accélérations des taux d'évolution que sur les parties distales de l'arbre, c'est à dire dans les populations. Ainsi, comme nous venons de le mentionner pour la cathepsine L, deux des trois populations européennes de *G. rostochiensis* présentent clairement une accélération de l'évolution du gène pectate lyase. Ceci pourrait être lié aux besoins adaptatifs de ces nouvelles populations lorsqu'elles ont été introduites en Europe. Les populations GPS7 et GPS8 ainsi que les populations GPE4 et GPE1 présentent également des taux de substitution différents des autres populations sans que nous puissions y attribuer de signification biologique. Enfin, comme pour la cathepsine L, l'évolution plus rapide détectée dans la sous-espèce *G. tabacum solanacearum* pourrait s'expliquer par un besoin d'adaptation particulier à une nouvelle gamme d'hôtes. Ne pouvant définir de patron d'évolution dans le cas du gène pectate lyase, ces évolutions observées dans des populations pourraient être liées à des adaptations très locales et ponctuelles sans lien quelconque avec l'histoire évolutive des nématodes à kyste. Nous mettons en évidence, cette fois, des types d'évolution complètement différents entre les deux gènes du pouvoir pathogène.

### 3. En quoi les informations sur les zones à fort potentiel évolutif sont-elles utiles ?

Les analyses réalisées montrent clairement une zone de mutation préférentielle de la cathepsine L chez les Heteroderidae. Cette zone définie globalement sous sélection positive est donc une zone à fort potentiel évolutif : c'est à dire que les mutations ont tendance et auront certainement tendance à s'accumuler davantage dans cette zone que dans les autres parties du gène. Nous pouvons donc tirer deux informations importantes de ces données :

1- cette portion du gène est donc une portion contre laquelle il vaut mieux éviter de générer de nouvelles résistances puisque cette forte tendance à muter pourrait rendre rapidement caduque les stratégies lutte. De plus, cette information pourrait être très utile pour les expérimentations sur la fonction du gène cathepsine L basée sur des techniques très spécifiques d'une portion de gène telle que le RNA interférence. En effet, une tentative d'extinction du gène *via* cette portion pourrait se révéler inefficace du fait de la diversité génétique de cette région eu sein des Heteroderidae. Il faudrait mettre en relation ces données avec celles des sites préférentiels d'action des inhibiteurs de protéases. En effet, nous pourrions imaginer que cette zone qui accumule beaucoup de mutations non synonymes soit la cible d'inhibiteurs de protéases au quel cas cette zone mute pour échapper à la reconnaissance de ces enzymes. Une autre hypothèse est l'accumulation de mutations pour permettre de reconnaître un plus large éventail de protéines de plante, dans le cas où cette cathepsine L joue un rôle dans l'alimentation du parasite, par exemple.

2- dans cet objectif d'avoir de meilleures connaissances des interactions fines entre le nématode et sa plante hôte, de nombreuses hypothèses peuvent être émises quand au rôle de cette portion du gène. En effet, cette zone très fortement soumise à des pressions de sélection diversificatrices, est potentiellement une zone d'interaction avec des protéines de plantes qui tentent de la reconnaître. En effet, les gènes codant des protéines directement impliquées dans les mécanismes de reconnaissance hôte pathogène semblent avoir des patrons dévolution moléculaires inhabituels (Thomas *et al.*, 2006). Si ces observations sont souvent faites *a posteriori*, lorsque les domaines des protéines sont bien connus, nous pouvons imaginer rechercher les gènes ou les portions de gènes soumis à sélection positive pour identifier de nouvelles interactions sans que le rôle fonctionnel des protéines ne soit élucidé (Thomas *et al.*, 2006 ; Ortiz *et al.*, 2006).



*Conclusion &  
Perspectives*

*Partie IV*

## PARTIE IV :

# CONCLUSION & PERSPECTIVES

Malgré les efforts déployés de par le monde pour tenter d'enrayer les infections des plantes cultivées par les nématodes à kyste de manière durable et respectueuse de l'environnement, aucune solution satisfaisante n'a pour le moment été proposée. Pourtant, l'émergence des biotechnologies couplées aux découvertes moléculaires du fonctionnement des interactions plante/nématode a permis d'imaginer de nouvelles stratégies de lutte qui sont actuellement en voie de développement. C'est le cas notamment des inhibiteurs de protéases qui constituent les seuls tests réalisés au champs mais qui n'ont pas montré de réduction de développement des nématodes supérieure à celle observée avec des variétés de pomme de terre partiellement résistantes (Urwin *et al.*, 2001). Il est donc nécessaire de poursuivre les investigations sur les gènes du pouvoir pathogène afin de mieux comprendre les mécanismes d'interaction entre les nématodes et leurs plantes hôtes, et également de proposer de nouvelles stratégies de lutte.

Au début de cette étude, nous nous posons les questions suivantes :

1- Quels sont les gènes du pouvoir pathogène des nématodes à kyste qui jouent un rôle clé dans le parasitisme ?

Nous avons apporté deux types de données : à l'échelle populationnelle, nous avons isolé et caractérisé des nouveaux gènes du pouvoir pathogène chez une population de *G. pallida*. Parmi eux, le gène *rbp-1* semble être un candidat de choix en tant que gène clé du parasitisme. Nous avons également apporté des données complémentaires sur les gènes pectate lyase et cathepsine L à un échelle beaucoup plus large, celle des nématodes à kyste.

2- Comment ces gènes ont-ils évolués par rapport à divers critères tels que les différences de pathogénicité, de spécificité ou de gamme d'hôtes, d'origine géographique ?

Nous avons clairement mis en évidence un mode évolutif particulier des gènes pectate lyase et cathepsine L en comparaison avec le gène de ménage (facteur d'élongation) qui passe par un taux de substitutions non synonymes élevé. Concernant le gène pectate lyase, l'arbre de gène montre un groupement des espèces du genre *Globodera* selon leur plante hôte qui n'est pas celui observé avec l'utilisation d'un marqueur neutre tel que le gène de ménage. Concernant le gène cathepsine, nous avons montré une relation entre les accélérations des tempos d'évolution et la spécialisation des sous espèces de *G. tabacum* sur différentes plantes hôtes, par exemple. Aucun lien n'a été mis en évidence entre l'origine géographique (indigène *versus* importée) des populations et des tempos d'évolution particuliers.

3- Comment peuvent-ils évoluer s'ils étaient soumis à des pressions de sélection liées à une nouvelle source de résistance (exemple des résistances artificielles) ?

Nous avons montré que les gènes pectate lyase et cathepsine L possèdent des régions sous sélection positive. Ces données suggèrent un potentiel évolutif relativement important de ces régions. De ce fait, nous avons conforté l'hypothèse de départ d'un risque de contournement des résistances mises en place par une évolution importante des gènes du pouvoir pathogène. Il semble donc impératif de cibler les régions géniques pour la création de résistances artificielles durables.

Enfin et surtout, mon travail de thèse, a permis de proposer un nouvelle approche d'étude des gènes du pouvoir pathogène. Les études fonctionnelles de ces gènes se font en général à l'échelle d'une population ou d'un espèce de nématode dans le cadre d'une interaction spécifique. C'est l'objectif de la génomique fonctionnelle : "après l'aire de la génomique, le challenge en biologie est de déterminer les fonctions de tous les gènes identifiés et de comprendre comment ils contribuent à la viabilité des organismes" (Greenbaum *et al.*, 2001). Nous proposons dans cette étude de compléter ce type de données avec des approches visant à prendre en compte la diversité des nématodes d'une même famille ayant des caractéristiques différentes telles que les plantes hôtes, les pathotypes, le caractère virulent ou avirulent de certaines populations..., par analyse de l'évolution moléculaire de gènes du pouvoir pathogène. Cette approche se situe à l'interface entre la génomique fonctionnelle et la phylogénie, dans le cadre de la biologie évolutive (Fig. 1). Le but de cette approche est de mieux comprendre l'évolution des interactions plante/nématode, et de définir le potentiel évolutif des gènes du pouvoir pathogène.

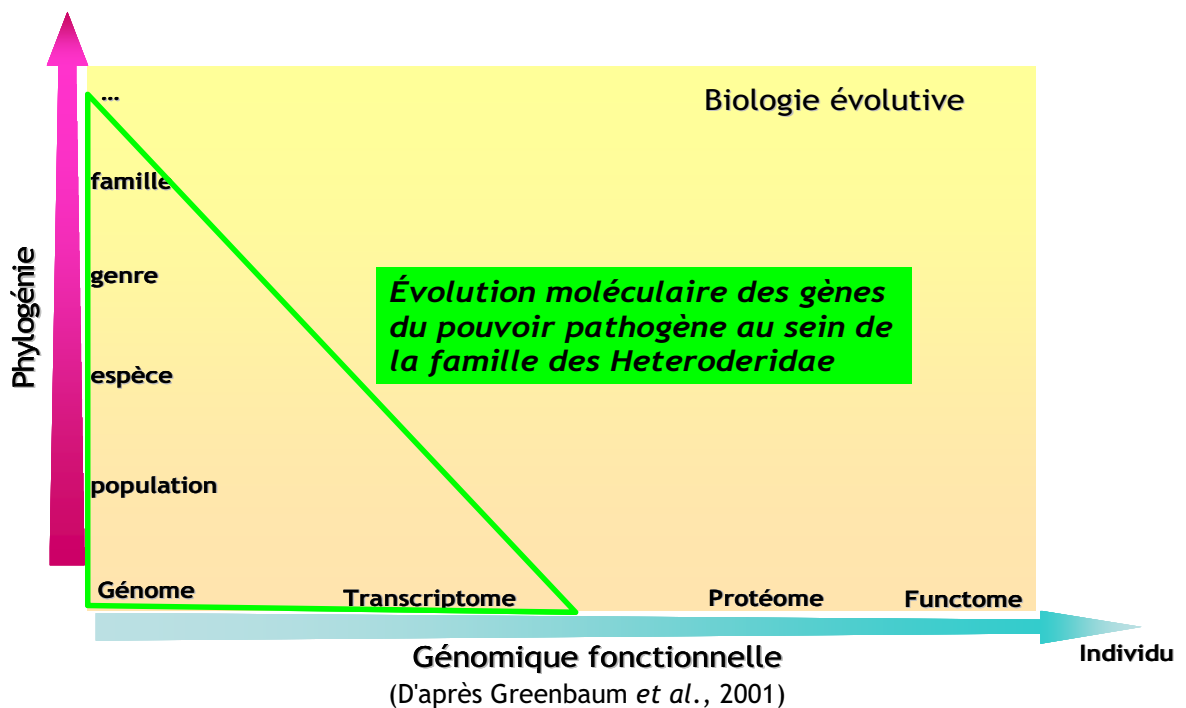


Figure 1 : Positionnement de notre approche d'étude du pouvoir pathogène des nématodes à kyste (cadre vert) à l'interface entre la génomique fonctionnelle, la phylogénie et la biologie évolutive.

## 1. IDENTIFICATION ET CARACTERISATION DES GENES DU POUVOIR PATHOGENE DES NEMATODES A KYSTE ET MISE EN EVIDENCE DE LEUR ROLE DANS L'INTERACTION PLANTE/NEMATODE

Cette étude nous a permis de caractériser 5 gènes du pouvoir pathogène chez *G. pallida*, dont 3 sont des nouveaux gènes. En effet, l'approche différentielle entre les nématodes à kyste *G. pallida* et *G. "mexicana"* a permis de caractériser 3 transcrits dont l'implication dans le parasitisme, si elle n'a pas été directement démontrée, a très fortement été supposée par l'accumulation d'indices tels que la localisation cellulaire des transcrits, l'analyse des séquences montrant la présence d'un signal de sécrétion ainsi qu'une variabilité intra spécifique relativement importante.

### 1.1. ETUDE DE LIGNEES HYBRIDES *G. PALLIDA* X *G. "MEXICANA"* : POLYMORPHISME DES GENES *IA7* ET *IVG9*

L'étude des mutations des gènes *la7* et *Ivg9* entre les parents *G. pallida* et *G. "mexicana"* et les lignées hybrides ayant des capacités variables à se développer sur pomme de terre et sur morelle noire, n'a pour le moment pas donné de résultats permettant de relier ces mutations à une spécificité d'hôte. La mise en évidence de mutations préférentiellement non synonymes constitue néanmoins un argument en faveur de l'implication des gènes *la7* et *Ivg9* dans le parasitisme. En effet, les gènes impliqués dans les interactions semblent être souvent soumis à des pressions diversifiantes comme nous l'avons observé pour d'autres gènes dont l'implication au niveau du pouvoir pathogène était plus certaine. Toutefois des investigations venant compléter les résultats présentés dans l'article 1 pourraient être réalisées afin de préciser l'implication de ces gènes dans la spécificité d'hôte par notamment :

1- l'amplification et le séquençage de ces mêmes gènes dans des lignées hybrides rétro croisées avec *G. "mexicana"* et sélectionnées sur morelle noire plutôt que sur pomme de terre. Nous pourrions alors mettre en évidence des différences entre ces gènes lorsque les lignées ne se développent que sur morelle ou que sur pomme de terre. Cependant ces lignées hybrides ne sont pour le moment pas disponibles.

2- l'étude des différences de niveau d'expression des ces gènes entre les espèces *G. pallida* et *G. "mexicana"* ainsi que sur les hybrides. La méthode la plus adaptée serait la PCR quantitative pour évaluer les différences d'expression, certainement très fines, qui existent entre les deux espèces de nématodes. Il serait également intéressant d'évaluer les variations du niveau d'expression à différents stades parasitaires afin de mettre en évidence l'étape à laquelle sont impliqués les gènes.

### 1.2. *IC5* DEVENU *RBP-1* : LA GRANDE FAMILLE DES GENES *RANBPM-LIKE*

Les études menées sur le transcrit *ic5* renommé par la suite *rbp-1*, ont permis de montrer l'implication probable de ce gène dans l'interaction plante/nématode, et plus particulièrement dans

l'étape de mise en place et de maintien du site nourricier. De nombreuses hypothèses sont émises quant à son rôle dans les interactions.

### 1.2.1. Rôle des RanBPM ?

Comme décrit dans la première partie du manuscrit, des protéines de type RanBPM sont présentes chez de nombreux organismes et ont des fonctions semble-t-il très diverses. Nous avons émis des hypothèses sur le rôle du gène *rbp-1* dans l'interaction parasitaire du fait des différents indices que nous avons mis en évidence : la localisation des transcrits de ce gène dans la glande dorsale et la présence d'un signal de sécrétion dans les séquences des transcrits *rbp-1*. Le seul moyen d'affirmer la sécrétion de protéines RBP-1 serait de les localiser dans la plante. Des anticorps spécifiquement dirigés contre cette protéine pourraient être utilisés pour rechercher la protéine dans des coupes cellulaires de plante, comme ce qui a été réalisé pour la calréticuline chez *M. incognita* (Jaubert *et al.*, 2005). Cependant, cette technique nécessite de faire des coupes de syncytium. Une première étape pourrait être la localisation de cette protéine dans les sécrétions salivaires du nématode (Jaubert *et al.*, 2002).

De nombreuses techniques pourraient maintenant être mises en place pour montrer clairement le rôle du gène *rbp-1* dans l'interaction. La technique certainement la plus efficace est le silencing par RNA interférence. Si cette technique est actuellement mise au point sur les nématodes à kyste et à galle (Chen *et al.*, 2005; Rosso *et al.*, 2005), certaines données pourraient être prises en compte pour affiner l'interprétation des résultats. En effet, la connaissance des familles multigéniques de gènes d'intérêt permettrait, selon la question posée, de ne cibler qu'une copie particulière du gène ou au contraire d'éteindre une famille entière. Ceci éviterait certainement des erreurs d'interprétation et et/ou des problèmes d'expérimentation (tels qu'une efficacité partielle de la technique comme décrit par Chen *et al.* (2005)) pouvant être expliqués par un effet compensatoire d'une autre copie dans une famille de gène. Maintenant que nous avons une vision plus large de la famille des gènes RanBPM-like, bien que probablement non exhaustive, nous aurions la possibilité de faire un choix de la ou les copies à inhiber. Au sein des 12 copies de RanBPM identifiées, nous ne sommes pas en mesure, pour le moment, de définir si elles sont toutes impliquées dans le pouvoir pathogène ou non. Une première étape serait de localiser les transcrits correspondants par hybridation *in situ*. L'intérêt porterait alors préférentiellement sur les copies localisées dans les glandes salivaires. Des études préliminaires sont réalisées en ce moment sur l'expression stade spécifique de quelques copies de gènes RanBPM-like. Les copies exprimées à tous les stades de développement du nématode seront potentiellement des copies liées à des fonctions plus ubiquitaires chez les nématodes en comparaison avec des copies dont l'expression apparaîtra différentielle entre les différents stades. Malgré tout, il ne semble pas improbable que toutes ces copies soient des gènes liés au parasitisme. De nombreux pathogènes possèdent plusieurs copies du même gène qui sont exprimées en fonction des besoins quantitatifs ou qualitatifs. C'est une stratégie évolutive qui permet d'anticiper des évolutions du milieu ou de l'hôte et d'être réactif très rapidement. Des variants d'un gène d'avirulence avec un haut niveau de variation (20% des acides aminés) ont été identifiés chez un champignon (Dodds *et al.*, 2006). Cette grande diversité de protéines d'avirulence serait la clé de la capacité du champignon à échapper aux défenses de l'hôte

qui constituent une forte pression de sélection sur ces gènes conduisant à leur évolution. Ainsi, une fois que les copies qui semblent être impliquées dans le parasitisme auront été identifiées, tel que *rbp-1*, l'étude de la variabilité de ces gènes au sein de diverses populations et espèces de nématodes à kyste présentant des pathogénies différentes (gamme d'hôtes, pathotypes) pourront être menées.

### 1.2.2. *rbp-1* : gène d'avirulence ?

Des études réalisées récemment par une équipe américaine sur le gène de résistance *Gpa2* ont récemment montré une interaction probable entre la protéine de ce gène et RBP-1 (Peter Moffett, com. pers.). En effet, lorsque le gène *rbp-1* est exprimé dans des feuilles de tabac contenant ce gène de résistance, une réaction d'hypersensibilité (HR) est observée. Ceci suggère une interaction spécifique entre GPA2 et RBP-1, le gène *rbp-1* pourrait donc être un gène d'avirulence. Cependant, la population de *G. pallida* chez laquelle *rbp-1* a été caractérisé n'est pas contrôlée par *Gpa2*. Il semble donc assez curieux d'observer une HR lorsque ces deux protéines se rencontrent. L'interaction entre GPA2 et RBP-1 demande donc à être démontrée. Ceci pourrait être réalisé par co-immunoprécipitation ou en utilisant un système double hybride. Il serait très intéressant d'amplifier le gène *rbp-1* chez les populations de *G. pallida* contrôlées par les gènes *Gpa2* afin de voir s'il existe des différences (structure/mutations) pour ce gène susceptibles d'expliquer les différences de pathogénicité observées entre les populations contrôlées ou non par *Gpa2*. L'analyse des séquences génomiques, de l'expression du gène *rbp-1* dans les différentes populations et des interactions protéiques devrait permettre de comprendre ces différences. Si *rbp-1* était un gène d'avirulence, ce serait le premier décrit chez un nématode phytoparasite. Dans le cas où aucune interaction directe entre GPA2 et RBP-1 ne serait mise en évidence, ceci suggérerait que *rbp-1*, bien que n'étant pas un facteur d'avirulence, participerait à la cascade d'activation qui conduit à la réaction HR. Dans tous les cas, le gène *rbp-1* en tant que gène impliqué dans les interactions incompatibles est un candidat de choix pour poursuivre les études fonctionnelles et évolutives.

### 1.3. UN TRANSFERT HORIZONTAL DES GENES DU PARASITISME ?

Les mécanismes responsables de l'évolution des génomes sont divers : duplication de gènes, diversification, altération du contrôle du métabolisme des gènes, adaptation de gènes existants pour coder des protéines avec de nouvelles fonctions et transferts horizontaux. L'origine du parasitisme des plantes pourrait être liée à des transferts horizontaux de gènes chez les nématodes. Ceux-ci seraient alors devenus plus adaptés à ce mode de vie parasitaire. L'évolution du génome de certains nématodes phytoparasites aurait bénéficié par exemple de transfert horizontal de gènes tels que les cellulases ou les pectinases (Jones *et al.*, 2005; Ledger *et al.*, 2006; Mitreva *et al.*, 2005). Néanmoins il est important que ces hypothèses soient basées sur des analyses des séquences et des reconstructions phylogénétiques et non sur de simples homologues identifiées dans les bases de données, ce qui avaient biaisées nombre d'interprétations de l'évolution de certains organismes (Kurland *et al.*, 2003). Nos données pourraient maintenant être analysées en s'intéressant notamment aux phases auxquelles sont insérés les introns (Ledger *et al.*, 2006) pour vérifier si nous

avons également des indications de transfert horizontal pour les gènes pectate lyase et cathepsine L.

## 2. POLYMORPHISME NUCLEOTIDIQUE ET EVOLUTION MOLECULAIRE DES GENES DU POUVOIR PATHOGENE

Pourquoi étudier le polymorphisme des gènes du pouvoir pathogène et leur évolution moléculaire ? Les études fonctionnelles ont jusqu'à présent été menées à partir des données générées généralement sur une espèce. Il manque actuellement des données sur la variabilité en terme de nombre de copies dans un génome et en terme de polymorphisme des copies orthologues au sein des populations et espèces de nématodes. En effet, ces données permettraient de pouvoir s'assurer d'étudier une seule copie d'une famille de gène et pour pouvoir transposer les données acquises sur une population ou espèce à d'autres espèces proches.

### 2.1. COMMENT S'ASSURER DE LA COMPARAISON DE GENES ORTHOLOGUES ?

L'étude du polymorphisme des gènes n'est pertinente que si les gènes comparés sont des copies orthologues. En effet les copies paralogues ne sont certainement pas sous les mêmes contraintes sélectives. Leur comparaison n'a donc pas de sens pour mettre en évidence des pressions de sélection différentielles.

Aucun critère fiable n'a actuellement été identifié pour être certain de comparer des orthologues. Les données les plus utilisées sont le pourcentage d'homologie entre les séquences et les structures introns/exons des gènes. Si de nombreuses études soutiennent la conservation de la structure intron/exon des gènes orthologues dans les clades (COGs) (Rokas *et al.*, 1999, Wada *et al.*, 2002), d'autres études ont cependant montré que ce critère n'est parfois pas suffisant (Goetze *et al.*, 2006). Néanmoins, les résultats obtenus pour les gènes pectate lyase et cathepsine L semblent intéressants. En effet, que ce soit pour la pectate lyase ou pour la cathepsine L, nous avons mis en évidence des structures communes de ces gènes au sein de la famille des Heteroderidae. Si les structures exoniques sont très conservées en terme de taille, les introns quant à eux sont plus variables. Cependant, nous avons montré que les introns pour ces deux gènes sont positionnés de la même façon à l'échelle des Heteroderinae et présentent des motifs très conservés à l'échelle intra spécifique et parfois interspécifique. Ceci nous laisse penser que nous comparons bien des orthologues. Il semblerait que les introns ne se soient pas insérés de façon aléatoire dans les génomes mais à des positions précises désignées sous le terme de "proto-splice sites" (Dibb, 1991). Si les introns se sont insérés chez leur ancêtre commun, des gènes orthologues auront donc une structure intron/exon très conservée, même s'ils ont gagné ou perdu un ou plusieurs introns par la suite. Ainsi, si la position de certains introns semble conservée à l'échelle de certains embranchements (Rogozin *et al.*, 2003), nous pouvons raisonnablement penser qu'à l'échelle d'une famille ce soit le cas de la majorité des introns. Les arrangements identiques que nous avons observés pour les gènes pectate lyase et cathepsine L suggèrent fortement que les copies étudiées dérivent d'un ancêtre commun proche. Ceci est un argument fort pour affirmer que nous avons

comparé des orthologues (Sanchez *et al.*, 2003). L'utilité de ces données a été confirmée lorsque nous avons recherché les orthologues de *rbp-1* chez différentes populations européennes et sud-américaines de *G. pallida*. En effet, les difficultés d'amplification de ce gène nous ont conduit à amplifier le transcrite correspondant. Nous avons comparé les séquences obtenues et caractérisé le polymorphisme au sein du genre *G. pallida*, en comparaison avec une séquence *G. "mexicana"* (données non montrées). Il s'est avéré que deux des séquences obtenues chez deux populations sud américaines sont nettement plus variables que les autres séquences. Les distances génétiques sont plus importantes entre ces deux populations et les *G. pallida* européennes que entre les espèces *G. pallida* et *G. "mexicana"*. Nous avons donc supposé que ces deux copies très divergentes sont des copies paralogues. Cependant, nous n'avons pour le moment pas été en mesure de confirmer cette hypothèse puisque nous ne disposons pas de l'information sur les introns pour le vérifier. Ces études préliminaires montrent les difficultés rencontrées pour différencier les paralogues des orthologues et que l'information des introns semble être un critère sur lequel il faut impérativement s'appuyer pour s'assurer de l'orthologie des séquences étudiées.

## **2.2. LES GENES DU POUVOIR PATHOGENE ETUDIES ONT-ILS EVOLUE DIFFEREMMENT DU GENE DE MENAGE ?**

L'analyse du polymorphisme des gènes du pouvoir pathogène a révélé un mode évolutif commun pour les deux gènes pectate lyase et cathepsine L, mais différent de celui du gène de ménage. En effet, ces deux gènes ont un tempo d'évolution qui semble être plus rapide puisqu'une accumulation plus importante de mutations est observée en comparaison avec le facteur d'élongation. Ces deux gènes sont donc moins contraints par la sélection purificatrice que le gène de ménage. Cependant, les comparaisons pourraient être affinées si nous avions accès à un référent plus proche des gènes du pouvoir pathogène. Ceci pourra être réalisé pour le gène *rbp-1* si une des copies paralogues identifiées correspond à un gène du fonctionnement cellulaire chez le nématode et non à un gène du parasitisme. Un autre moyen d'affiner notre connaissance de l'évolution des gènes du pouvoir pathogène serait de comparer des copies paralogues au sein d'une même population qu'elles aient des activités communes ou non. Dans le cas où elles ont des fonctions différentes, nous posséderons un référent plus proche, tel que nous venons de l'évoquer pour *rbp-1*, sinon, dans le cas de fonctions similaires, nous pourrions mettre en évidence d'éventuelles pressions de sélection différentielles entre différentes copies.

## **2.3. TOUS LES GENES DU POUVOIR PATHOGENE EVOLUENT-ILS SUR LE MEME MODE ?**

Des études préliminaires ont été réalisées sur le polymorphisme d'une portion du gène *rbp-1* au sein des nématodes de l'espèce *G. pallida* (article 3). A l'image de ce qui a été observé pour les autres gènes du pouvoir pathogène, tels que *la7*, *IVg9*, la pectate lyase et la cathepsine L, les substitutions non synonymes représentent une partie importante des mutations réparties de façon homogène le long de la séquence. Le gène *rbp-1* semble avoir évolué sur le même mode que les autres gènes du parasitisme. Des études complémentaires sont nécessaires afin : 1- d'obtenir les séquences génomique qui nous permettront de nous assurer de comparer des orthologues, 2-



d'obtenir des séquences sur un plus large éventail de populations et d'espèces. Enfin, nous pourrions ainsi engager des analyses plus poussées sur l'évolution de ces gènes chez les Heteroderidae.

Nos résultats ont montré que les cinq gènes du pouvoir pathogène étudiés semblent évoluer sur le même mode, et ce, bien qu'ils soient impliqués à différentes phases parasitaires et soient exprimés dans différents organes. Il se pourrait donc bien que ces données soient généralisables à l'ensemble des gènes du pouvoir pathogène. Il semble assez peu probable que des pressions de sélection identiques aient agi sur un ensemble de gènes dispatchés dans le génome de façon simultanée au cours de l'évolution des nématodes. Ce mode évolutif commun pourrait alors être expliqué par un regroupement physique des gènes du parasitisme dans le génome, à l'instar des îlots de pathogénicité décrits chez les bactéries qui semblent constituer des unités de sélection (Araki *et al.*, 2006).

### 3. PERSPECTIVES

#### 3.1. APPORTS DES RESULTATS POUR L'ETUDE DES GENES DU POUVOIR PATHOGENE

Toute l'originalité de notre approche d'étude du polymorphisme et de l'évolution moléculaire repose sur le fait qu'au-delà de l'aspect mécanistique pur, les gènes du pouvoir pathogène ont été étudiés en prenant en compte le cadre coévolutif dans lequel les nématodes phytoparasites évoluent. Notre étude constitue une nouvelle approche pour élucider le fonctionnement des gènes du pouvoir pathogène qui font souvent partie du "unknown" comme décrit par Greenbaum *et al.* (2001). En effet, cette méthode nous permet de définir si le gène identifié est impliqué dans les interactions sans avoir besoin d'y associer un rôle fonctionnel bien établi.

Les données que nous avons obtenues spécifiquement pour quelques gènes du pouvoir pathogène doivent maintenant être transposées à d'autres gènes et à d'autres échelles taxonomiques. Nous avons montré que les gènes du pouvoir pathogène semblent avoir un mode évolutif commun quelle que soit la fonction du gène dans l'interaction. Les gènes du pouvoir pathogène sont donc des gènes qui en règle générale présentent des niveaux de variabilité importants répercutés au niveau protéique. Les risques que nous avons mentionnés au début de cette étude quant aux problèmes de contournement potentiel des résistances artificielles créées sans s'être préoccupé de la variabilité de la cible sont donc réels. C'est le cas actuellement des techniques de silencing par RNA interférence, par exemple, qui sont utilisées sans connaître la variabilité des gènes cibles entre les populations ou espèces étudiées. En effet, supposons un gène dont la zone fonctionnelle est très conservée au sein des Heteroderidae, celle-ci serait très intéressante pour empêcher l'expression du gène chez tous les nématodes à kyste. Au contraire, une zone très variable peut permettre de s'assurer de la spécificité vis-à-vis d'une population ou d'une espèce. C'est le cas des gènes que nous avons étudiés qui montrent que seules certaines zones des gènes pectate lyase et cathepsine L sont soumises à sélection positive.

Notre approche peut également permettre de mettre en évidence des domaines d'interaction ou de reconnaissance en fonction du niveau de polymorphisme et du type de sélection auxquels les gènes du pouvoir pathogène sont soumis. Le peu de données auxquelles nous avons accès sur les domaines d'activités des protéines pectate lyase et cathepsine L ne nous permet pas de déterminer l'impact des mutations dans ces régions. Cependant, la sélection positive détectée sur ces gènes est une indication d'un rôle potentiel dans l'interaction plante/nématode et peut ainsi être utilisée pour identifier comme cible les portions/domaines d'un gène ayant plus particulièrement un rôle dans l'interaction. L'approche que nous avons développée est donc complémentaire des analyses fonctionnelles classiques, puisqu'elle permet, par l'analyse des séquences d'un gène à une échelle taxonomique large, de mettre en évidence des zones préférentielles d'activité de la protéine. De plus, le fait de se situer à une échelle large de populations, d'espèces, de genres voire de familles d'organismes, permet une transposition immédiate des conclusions quant aux implications fonctionnelles d'une protéine à l'échelle entière des organismes étudiés et non pas restreinte à une population ou une espèce, comme ce qui est couramment fait.

### **3.2. APPORTS DES RESULTATS POUR LA GESTION DES GENES DE RESISTANCE**

Les données acquises au cours de ce projet ont permis d'apporter de nouvelles informations sur les gènes du parasitisme : leur nature, leur polymorphisme et leur potentiel évolutif. Quelles en seront les applications possibles ?

Pour la lutte génétique artificielle : les nouveaux gènes du parasitisme identifiés peuvent constituer des cibles potentielles pour créer des variétés de plantes résistantes. La connaissance de leur polymorphisme permet d'identifier les régions du gène ou de la protéine correspondante à cibler. Les données sur le potentiel évolutif permettent de définir si le gène choisi pour cible risque d'évoluer au point de rendre la résistance créée rapidement caduque. Seule cette connaissance du côté évolutif des gènes ciblés permettra de s'assurer de la durabilité des résistances mises en place. En effet, nous avons montré par notre étude que les gènes du pouvoir pathogène présentent un mode évolutif particulier qui tend à générer de la diversité dans certaines régions géniques. Maintenant que nous avons montré que les gènes du pouvoir pathogène des nématodes sont réellement soumis à des pressions de sélection particulières, il faudrait s'attacher à l'avenir à définir le potentiel évolutif des gènes pressentis comme cibles de résistances artificielles. En effet, l'identification de régions géniques soumises à sélection positive, c'est à dire possédant un fort potentiel d'évolution, pourront être évitées pour le développement de biotechnologies qui visent à cibler une portion de gène ou de la protéine correspondante, et ce, afin de s'assurer de la durabilité de ces nouvelles résistances.

Pour la lutte génétique naturelle : Nous avons mis en évidence une sélection positive des gènes du pouvoir pathogène. Si ces gènes sont soumis à la sélection positive, c'est parce qu'ils y sont contraints par la plante et inversement. Dans ce cas, nous pouvons supposer que la plante est capable de reconnaître les protéines de nématodes, ce qui force ces derniers à accumuler des

mutations pour échapper au système de reconnaissance de la plante. En contre partie, la plante va évoluer de manière à reconnaître un maximum de protéines de nématodes également par accumulation de mutations. Ainsi il y a un intérêt manifeste à développer une étude miroir qui permettrait un enrichissement réciproque puisque :

- les gènes de nématodes pourraient être utilisés pour aller chercher chez la plante la protéine avec laquelle le produit de ce gène interagit. Cette manipulation pourrait se faire par co-immunoprécipitation des protéines. Lorsque la protéine de plante sera identifiée dans le cadre d'une réaction compatible, il sera alors possible de chercher les mêmes protéines de plante dans le cadre d'une réaction incompatible et voir quelles sont les mutations qui font que la plante passe du statut de résistant à sensible. Ainsi, de nouvelles sources de résistance pourraient être identifiées par criblage de génotypes de plantes portant ces gènes.
- les gènes de plantes également soumis à la sélection positive pourraient être utiles pour cibler les gènes de nématodes sur lesquels porter les futures investigations de variabilité et d'évolution.

Pour la lutte génétique par association de gènes de résistance : qu'ils soit naturels ou artificiels, les gènes de résistance peuvent être combinés pour obtenir une résistance plus durable. Deux types d'associations peuvent être envisagés : soit les gènes de résistance sont accumulés dans une même plante, soit des plantes portant divers gènes de résistance se succèdent au champ lors des rotations de culture. Dans les deux cas, il est impossible de prévoir réellement la durabilité du système sans tests en conditions de cultures au champ. C'est pourquoi il est nécessaire de développer des connaissances qui permettront d'orienter *a priori* les associations de gènes à privilégier pour assurer une plus grande durabilité de ces résistances. Cette stratégie semble prometteuse puisque l'on peut considérer que plus le pathogène devra mettre en place de stratégies parasites pour se développer sur ce type de plantes, plus ce sera coûteux en terme d'énergie déployée à cela et plus il aura de difficultés à s'adapter.

### 3.3. APPORT DES RESULTATS POUR L'ETUDE DU FONCTIONNEMENT DES POPULATIONS

Les pressions de sélection diversifiantes auxquelles sont soumis les gènes du pouvoir pathogène sont observables à un temps "t", temps auquel nous avons décidé de les étudier. Nous avons ainsi déterminé à cet instant "t" que des acides aminés étaient sous sélection positive, c'est-à-dire qui ont été mutés et conservés au cours des générations. Comment vont évoluer ces gènes ? Est-ce que les mutations à venir vont être conservées ou vont-elles être contre sélectionnées au cours des générations ? Alors que notre approche se plaçait essentiellement à l'échelle inter ou intra spécifique, il est maintenant possible d'étendre ces investigations à une échelle plus fine, intra populationnelle pour suivre le devenir des allèles d'un gène au sein des populations (au fil des générations). Nous abordons alors une nouvelle approche d'étude de la génétique des populations basée sur les gènes du pouvoir pathogène. Nous avons maintenant les moyens d'amplifier différents gènes dans une large gamme de populations et d'espèces. Nous savons quelles sont les zones polymorphes et nous savons également quels sont les acides aminés sur lesquels il faut concentrer l'attention pour voir si les mutations se fixent dans les populations. Toutes les données préalables sont maintenant acquises et devraient permettre de donner une nouvelle dimension aux études de

génétique des populations qui pourront dépasser la seule utilisation de marqueurs neutres et utiliser des marqueurs soumis à sélection du fait de l'influence de la plante hôte.

# Références Bibliographiques

## REFERENCES BIBLIOGRAPHIQUES

(HORS REFERENCES CITEES DANS LES ARTICLES)

**Adoutte A, Balavoine G, Lartillot N, de Rosa R.** 1999. Animal evolution: the end of the intermediate taxa?. *Trends Genet.* 15:104-108.

**Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA.** 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature.* 387:489-493.

**Anthoine G.** Polymorphismes biologiques et moléculaires chez le complexe d'espèces *Nacobbus aberrans* (Thorne, 1935) Thorne & Allen, 1944 (Nematoda:Pratylenchidae). Université de Rennes 1/ENSAR. Thèse. 178p.

**Araki H, Tian D, Goss EM, Jakob K, Halldorsdottir SS, Kreitman M, Bergelson J.** 2006. Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a pathogen of Arabidopsis. *Proc Natl Acad Sci.* 103:5887-5892.

**Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP.** 1997. Signaling in Plant-Microbe Interactions. *Science.* 276:726-733.

**Bakhtia M, Charlton WL, Urwin PE, McPherson MJ, Atkinson HJ.** 2005. RNA interference and plant parasitic nematodes. *Trends Plant Sci.* 10:362-367.

**Barras F, Gisjsegem FV, Chatterjee AK.** 1994. Extracellular enzymes and pathogenesis of soft-rot *Erwinia*. *Annu Rev Phytopathol.* 32:201-234.

**Barret AJ, Rawling ND, Woessner JF.** 1998. *Hand-book of proteolytic enzymes.* Eds, Academic press, London.

**Baum TJ, Hiatt A, Parrott WA, Pratt LH, Hussey RS.** 1996. Expression in tobacco of a functional monoclonal antibody specific to stylet secretions of the root-knot nematode. *Mol Plant-Microbe Interact.* 9:382-387.

**Bekal S, Niblack TL, Lambert KN.** 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Mol Plant-Microbe Interact* 16:439-446.

**Blaxter M.** 1998. *Caenorhabditis elegans* is a Nematode. *Science* 282:2041-2046.

**Blumenthal T, Davis RE.** 2004. Exploring nematode diversity. *Nature Genet.* 36:1246-1247.

**Bossis M, Mugniéry D.** 1993. Specific status of six *Globodera* parasites of solanaceous plants studied by means of two-dimensional gel electrophoresis with a comparison of gel patterns by a computed system. *Fundam Appl Nematol.* 16:47-56.

**Bouarab K, Melton R, Peart J, Baulcombe D, Osbourn A.** 2002. A saponin-detoxification enzyme mediates suppression of plant defences. *Nature.* 418:889-892.

**Campos-Vela A.** 1967. Taxonomy, life cycle and host range of *Heterodera mexicana*. Sp. (Nematoda: Heteroderidae). University of Wisconsin. Thèse. 70p.

**Castadena O, Sotolongo V, Amor AM, Stöcklin R, Anderson AJ, Harvey AL, Engström A, Wernstedt C, and Karlson E.** 1995. Characterization of a potassium channel toxin from the caribbean sea anemaone *Stichodactyla helianthus*. *Toxicon* 33:603-613.

**Chen Q, Rehman S, Smant G, Jones JT.** 2005. Functional analysis of pathogenicity proteins of the potato cyst nematode *Globodera rostochiensis* using RNAi. *Molec Plant-Microbe Interact.* 18:621-625.

**Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D., Bieri T, Meyer RR, Ozersky P, Armstrong JR, Fulton RS, Latreille JP, Spieth J, Hooton TM, Mardis ER, Hultgren SJ, Gordon JI.** 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. *Proc Nat Acad Sci.* 103:5977-5982.

**Chisholm ST, Coaker G, Day B, Staskawicz BJ.** 2006. Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. *Cell.* 124:803-814.

**Combes C.** 1992. Etre un parasite et transmettre ses gènes. *Pour la science.* 174p.

**Combes C.** 2000. Pressions sélectives dans les systèmes parasites-hôtes. *Journal de la Société de Biologie.* 194:19-23.

**Combes C.** 2001. L'art d'être parasite. Les associations du vivant. Champs-Flammarion Eds. Paris. 362p.

**Dautova M, Rosso MN, Abad P, Gommers FJ, Bakker J, Smant G.** 2001. Single pass cDNA sequencing - a powerful tool to analyse gene expression in preparasitic juveniles of the southern root-knot nematode *Meloidogyne incognita*. *Nematol.* 3:129-139.

**Davis EL, Hussey RS, Baum TJ, Bakker J, Schots A, Rosso MN, Abad P.** 2000. Nematode parasitism genes. *Annu Rev Phytopathol.* 38:365-396.

**Davis EL, Hussey RS, Baum TJ.** 2004. Getting to the roots of parasitism by nematodes. *Trends Parasitol.* 20:134-141.

- De Boer JM, McDermott JP, Davis EL, Hussey RS, Popeijus H, Smant G, Baum JT.** 2002. Cloning of a putative pectate lyase gene expressed in the subventral esophageal glands of *Heterodera glycines*. *J Nematol.* 34:9-11.
- De Guiran G.** 1983. Les nématodes parasites des cultures. La littorale S.A. Eds. Bézier. 42p.
- Dibb NJ Newman AJ.** 1989. Evidence that introns arose at proto-splice sites. *EMBO J.* 8:2015-2021.
- Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CIA, Ayliffe MA, Kobe B, Ellis JG.** 2006. From the cover: direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Nat Acad Sci.* 103:8888-8893.
- Dorris M, De Ley P, Blaxter ML.** 1999. Molecular analysis of nematode diversity and the evolution of parasitism. *Parasitol Today.* 15:188-193.
- Doyle EA, Lambert KN.** 2002. Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. *Mol Plant-Microbe Interact.* 15:549-556.
- Duncan LH, Robertson L, Robertson WM, Kusel JR.** 1997. Isolation and characterization of secretions from the plant-parasitic nematode *Globodera pallida*. *Parasitol.* 115:429-438.
- Emberley E, Gietz RD, Campbell JD, HayGlass K, Murphy L, Watson P.** 2002. RanBPM interacts with psoriasin in vitro and their expression correlates with specific clinical features in vivo in breast cancer. *BMC Cancer* 2:28.
- Ferraz LCCB, Brown DJF.** 2002. An introduction to nematodes: plant nematology. Derek JF, Brown BA. Eds. 221p.
- Gao BL, Allen KD, Davis EL, Baum TJ, Hussey RS.** 2004. Developmental expression and biochemical properties of a beta-1,4-endoglucanase family in the soybean cyst nematode, *Heterodera glycines*. *Mol Plant Pathol.* 5:93-104.
- Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS.** 2001. Identification of putative parasitism genes expressed in the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. *Mol Plant-Microbe Interact.* 14:1247-1254.
- Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS.** 2002. Identification of a new beta-1,4-endoglucanase gene expressed in the esophageal subventral gland cells of *Heterodera glycines*. *J Nematol.* 34:12-15.
- Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS.** 2003. The parasitome of the phytonematode *Heterodera glycines*. *Mol Plant-Microbe Interact.* 16:720-726.
- George HL, VanEtten HD.** 2001. Characterization of pisatin-inducible cytochrome p450s in fungal pathogens of pea that detoxify the pea phytoalexin pisatin. *Fungal Genet Biol.* 33:37-48.



- Gheysen G, Fenoll C.** 2002. Gene expression in nematode feeding sites. *Annu Rev Phytopathol.* 40:191-219.
- Goellner M, Wang XH, Davis EL.** 2001. Endo-beta-1,4-glucanase expression in compatible plant-nematode interactions. *Plant Cell.* 13:2241-2255.
- Goellner M, Smant G, De Boer JM, Baum TJ, Davis EL.** 2000. Isolation of beta-1,4-endoglucanase genes from *Globodera tabacum* and their expression during parasitism. *J Nematol.* 32:154-165.
- Goetze E.** 2006. Elongation factor 1-[alpha] in marine copepods (Calanoida: Eucalanidae): Phylogenetic utility and unique intron structure. *Mol Phylogenet Evol.* 40:3, pp. 880-886.
- Goldman N, Yang Z.** 1994. A codon-based model of nucleotide substitution for protein coding DNA sequences. *Mol Biol Evol.* 11:725-736.
- Gouyon PH.** 2001. Les harmonies de la nature à l'épreuve de la biologie. Evolution et biodiversité. INRA Eds. Paris. 91p.
- Goverse A, de Almeida Engler J, Verhees J, van der Krol S, Helder J, Gheysen G.** 2000. Cell cycle activation by plant parasitic nematodes. *Plant Mol Biol.* 43:747-761.
- Greenbaum D, Luscombe NM, Jansen R, Qian J, Gerstein M.** 2001. Interrelating Different Types of Genomic Data, from Proteome to Secretome: 'Oming in on Function. *Genome Res.* 11:1463-1468.
- Grenier E, Blok VC, Jones JT, Fouville D, Mugniéry D.** 2002. Identification of gene expression differences between *Globodera pallida* and *G. mexicana* by suppression subtractive hybridization. *Mol Plant Pathol.* 3:217-226.
- Guiliano DB, Hong X, McKerrow JH, Blaxter M, Oksov Y, Liu J, Ghedin E, Lustigman S.** 2004. A gene family of cathepsin L-like proteases of filarial nematodes are associated with larval molting and cuticle and eggshell remodeling. *Mol Biochem Parasitol.* 136:227-242.
- Hafizi S, Gustafsson A, Stenhoff J, Dahlback B.** 2005. The Ran binding protein RanBPM interacts with Axl and Sky receptor tyrosine kinases. *The International Journal of Biochemistry & Cell Biol.* 37:2344-2356.
- Halton DW.** 1997. Nutritional adaptations to parasitism within the platyhelminthes. *Int J Parasitol.* 27:693-704.
- Haq SK, Atif SM, Khan RH.** 2004. Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. *Arch Biochem Biophys.* 431:145-159.
- Heald R, Weis K.** 2000. Spindles get the Ran around. *Trends Cell Biol.* 10:1-4.

- Hofmann J, Grundler FMW.** 2006. Females and males of root-parasitic cyst nematodes induce different symplasmic connections between their syncytial feeding cells and the phloem in *Arabidopsis thaliana*. *Plant Physiol Biochem.* 44:430-433.
- Holterman M, van der Wurff A, van den Elsen S, van Megen H, Bongers T, Holovachov O, Bakker J, Helder J.** 2006. Phylum-wide analysis of ssu rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Mol Biol Evol.* 23:1792-1800.
- Hu KJ, Leung PC.** 2006. Complete, precise, and innocuous loss of multiple introns in the currently intronless, active cathepsin L-like genes, and inference from this event. *Mol Phylogenet Evol.* 38:685-696.
- Huang G, Dong R, Allen R, Davis EL, Baum TJ, Hussey RS.** 2005. Developmental expression and molecular analysis of two *Meloidogyne incognita* pectate lyase genes. *Int J Parasitol.* 35:685-692.
- Huang GZ, Gao BL, Maier T, Allen R, Davis EL, Baum TJ, Hussey RS.** 2003. A profile of putative parasitism genes expressed in the esophageal gland cells of the root-knot nematode *Meloidogyne incognita*. *Mol Plant-Microbe Interact.* 16:376-381.
- Hugot JP, Baujard P, Morand S.** 2001. Biodiversity in helminths and nematodes as a field of study: an overview. *Nematol.* 3:199-208.
- Jaubert S, Ledger TN, Laffaire JB, Piotte C, Abad P, Rosso MN.** 2002. Direct identification of stylet secreted proteins from root-knot nematodes by a proteomic approach. *Mol Biochem Parasitol* 121:205-211.
- Jaubert S, Milac AL, Petrescu AJ, de Almeida-Engler J, Abad P, Rosso MN.** 2005. *In planta* secretion of a calreticulin by migratory and sedentary stages of root-knot nematode. *Mol Plant-Microbe Interact.* 18:1277-1284.
- Jones TJ, Furlanetto C, Kudla T.** 2005. Horizontal gene transfer from bacteria and fungi as a driving force in the evolution of plant parasitism. *Nematology.* 7:641-646.
- Jones DA, Takemoto D.** 2004. Plant innate immunity - direct and indirect recognition of general and specific pathogen-associated molecules. *Curr Opin Immunol.* 16:62.
- Jones JT, Furlanetto C, Bakker E, Banks B, Blok VC, Chen Q, Phillips MS, Prior A.** 2003. Characterization of a chorismate mutase from the potato cyst nematode *Globodera pallida*. *Mol Plant Pathol.* 4:43-50.
- Juergensen K, Joachim SS, Sauer N, Hess P, Van Bel AJE, Grundler FMW.** 2003. The companion cell-specific *Arabidopsis* disaccharide carrier AtSUC2 is expressed in nematode-induced syncytia. *Plant Physiol.* 131:61-69.

- Kevan DKM.** 1965. The soil fauna-its nature and biology. P.33-51 in: Ecology of soil-borne plant pathogens. Baker KF and Snyder WC, Eds. University of California Press. Berkley.
- Kikuchi T, Shibuya H, Aikawa T, Jones J.** 2006. Cloning and Characterization of pectate lyases expressed in the esophageal gland of the pine wood nematode *Bursaphelenchus xylophilus*. Mol Plant-Microbe Interact. 19:280-287.
- Kimura M.** 1968. Evolutionary rate at the molecular level. Nature. 217: 624-626.
- Kimura M.** 1983. The Neutral theory of molecular evolution. Cambridge University Press, Cambridge.
- Koritas VM, Atkinson HJ.** 1994. Proteinases of females of the phytoparasitic *Globodera pallida* (potato cyst nematode). Parasitol, 109:357-365.
- Kort J, Ross H, Rumpfenhorst HJ, Stone AR.** 1977. An international scheme for identifying and classifying pathotypes of potato cyst nematodes *Globodera rostochiensis* and *G. pallida*. Nematologica. 23:333-339.
- Kudla U, Qin L, Milac A, Kielak A, Maissen C, Overmars H, Popeijus H, Roze E, Petrescu A, Smart G.** 2005. Origin, distribution and 3D-modeling of Gr-EXPB1, an expansin from the potato cyst nematode *Globodera rostochiensis*. FEBS Letters. 579:2451-2457.
- Kurland CG, Canback B, Berg OG.** 2003. Horizontal gene transfer: A critical view. Proc Nat Acad Sci. 100:9658-9662.
- Lambert KN, Allen KD, Sussex IM.** 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. Mol Plant-Microbe Interact. 12:328-336.
- Lambert KN, Bekal S, Domier LL, Niblack TL, Noel GR, Smyth CA.** 2005. Selection of *Heterodera glycines* chorismate mutase-1 alleles on nematode-resistant soybean. Mol Plant-Microbe Interact. 18:593-601.
- Ledger TN, Jaubert S, Bosselut N, Abad P, Rosso MN.** 2006. Characterisation of a new beta-1,4-endoglucanase gene from the root-knot nematode *Meloidogyne incognita* and evolutionnary scheme for phytonematode family 5 glycosyl hydrolases. Gene 382:121-128.
- Lilley CJ, Urwin PE, McPherson MJ, Atkinson HJ.** 1996. Characterization of intestinally active proteinases of cyst nematodes. Parasitol. 113:415-424.
- Lownsbery BF, Lownsbery JW.** 1954. *Heterodera tabacum* new species, a parasite of Solanaceous plant in Connecticut. Proc Helminth Soc Wash. 21:42-47.

- Marché L, Valette S, Grenier E, Mugniéry D.** 2001. Intra-species DNA polymorphism in the tobacco cyst-nematode complex (*Globodera tabacum*) using AFLP. *Genome*. 44:941-946.
- Mayr E.** 1974. Populations, espèces et évolution. Hermann, Paris, p. 496.
- Menard R, Carmona E, Takebe S, Dufour E, Plouffe C, Mason P, Mort JS.** 1998. Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro. *J Biol Chem*. 273:4478-4484.
- Merrick WC.** 1992. Mechanism and regulation of eukaryotic protein synthesis. *Microbiol Rev*. 56:291-315.
- Michaud D, Cantin L, Bonade-Bottino M, Jouanin L, Vrai TC.** 1996. Identification of stable plant cystatin/nematode proteinase complexes using mildly denaturation gelatin/polyacrylamide gel electrophoresis. *Electrophoresis*. 17:1373-1379.
- Mitreva M, Wendl M, Martin J, Wylie T, Yin Y, Larson A, Parkinson J, Waterston R, McCarter J.** 2006. Codon usage patterns in Nematoda: analysis based on over 25 million codons in thirty-two species. *Genome Biol*. 7:R75.
- Moczon T.** 1994. A cysteine proteinase in the cercariae of *Diplostomum pseudospathaceum* (Trematoda, Diplostomatidae). *Parasitol Res*. 80:680-683.
- Mugniéry D, Bossis M, Pierre JS.** 1992. Hybridations entre *Globodera rostochiensis* (Wollenweber), *G. pallida* (Stone), *G. virginiae* (Miller & Gray), *G. solanacearum* (Miller & Gray) et *G. "mexicana"* (Campos-Vela). Description et devenir des hybrides. *Fundam Appl Nematol*. 15:375-382.
- Mugniéry D, Fouville D, Dantec JP, Pellé R, Rousselle-Bourgeois F, Ellissèche D.** 2001. Résistance à *Globodera pallida* Pa2/3 chez *Solanum sparsipilum*. *Nematol*. 3:619-626.
- Mugniéry D, Phillips M.** Sous presse. The nematodes parasites of potato. Dans: Potato biology and biotechnology. D.Vrengdenbil Eds. Elsevier B.V. publisher.
- Müller J.** 1992. Detection of pathotypes by assessing the virulence of *Heterodera schachtii* populations. *Nematologica*. 38:50-64.
- Neveu C, Abad P, Castagnone-Sereno P.** 2003. Molecular cloning and characterization of an intestinal cathepsin L protease from the plant-parasitic nematode *Meloidogyne incognita*. *Physiol Mol Plant Pathol*. 63:159-165.
- Nielsen R, Yang Z.** 1998. Likelihood models for detecting positive selected amino acids sites and applications to the HIV-1 envelop gene. *Genetics* 148:929-936.

- Nishitani H, Hirose E, Uchimura Y, Nakamura M, Umeda M, Nishii K, Mori N, Nishimoto T. 2001. Full-sized RanBPM cDNA encodes a protein possessing a long stretch of proline and glutamine within the N-terminal region, comprising a large protein complex. *Gene*. 272:25-33.
- Nomura K, Melotto M, He SY. 2005. Suppression of host defense in compatible plant-*Pseudomonas syringae* interactions. *Curr Opin Plant Biol*. 8:361-368.
- Ortiz M, Bleiber G, Martinez R, Kaessmann H, Telenti A. 2006. Patterns of evolution of host proteins involved in retroviral pathogenesis. *Retrovirol*. 3:11.
- Picard D, Plantard O, Scurrah M, Mugniery D. 2004. Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru). *Mol Ecol*. 13:2899-2908.
- Picard D, Plantard O. 2006. What constitutes a population for the plant parasitic nematode *Globodera pallida* in its native area (Peru)? *Int J Parasitol*. 36:115-122.
- Picard D. 2005. Génétique des populations et phylogéographie du nématode à kyste de la pomme de terre (*Globodera pallida*) au Pérou. Université de Rennes 1/Agrocampus. Thèse. 185p.
- Piotte C, Arthaud L, Abad P, Rosso MN. 1999. Molecular cloning of an acetylcholinesterase gene from the plant parasitic nematodes, *Meloidogyne incognita* and *M. javanica*. *Mol Biochem Parasitol*. 99:247-256.
- Popeijus H, Overmars HA, Jones J, Blok VC, Goverse A, Helder J, Schots A, Bakker J, Smant G. 2000. Degradation of plant cell walls by a nematode. *Nature*. 406:36-37.
- Pouttu R, Puustinen T, Virkola R, Hacker J, Klemm P, Korhonen TK. 1999. Amino acid residue Ala-62 in the *FimH* fimbrial adhesin is critical for the adhesiveness of meningitis-associated *Escherichia coli* to collagens. *Mol Microbiol*. 31:1747-1757.
- Qin L, Kudla U, Roze EHA, Goverse A, Popeijus H, Nieuwland J, Overmars H, Jones JT, Schots A, Smant G, Bakker J, Helder J. 2004. A nematode expansin acting on plants. *Nature*. 427:30.
- Qin L, Overmars B, Helder J, Popeijus H, Rouppe van der Voort JN, Groenink W, Van Koert PH, Schots A, Bakker J, Smant G. 2000. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. *Mol Plant-Microbe Interact*. 13:830-836.
- Qin L, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J. 2001. GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. *Nucleic Acids Res*. 29:1616-1622.
- Robertson L, Robertson WM, Jones JT. 1999. Direct analysis of the secretions of the potato cyst nematode *Globodera rostochiensis*. *Parasitol*. 119:167-176.

**Robertson L, Robertson WM, Sobczak M, Helder J, Tetaud E, Ariyanayagam MR, Ferguson MA, Fairlamb A, Jones JT.** 2000. Cloning, expression and functional characterisation of a peroxiredoxin from the potato cyst nematode *Globodera rostochiensis*. *Mol Biochem Parasitol.* 111:41-49.

**Rogers LM, Kim YK, Guo W, Gonzalez-Candelas L, Li D, Kolattukudy PE.** 2000. Requirement for either a host- or pectin-induced pectate lyase for infection of *Pisum sativum* by *Nectria haematococca*. *Proc Nat Acad Sci.* 97:9813-9818.

**Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV.** 2003. Remarkable Interkingdom Conservation of Intron Positions and Massive, Lineage-Specific Intron Loss and Gain in Eukaryotic Evolution. *Curr Biol.* 13:1512-1517.

**Rokas A, Kathirithamby J, Holland PW.** 1999. Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. *Insect Mol Biol* 8:527-530.

**Rokas A, Nylander JAA, Ronquist F, Stone GN.** 2002. A maximum-likelihood analysis of eight phylogenetic markers in gallwasps (hymenoptera: cynipidae): implications for insect phylogenetic studies. *Mol Phylogenet Evol.* 22:206-219.

**Rosso MN, Dubrana MP, Cimblolini N, Jaubert S, Abad P.** 2005. Application of RNA interference to root-knot nematode genes encoding esophageal gland proteins. *Mol Plant Microbe Interact.* 18:615-620.

**Rosso MN, Scouten A, Roosien J, Borst-vernessen T, Hussey RS, Gommers FJ, Schots A, Abad P.** 1996. Expression and functional characterization of a single chain FV antibodies directed against secretions involved in plant nematode infection process. *Biochem Biophys Res Commun.* 220:255-263.

**Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R.** 2003. DnaSP: Polymorphism analyses of the coalescent and other methods. *Bioinformatics.* 19: 2496-2497.

**Sajid M, McKerrow JH.** 2002. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol.* 120:1-21.

**Sanchez D, Ganfornina MD, Gutierrez G, Marin A.** 2003. Exon-intron structure and evolution of the lipocalin gene family. *Mol Biol Evol.* 20:775-783.

**Sasser JN, Freckman DW.** 1987. A world perspective on nematology: the role of the society. Dans : *Visitas on nematology.* Veech, J.A. et Dickson, D.W. Eds. p. 7-14. Hyattsville, USA, society of nematologist inc.

**Sawires YS, Songer JG.** 2006. *Clostridium perfringens*: Insight into virulence evolution and population structure. *Anaerobe.* 12:23-43.

**Scholte K.** 2000. Screening of non-tuber bearing Solanaceae for resistance to and induction of juvenile hatch of potato cyst nematodes and their potential for trap cropping. *Annu Appl Biol* 136:239-246.

**Shevchik VE, Robert-Baudouy J, Hugouvieux-Cotte-Pattat N.** 1997. Pectate lyase Pell of *Erwinia chrysanthemi* 3937 belongs to a new family. *J Bacteriol.* 179:7321-7330.

**Shingles J, Lilley CJ, Atkinson HJ, Urwin PE.** 2006. *Meloidogyne incognita*: Molecular and biochemical characterisation of a cathepsin L cysteine proteinase and the effect on parasitism following RNAi. *Exp Parasitol.* In Press, Corrected Proof.

**Sijmons PC, Atkinson HJ, Wyss U.** 1994. Parasitic strategies of root nematodes and associated host cell responses. *Annu Rev Phytopathol.* 32:235-259.

**Silva FB, Batista JAN, Marra BM, Fragoso RR, Monteiro ACS, Figueira ELZ, Grossi-de-Sa MF.** 2004. Pro domain peptide of HGCP-Iv cysteine proteinase inhibits nematode cysteine proteinases. *Genet Mol Res.* 3:342-355.

**Slobin LI.** 1980. The role of eucaryotic factor Tu in protein sunthesis. The mesurement of the elongation factor Tu content of rabbit reticulocytes and other mammalian cells by a sensitive radioimmunoassay. *Eur J Biochem.* 110:555-563.

**Smant G, Stokkermans JP, Yan Y, de Boer JM, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, Bakker J.** 1998. Endogenous cellulases in animals: Isolation of beta -1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc Nat Acad Sci.* 95:4906-4911.

**Stone AR.** 1972. *Heterodera pallida* n. sp. (Nematode: Heteroderidae), a second species of potato cyst nematode. *Nematologica.* 18:591-606.

**Suarez CE, Norimine J, Lacy P, McElwain TF.** 2006. Characterization and gene expression of *Babesia bovis* elongation factor-1[alpha]. *Int J Parasitol.* 36:965-973.

**Subbotin SA, Sturhan D, Chizov V, Volvas N, Baldwin JG.** 2006. Phylogenetic analysis of *Tilenchida* Thorne, 1949 as inferred from D2 and D3 expansion fragments of the 28S rRNA gene sequence. *Nematol.* 8:455-474.

**Subbotin SA, Vierstraete A, De Ley P, Rowe J, Waeyenberge L, Moens M, Vanfleteren JR.** 2001. Phylogenetic Relationships within the Cyst-Forming Nematodes (Nematoda, Heteroderidae) Based on Analysis of Sequences from the ITS Regions of Ribosomal DNA. *Mol Phylogenet Evol.* 21:1-16.

**Thiery M, Mugniéry D, Bossis M, Sosa-Moss C.** 1997. Résultats de croisements entre *Globodera pallida* Stone et *G. "mexicana"* Campos-Vela : héritabilité du développement sur pomme de terre et notion d'espèce. *Fundam Appl Nematol.* 20:551-556.

- Thomas JH.** 2006. Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Res.* 16:1017-1030.
- Tsai YH, Orsi RH, Nightingale KK, Wiedmann M.** 2006. *Listeria monocytogenes* internalins are highly diverse and evolved by recombination and positive selection. *Infect Genet Evol.* 6:378-389.
- Tudor JE, Pennington MW, Norton R.** 1998. Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin. *Eur J Biochem.* 251:133-141.
- Tytgat T, Vanholme B, De Meutter J, Claeys M, Couvreur M, Vanhoutte I, Gheysen G, Criekinge WV, Borgonie G, Coomans A, Gheysen G.** 2004. A new class of ubiquitin extension proteins secreted by the dorsal pharyngeal gland in plant parasitic cyst nematodes. *Mol Plant-Microbe Interact.* 17:846-852.
- Urwin PE, Lilley CJ, McPherson MJ, Atkinson HJ.** 1997. Characterization of two cDNAs encoding cystéine proteinases from the soybean cyst nematode *Heterodera glycines*. *Parasitol.* 114:605-613.
- Urwin PE, Troth KM, Zubko EI, Atkinson HJ.** 2001. Effective transgenic resistance to *Globodera pallida* in potato field trials. *Mol Breeding.* 8:95-101.
- Vanholme B, De Meutter J, Tytgat T, Van Montagu M, Coomans A, Gheysen G.** 2004. Secretions of plant-parasitic nematodes: a molecular update. *Gene.* 332:13-27.
- Van Valen L.** 1973. A new evolutionnary law. *Evol Theory.* 1:1-30.
- Vercauteren I, de Almeida Engler J, De Groot R, Geyhsen G.** 2002. An *Arabidopsis thaliana* pectin acetylerase gene is upregulated in nematode feeding sites induced by root-knot and cyst nematodes. *Mol Plant-Microbe Interact.* 14:404-407.
- Wada H, Kobayashi M, Sato R, Satoh N, Miyasaka H, Shirayama Y.** 2002. Dynamic insertion-deletion of introns in deuterostome EF-1 alpha genes. *J Mol Evol.* 54:118-128.
- Wang D, Li Z, Messing EM, Wu G.** 2002. Activation of Ras/Erk pathway by a novel MET-interacting protein RanBPM. *J Biol Chem.* 277:36216-36222.
- Wang X, Meyers D, Yan T, Baum T, Smant G, Hussey R, Davis E.** 1999. *In planta* localization of a beta-1,4-endoglucanase secreted by *Heterodera glycines*. *Mol Plant-Microbe Interact.* 12:64-67.
- Wang XIAO, Mitchum MG, Gao BING, Li CHUN, Diab HANA, Baum TJ, Hussey RS, Davis EL.** 2005. A parasitism gene from a plant-parasitic nematode with function similar to CLAVATA3/ESR (CLE) of *Arabidopsis thaliana*. *Mol Plant Pathol.* 6:187-191.
- Weber E, Koebnik R.** 2006. Positive selection of the Hrp pilin HrpE of the plant pathogen *Xanthomonas*. *J Bacteriol.* 188:1405-1410.



**Wieczorek K, Golecki B, Gerdes L, Heinen P, Szakasits D, Durachko DM, Cosgrove DJ, Kreil DP, Puzio PS, Bohlmann H, Grundler FMW.** 2006. Expansins are involved in the formation of nematode-induced syncytia in roots of *Arabidopsis thaliana*. *Plant J.* 48:98-112.

**Williamson VM, Kumar A.** 2006. Nematode resistance in plants: the battle underground. *Trends Genet.* 22:396-403.

**Wollenweber HW.** 1923. Krankheiten und beschädigungen der kartoffeln. *Abh Forschungsinst Kartoffelbau Heft.* 7, Berlin, 56 p.

**Wyss Y, Gründler FMW.** 1992. Feeding behaviour of sedentary plant parasitic nematodes. *Netherlands J Plant Pathol.* 98:165-173.

**Yan Y, Smant G, Davis E.** 2001. Functional screening yield a new beta-1,4-endoglucanase gene from *Heterodera glycines* that may be the product of recent gene duplication. *Mol Plant-Microbe Interact.* 14:63-71.

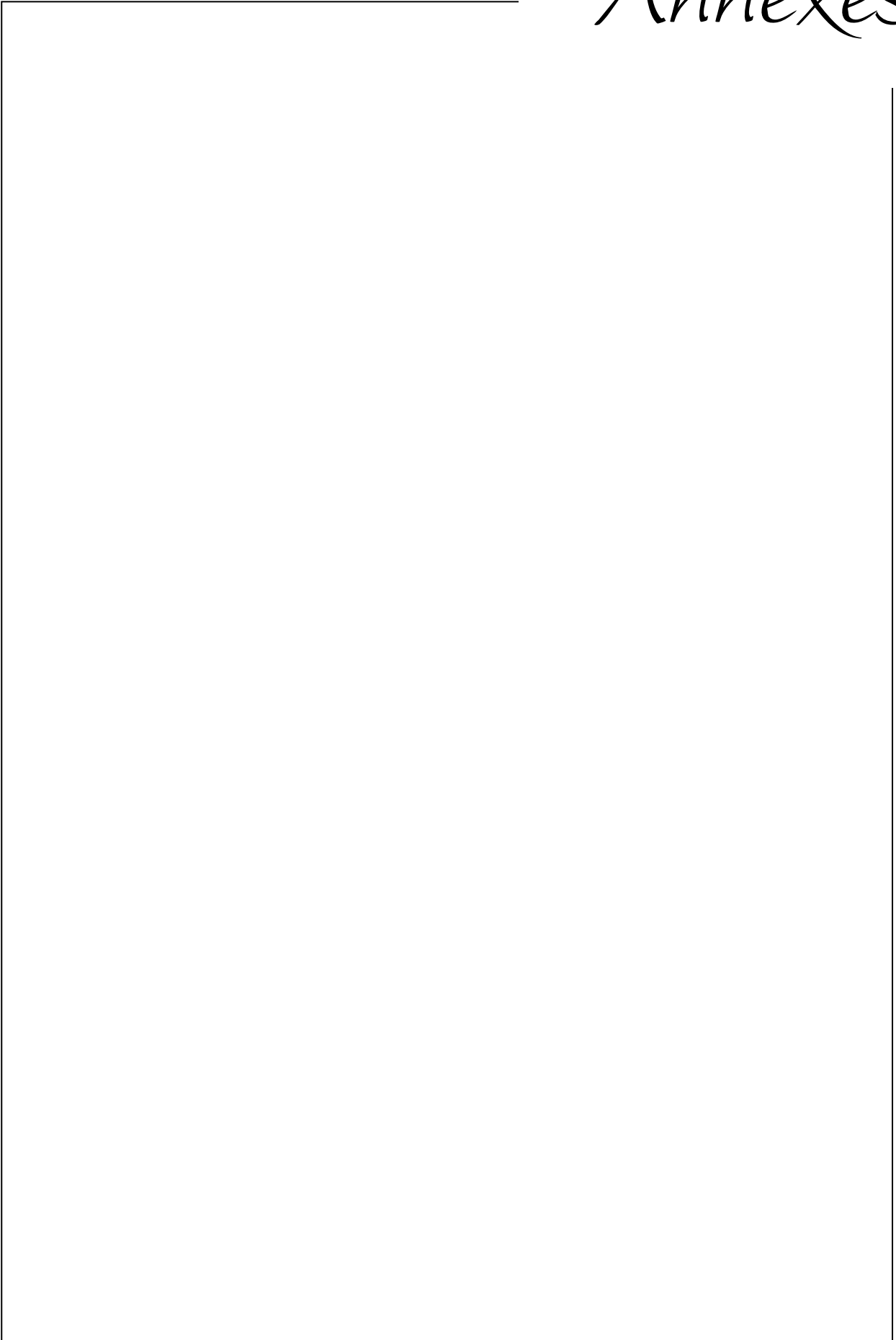
**Yang Z.** 1997. Phylogenetic analysis by maximum likelihood (PAML). Version 1.3. Department of integrative biology. University of California at Berkeley.

**Yang Z.** 1998. Likelihood ratio test for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol.* 15:568-573.

**Yang Z, Nielsen R.** 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J Mol Evol.* 46:409-418.

**Yang Z, Nielsen R.** 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineage. *Mol Biol Evol* 19:908-917.

# *Annexes*



*Annexe 1:*  
*Amorces*

GENE	NOM DES AMORCES	SEQUENCE DES AMORCES 5'→3'
<b>IA7</b>	IA7Fwd2 IA7Rev 5'IA7 3'IA7 GPLIA7Rev2	CGTCGGCCTCTAGCTGTA TGCCTCTTTATCTTTGAAAATG AGCCGTCTGAGTCCAATG TTGATTTCTGTCTTCAGCAAAA GTTTTTGGGCACATGGTTG
<b>IB3</b>	IB3Fwd IB3Rev	TGATATGCCGACAGCTCAAC ACCGACCACACCATCATCTT
<b>IVG9</b>	IVG9Fwd IVG9Rev 5'IVG9.2 3'IVG9.2	AAACAACGTTTGCCGAATG GCGCAACATTCAACGACA ATGACAAAAATTGTGTTTTTCCT CGAATGTACACATATACCAAT
<b>Ic5</b>	GPLIC5Fwd GPLIC5Rev2 GPLIC5Rev3 5'IC5 5'IC5.2 3'IC5.3 3'IC5.2	GGCTATTTTTGGGGTCAC CGTATGAGTAAGTGCCT TCACTTCATAGTAGAAAAATGCC TTTTTATTGCCCCCAAAATG AAATGCGCACCTTTCTCTTC GCAAACCCATCATAAATTCTCG CATGATCAATAGGCAAACCG
<b>PECTATE LYASE</b>	PLFwd.2 PLRev PLRev.2 PLFwd.4 PLRev.3 PLFwd5 PLRev4	GGCGGTTCCGGCATT CCCGAAAGTGACATTTTTGT CATTTTGTACCAGTAGTTGGA ATGCTTTTTGTTATCATTTCAAT TAGTTGACAATTTTAAAYWGCCGA ATGCTTTTCATTCTACTGGTCAT TTAGCTCGCAATTTTAACTGCC
<b>CATHEPSINE L</b>	CathepFwd CathepFwd5 CathepRev5	CGCGAGCGTGAATTGA CGTGATCGTGCCATAGA CAACTGTTCTTACAATCCA
<b>EF</b>	EF1wd EF1Rev	GCTTGGGTGYTKACAAR CATCYTRTCAGACGGCTCCA
<b>AUTRES AMORCES</b>	AUAP AAP GPFARF 3'GPFAR E11Fwd E11Rev D10Fwd D10Rev H09Fwd H09 Rev	GGCCACGCGTCGACTAGTAC GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG TCGCTGCCGCCATTGAT GGCGCCCTCAGTCTTCAGCA TTGTCAAACATAATGGAAGGA TTCAATTTTGTGCCAAGC CCAAAACCGTTAAACACCAA CCTAGTCCAATCGCTCTTCC CACTTACGCATACGGCAATC TTTCTCCGAGGTCTTTGCAT







*Annexe 2 :*  
*Matériel et Méthodes*



# CREATION ET CIBLAGE DE LA BANQUE SSH

## 1. ELABORATION DE LA BANQUE cDNA

Cette banque a été générée au SCRI à partir d'un pool de seconds stades larvaires de *G. pallida* de la population Bedale. Les ARNm ont été extraits à l'aide du kit Micro-Fast Track mRNA isolation 2.0 (INVITROGEN) en suivant les instructions du fabricant. Le culot contenant les ARNm a été séché pendant 10 min à 55°C. Les cDNA double brin ont alors pu être synthétisés grâce au kit Creator SMART cDNA library construction (CLONTECH). Tous les cDNA obtenus ont été clonés dans des plasmides pBluescript II SK (+) afin de générer la banque par transformation (électroporation) de bactéries électrocompétentes.

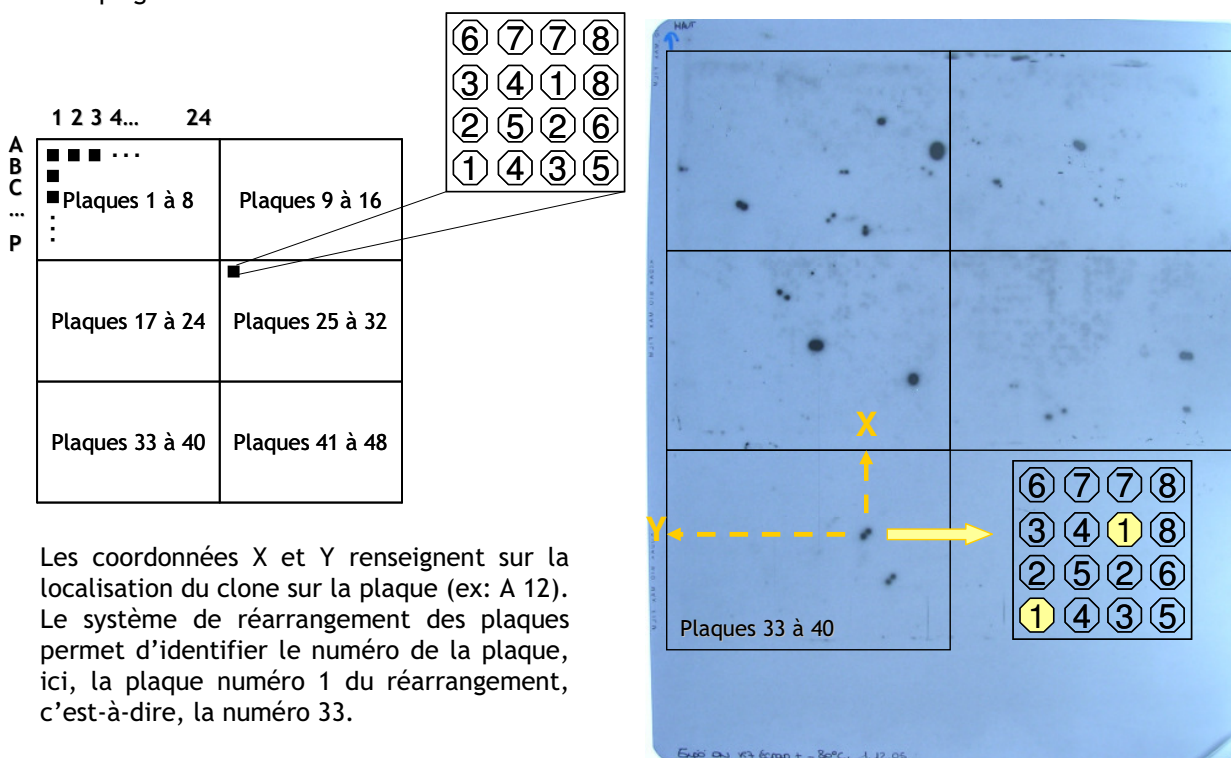
La banque est constituée de 33 plaques de 384 puits soit 12672 clones bactériens stockés dans du milieu de conservation (freezing solution) contenant de l'ampicilline.

Freezing solution : 360mM K<sub>2</sub>HPO<sub>4</sub>, 132 mM KH<sub>2</sub>PO<sub>4</sub>, 17mM Sodium citrate, 4mM MgSO<sub>4</sub>, 68mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, Glycérol 44mL pour 100mL de solution

## 2. SCREENING DE LA BANQUE cDNA

### a- Transfert des colonies sur membrane de nylon

Les 33 plaques sont repiquées sur une seule membrane par le robot Q-Bot. La membrane sur laquelle les colonies sont repiquées est posée sur un milieu LB/ampicilline. Les colonies sont repiquées en doublons et ordonnées de manière à pouvoir identifier la position des clones positifs après marquage.

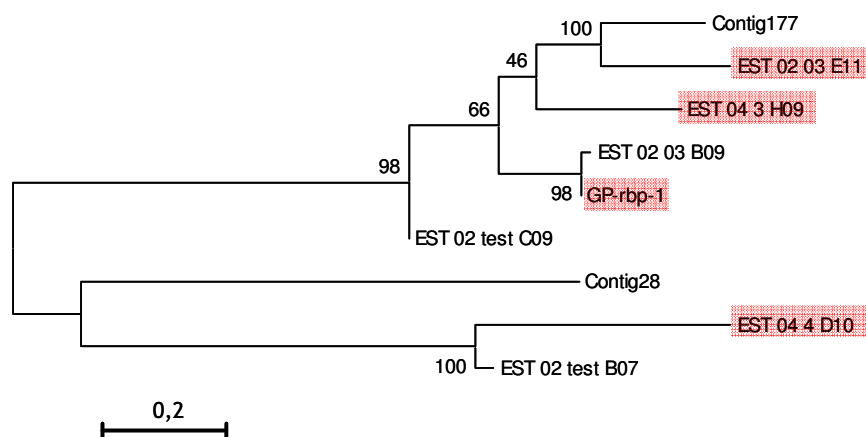


Agencement des colonies déposées en doublon sur les membranes de nylon

Après une nuit de culture, les membranes sont récupérées et sont traitées de manière à fixer l'ADN plasmidique sur la membrane (SDS 1% 5min, Dénaturation - 1,5M NaCl, 0,5M NaOH - 5min, Neutralisation - 3.0M NaCl, 0,5M Tris - 5min). La membrane est ensuite lavée dans du SSC 2X côté colonies pendant 5 min en agitant un peu pour éliminer les débris cellulaires bactériens. L'ADN plasmidique est ensuite fixé sur les membranes par cuisson à 80°C pendant 1 à 2h. Les membranes doivent être ensuite complètement séchées avant stockage.

### b- Choix et marquage des sondes

Parmi les séquences de RanBPM disponibles chez *G. pallida* (J. Jones, comm. Pers.), nous avons choisi les 3 plus divergentes ainsi que *rbp-1* pour générer les sondes qui nous ont servi au criblage de la banque cDNA (voir ci-dessous).



Les sondes ont été amplifiées par PCR à l'aide d'amorces spécifiques à partir d'ADNc de *G. pallida* de la population Bedale. Les produits PCR ont été mixés et utilisés pour le marquage radioactif. Pour cela, le volume total des sondes est ajusté à 45µl puis le tube est placé à 100°C pendant 5 min. Un kit de marquage au <sup>32</sup>P est utilisé selon les indications du fournisseur (Amersham). Le marquage se fait à 37°C pendant 1h.

### c- Hybridation des membranes

Pour l'étape de pré hybridation, le tampon d'hybridation (Pour 50mL : 12,5 mL SSC20X, 2,5mL Denhardt's 100X, 2,5 mL SDS 10%, 32,5mL eau distillée stérile) est préchauffé à 65°C. L'ADN de poisson soniqué est dénaturé et est ajouté au tampon d'hybridation. Le mélange est ajouté dans les tubes contenant les membranes et est incubé 1h à 65°C.

L'hybridation se fait par ajout des sondes marquées et dénaturées dans le tube contenant les membranes recouvertes de tampon d'hybridation. L'hybridation est réalisée pendant la nuit à 65°C.

Après hybridation, les lavages sont ensuite réalisés de la façon suivante (les étapes de lavage nécessitent de préchauffer les solutions à 55°C) :

- 5 à 10 min dans SSC 2X + SDS 0,1%
- 2 lavages de 15 min avec SSC 1X + SDS 0,1%
- 2 lavages de 15 min avec SSC 0,1X + SDS 0,1%

Enfin, l'hybridation est ensuite révélée par exposition à des films X-ray.

# 5'RACE

## 1. REVERSE TRANSCRIPTION

Amorce 10 $\mu$ M	1 $\mu$ l	} 70° C 10 min puis 1 minute sur la glace
ARN	6 $\mu$ l	
Eau RNase free	8,5 $\mu$ l	

- Ajouter : 2,5 $\mu$ l de tampon 10X, 2,5 $\mu$ l de MgCl<sub>2</sub> 25mM, 1 $\mu$ l de dNTP 10mM chaque, 2,5 $\mu$ l de DTT 0,1M
- Incuber 1 minute à 42° C
- Ajouter 1 $\mu$ l de Superscript III
- Incuber 50 minutes à 42° C
- Inhiber l'activité de l'enzyme par incubation 15 minutes à 70° C
- Centrifuger 10 à 20 secondes et placer à 37° C
- Ajouter 1 $\mu$ l de RNase Mix et incuber 30 minutes à 37° C
- Centrifuger et placer sur la glace
- Conservation : -20° C

## 2. PURIFICATION

Kit : GLASSMAX DNA Isolation Spin Cartidge Purification of cDNA

- Placer la solution « binding » à température ambiante (15 à 20 min)
- Pour chaque échantillon, placer 100 $\mu$ l d'eau distillée stérile à 65° C
- Ajouter 120 $\mu$ l de solution « binding » (6M NaCl) à la réaction de RT
- Transférer dans la cartouche et centrifuger 20 sec à 13000 rpm
- Récupérer le liquide dans un tube neuf (jusqu'à la vérification de l'obtention d'un cDNA)
- Ajouter 400 $\mu$ l de tampon de lavage 1X à 4° C et centrifuger 20sec à 13000 rpm. Renouveler l'opération
- 3 fois
- 2 lavages avec 400 $\mu$ l d'éthanol 70%
- Centrifuger à vide 1 min à 13000 rpm
- Transférer la colonne dans un nouveau tube. L'élution est réalisée avec 2x30 $\mu$ l d'eau distillée stérile, par centrifugation 20 sec à 13000rpm

## 3. ACCROCHAGE DE LA QUEUE POLYC

- Ajouter 5 $\mu$ l de tampon « tailing » 5X, 2,5 $\mu$ l de dCTP 2mM, 16,5 $\mu$ l de cDNA purifié
- Incuber 3 min à 94° C
- Ajouter 1 $\mu$ l TdT (Terminal deoxynucleotidyl transferase) et incuber 10 min à 37° C

- Inactiver la TdT par passage 10 min à 65° C
- Centrifuger et placer sur la glace
- Conservation à -20° C

#### 4. CONDITIONS PCR

Tampon PCR 10X	5µl
MgCl <sub>2</sub> 25mM	3µl
Mix dNTPs 10mM chaque	1µl
Amorce interne 10µM	2µl
AAP 10µM	2µl
cDNA_polyC	5µl
Taq polymérase 5U/µl	0,5µl
Eau distillée stérile	31,5µl
	Total 50µl

#### Programme PCR HOT START

96° C 5 min (Chauffer à vide pendant 4 min puis ajouter les tubes = dénaturation 1 min)

96° C 20s	} X 35
55° C 30s	
72° C 30s + 1s à chaque cycle	
72° C 5 min	
10° C ∞	

# EXTRACTION D'ADN AU CTAB

## (CETYLTRIMETHYLAMMONIUM BROMIDE)

### (REINEKE ET AL. 1998 ET MOELLER ET AL. 1992)

#### 1. RECUPERATION DES ŒUFS DE NEMATODES

- 200 kystes sont réhydratés une nuit dans des tubes eppendorf contenant de l'eau stérile, à 4°C
- Ouvrir les kystes avec 2 pointes dans un verre de montre
- Placer le tout dans un bécher et agiter avec un agitateur magnétique de manière à séparer les débris de kystes des œufs
- Filtrer sur un tamis de 250 µm placé sur un petit entonnoir en verre, et récupérer dans des tubes à hémolyse
- Laisser sédimenter 1h à température ambiante ou une nuit à 4°C
- Retirer le surnageant à la pipette pasteur
- Remettre en suspension le culot et répartir dans 2 tubes eppendorf de 2ml (bleus)
- Centrifuger 2 min à 6500rpm et éliminer le surnageant
- Ajouter de l'eau distillée stérile afin de remplir le tube. Remettre le culot en suspension.
- Centrifuger 2 min à 6500 rpm
- Éliminer un maximum d'eau

#### 2. EXTRACTION D'ADN

- Plonger les tubes dans de l'azote liquide et placer aussitôt dans une étuve à 65°C (four à hybridation) jusqu'à décongélation
- Les culots sont remis en suspension en frottant les tubes contre un portoir.
- L'opération de congélation/décongélation est répétée une fois
- Centrifuger 5 min à 6500 rpm
- Congeler les tubes minimum 30 min à -20°C
- Ecraser les œufs congelés avec un piston
- Ajouter 500µl de tampon de lyse (0,1M Tris pH 8, 10mM EDTA, 2% SDS) et 20µl de protéinase K 10mg/ml
- Vortexer et incuber 1h à 65°C (vortexer de nouveau après 30 min d'incubation)
- Centrifuger 3 min à 4500rpm
- Récupérer le surnageant en transférant de tube à tube
- Ajouter au surnageant 140µl de NaCl 5M et 64µl de CTAB 10% (soit 1/10<sup>ème</sup> du volume)
- Agiter et incuber 10 min à 65°C
- Réaliser une extraction phénol/chloroforme avec 700µl de ce mélange  
*(Remarque : Le phénol doit impérativement être à pH 8 pour les extractions d'ADN. Le mélange doit être récent < 1mois)*
- Agiter par inclinaison pendant 20 min (table agitante avec portoir à la verticale).

- Centrifuger 10 min à 4500rpm.
- Récupérer le surnageant avec un cône à pointe coupée.
- Ajouter 200µl d'éther
- Centrifuger 2 min à 6000rpm
- Eliminer les 200µl d'éther
- Ajouter 225µl de NH<sub>4</sub>Ac 5M
- Agiter et laisser une nuit à 4°C
  
- Centrifuger 20min à 13000rpm
- Récupérer le surnageant et ajouter ¼ de volume de PEG 6000 30% (225µl) et un volume d'isopropanol 100% froid (900µl) (-20°C)
- Homogénéiser et laisser 45 min à -70°C
- Centrifuger 5 min à 11000rpm à 4°C
- Laver le culot avec 1mL d'éthanol 70% froid (-20°C)
- Centrifuger 5 min à 11000rpm à 4°C
- Eliminer le surnageant
- Sécher le culot (cloche à vide, environ 15 min)
- Reprendre le culot dans 10µl d'eau UP stérile
- Ajouter 1µl de RNase A (500µg/mL)
- Laisser réhydrater 2h à 37°C ou 15 min à 37°C + une nuit à 4°C

# EXTRACTION D'ARNs TOTAUX A PARTIR D'UN JUVENILE DE NEMATODE A KYSTE

## 1. EXTRACTION DES ARNS TOTAUX

- La paillasse est traitée à la RNase Zap, ainsi que tous les ustensiles utilisés
- Déposer un nématode sur une lame de verre, dans une goutte d'eau RNase free
- Attendre que la goutte sèche et broyer le nématode avec la pipette boule. Récupérer immédiatement le broyat avec un mix : 13,5µl d'eau RNase free + 1µl d'amorce dT25 10µM + 1µl RNasine
- Tremper la pipette boule dans le tube contenant le mix
- Placer sur la glace
- Placer 10 min à 70°C
- Ajouter : 2,5µl tampon 10X, 2,5µl MgCl<sub>2</sub> 25mM, 1µl dNTPs 10mM chaque et 2,5µl DTT 0,1M
- Placer 1 min à 42°C
- Ajouter 1µl de Superscript II ou III
- Placer 50 min à 42°C
- Puis 15 min à 70°C
- Conserver à -20°C

## 2. AMPLIFICATION PAR PCR

Tampon PCR 10X	5µl
MgCl <sub>2</sub> 25mM	3µl
dNTPs 10mM chaque	1µl
Amorce 3' 10µM	2µl
Amorce 5' 10µM	2µl
Taq 5U/µl	0,5µl
Eau UP stérile	31,5µl
cDNA	5µl

### Programme :

96°C 5min  
 96°C 20 sec  
 Tm 30 sec  
 72°C 30 sec  
 72°C 5min  
 10°C ∞

} x 35

## EXTRACTION D'ADN GENOMIQUE A PARTIR D'UN JUVENILE DE NEMATODE A KYSTE

- Une goutte d'eau ultra pure est déposée sur la lame de verre. Un individu y est déposé.
- Lorsque la goutte est évaporée, placer la lamelle sur la larve et écraser par pression sur la lamelle avec la pince.
- La lamelle est ensuite retournée et nettoyée avec du tampon EX4 qui est immédiatement récupéré et placé dans un tube eppendorf de 200 $\mu$ l, placé sur la glace. La lame est également lavée avec du tampon EX4.

Au final, l'ADN est récupéré dans 10 $\mu$ l de tampon EX4.

- Congeler le culot à -80°C, pendant 15 minutes.
- Décongeler le culot et placer 1h à 60°C, puis 15 min à 95°C (Thermocycleur Hybaid).
- Un traitement à la RNase peut être réalisé : 1 $\mu$ l de RNase A (500 $\mu$ g/ $\mu$ l). Placer à 37°C, durant 30 min (Bain marie).
- Conservation à -20°C.

### Composition du tampon d'extraction :

Tris 100mM pH 8,8 (10%)	10mL
EDTA (1mM)	373mg
Monidet P40-Tergitol (1%)	1mL
Protéinase K (100 $\mu$ g/mL de tampon)	100 $\mu$ l
Eau ultra pure qsp	100mL



# SOUTHERN BLOT

## 1. DIGESTION DES ADNS

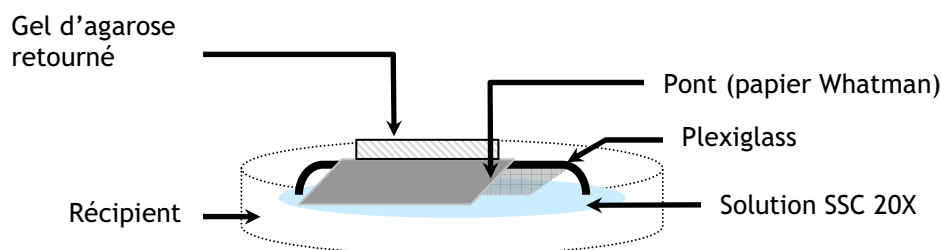
- Mix :

Tampon de l'enzyme 10X	2µl
Enzyme (10U/µl)	1µl
ADN	5µg
H <sub>2</sub> O UP stérile	qsp 20µl

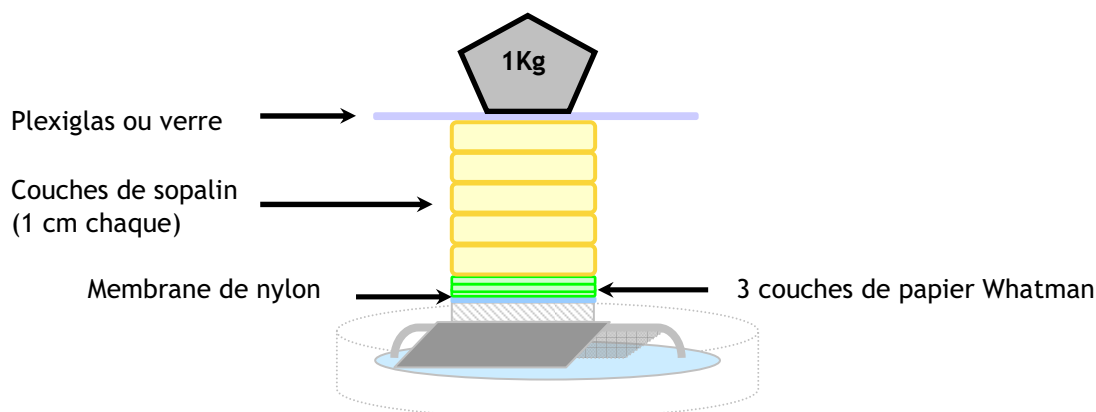
- Placer 4h à 37° C
- Déposer la totalité sur agarose 1,5%, en grande plaque  
Utilisation du 100pb comme marqueur de taille
- Migration 30V (soit 20 à 30mA) sur la nuit  
*Attention aux paramètres de temps sur le générateur !*

## 2. TRANSFERT SUR MEMBRANE DE NYLON

- Observer le gel sous UV pour contrôler la migration
- Remettre à migrer si besoin avec un voltage plus élevé
- Prendre une photo du gel
- Réduire le gel : retirer les bandes d'agarose sans ADN (UV)
- Marquer les puits et les bandes du marqueur de taille à l'encre, avec une aiguille
- Préparation du gel :
  - ♦ Lavage 15 min avec du HCl 0,25N (recouvrir le gel et agiter de temps en temps) (=cassure ADN >10Kb)
  - ♦ Rinçage léger à H<sub>2</sub>O
  - ♦ 2 lavages de 15 min avec la solution de dénaturation (NaOH 0,5M + NaCl 1,5M)
  - ♦ Rinçage léger à H<sub>2</sub>O
  - 2 lavages de 30 min avec la solution de neutralisation (NaCl 1,5M + Tris HCl 0,5M pH 7,4 + Na<sub>2</sub>EDTA 0,001M)
- Découper la membrane de nylon (Hybridation N, chargée positivement) à l'exacte dimension du gel. Reproduire l'encoche.
- Découper 3 feuilles de papier Whatman à la dimension du gel + 1 feuille 1 cm plus large et beaucoup plus longue pour faire le pont (voir montage)

- Montage :

- ♦ Dans un bac, placer un moule retourné en plexiglas préalablement lavé à l'eau puis à l'éthanol.
- ♦ Laver également le bac à l'eau puis à l'éthanol.
- ♦ Mettre en place le pont avec le papier Whatman
- ♦ Ajouter dans le bac du SSC 20X et imbiber le pont avec cette même solution. *Attention à ne pas faire de bulles entre le plexi et le papier. Pour cela, faire rouler une pipette de 10 mL dessus.*
- ♦ Placer le gel **RETOURNE** sur le pont
- ♦ Au dessus du gel :



- ♦ Placer la membrane de nylon au dessus du gel, après l'avoir trempée dans de l'eau ultra pure puis 2 min dans du SSC 2X. *Attention à bien manipuler la membrane avec des pinces. Faire rouler la pipette dessus pour éviter les bulles.*
- ♦ Placer les 3 feuilles de papier Whatman, une à une, en faisant rouler la pipette à chaque fois. Les 3 feuilles sont placées extemporanément dans du SSC 2X.
- ♦ Placer les couches de sopalin (5) (préalablement préparées par enroulement des feuilles Whatman). *Ces couches doivent également avoir la même taille que le gel.*
- ♦ Poser une plaque de plexi ou de verre et ajouter un poids d'environ 1kg.
- ♦ Avant la nuit, retirer les couches de sopalin mouillées et les remplacer, SAUF celle qui est au contact du Whatman.
- ♦ Laisser une nuit.

- Après la nuit, retirer les différentes couches, sauf la membrane de nylon.
- Refaire le marquage des poids moléculaires et des puits à l'encre (percer la membrane).
- Récupérer la membrane avec des pinces et la placer dans du SSC 2X, 30 min. Renouveler le lavage 1 fois.
- Vérifier le gel sous UV.
- Egoutter et sécher la membrane sur du papier Whatman.
- Dans du Whatman plié en deux, placer la membrane et laisser cuire 2h à 80°C.
- Récupérer la membrane et la placer dans un nouveau Whatman plié en 2. Placer dans un sachet hermétique et conserver à température ambiante, à l'abri de la lumière.

### 3. HYBRIDATION

#### a. Préparation des sondes

##### Mix :

Insert	15ng
Amorce Fwd 10µM	1µl
Amorce Rev 10 µM	1µl
Tampon Taq 10X	2,5µl
dNTPs 2,5mM chaque	2µl
MgCl <sub>2</sub> 25mM	1,5µl
Taq	0,2µl
Eau UP stérile	qsp 25µl

##### Programme :

94°C 4 min  
 94°C 30 sec  
 Tm 30 sec  
 72°C 1 min  
 72°C 4 min  
 10°C Infini

} X 34

Purification des sondes : kit sigma de purification de produits PCR

#### b. Préhybridation

- La membrane est placée dans un tube à hybridation, enroulée. *Possibilité de mettre jusqu'à 6 à 7 membranes en même temps.*
- Ajout dans le tube de : 15 mL de tampon d'hybridation + 150µl de sperme de saumon
- Le tube est placé dans le four à hybridation à 65°C durant 30 min minimum sous agitation. (*Attention à bien équilibrer le rotor du four*)

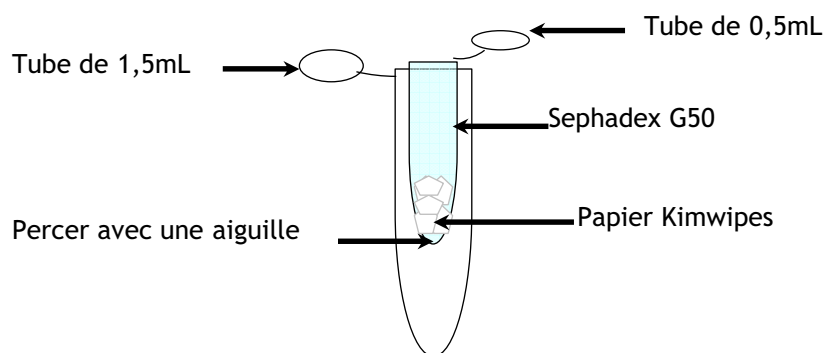
### c. Préparation de la sonde

- Préparer 45µl de produit PCR purifié à 2ng/µl final, dans du TE (Tris 10mM/EDTA 1mM).
- Dénaturer : 2 min à 100°C et placer sur la glace (1min).
- Incorporer le  $^{32}\text{P}$  selon le manuel du kit : Ready to Go. DNA Labelling Beads (-dCTP) (Amersham pharmacia biotech).

Mix : DNA dénaturé 45µl

[ $\alpha$ - $^{32}\text{P}$ ]dCTP (3000Ci/mmol) 5µl

- Mélanger en pipetant ou en vortexant
- Eliminer les éventuelles bulles par centrifugation
- Incuber à 37°C 20 à 30 min
- Purifier la sonde sur Sephadex G50 :



- Une fois la matrice sephadex déposée, centrifuger. *Attention à ne pas décoller la matrice des parois du tube.*
- Déposer la solution contenant la sonde au centre de la matrice et centrifuger.
- La sonde peut alors être dénaturée : 2 min à 100°C puis 1min sur glace
- Utiliser des tubes avec bouchons à vis de manière à ce que le bouchon ne saute pas lors du passage à 100°C (idem dans l'étape précédente de dénaturation). Ne pas percer le bouchon, la radioactivité s'évapore !

### d. Hybridation

- Après avoir retiré la solution de pré hybridation, ajouter la sonde, 2,5mL de tampon d'hybridation et du sperme de saumon dans le tube avec la membrane.
- Laisser en hybridation, sous agitation à 65°C, tout la nuit.

## 4. AUTORADIOGRAPHIE

### a. Lavages

- Le mélange d'incubation radioactif est récupéré et conservé dans un flacon en verre.
- Lavages :
  - ♦ SSC 2X/0,1% SDS (1 lavage rapide, 2 lavages de 20 min sous agitation à 65°C)

- ♦ SSC 1X/0,1%SDS (1 lavage de 20 min)
  - La membrane est emballée dans du film alimentaire
  - Placer la membrane dans une cassette
  - En lumière rouge :
- Sur l'écran amplificateur de la cassette, placer le film photo (pure) sur lequel on a reproduit l'encoche de la membrane. Par dessus le film, placer la membrane. Refermer la cassette.
- Garder la cassette à température ambiante ou placer à -80°C pour une exposition plus longue.

#### **b. Révélation**

- 5 min dans le révélateur (1/9) en agitant
- Passage rapide sous eau du robinet
- 5min dans le fixateur (1/4) en agitant
- Passage rapide sous eau du robinet
- Laisser sécher

### **5. DESHYBRIDATION**

Réaliser les lavages sous agitation à 45°C, deux fois dans l'ordre suivant :

- NaOH 0,4M 30 min
- 0,1X SSC/0,1%SDS/Tris 0,2M pH 7,5 30min

#### **SOLUTIONS :**

- **SSC 20X** : NaCl 3M, Tri citrate de sodium 0,3M
- **Solution de dénaturation** : NaOH 5M, NaCl 1,5M
- **Denhart's 50X** : Ficoll 0,01%, Polyvinyl pyrrolidone 10µg/ml, BSA 10µg/ml
- **Tampon d'hybridation (sondes radioactives)** : SSC 6X, Denhart's 5X, SDS 0,5%
- **Tampon de déshybridation** : SSC 0,1X, SDS 0,1%, Tris 0,2M pH 7,4,
- **Solution de neutralisation** : NaCl 1,5M, TrisHCl 0,5M pH 7,4, Na<sub>2</sub>EDTA 0,001M

# HYBRIDATION *IN SITU*

## 1. PREPARATION DE LA SONDE

### *a. Amplification des sondes :*

Mix PCR : 1 à 15 ng d'insert  
 2µl d'amorce forward 10µM  
 2µl d'amorce reverse 10µM  
 5 µl tampon Taq 10X (Promega)  
 3 µl MgCl<sub>2</sub> 25mM  
 4 µl dNTP 2,5mM chaque  
 0.4 µl Taq (Promega)  
 qsp 50µl H<sub>2</sub>O RNase free

Programme : 94° C 4 min  
 94° C 30 s  
 56° C à 58° C 30s  
 72° C 1 min  
 72° C 4 min  
 10° C

} 34 à 40 cycles

La taille des fragments est contrôlée par électrophorèse en gel d'agarose 1%.

### *b. Marquage des sondes à la Digoxygénine :*

Mix PCR : Produit PCR  
 10µl d'amorce Forward ou Reverse 10µM  
 4µl Tampon Taq 10 X  
 2,4 µl MgCl<sub>2</sub> 25mM  
 3µl DIG-dUTP/dNTP mix (dangereux : sous hotte)  
 0.4 µl Taq  
 Eau ultra pure : qsp 40 µl

Programme : 94° C 4 min  
 94° C 30 s  
 Tm 60s  
 72° C 90 s  
 72° C 4 min  
 10° C

} 34 cycles

Purification des sondes : kit de purification Sigma, GenElute™ PCR Clean-Up

Les sondes sont conservées au congélateur.

## 2. PURIFICATION ET FIXATION DES NEMATODES

### a. Purification

Les nématodes sont répartis dans 2 tubes, puis centrifugés à 12000 rpm, 2 min (pour culoter la totalité des nématodes) (*N.B. : utiliser des tubes en verre sinon les nématodes se collent à la paroi*). Retirer le surnageant au maximum (pipette puis pipette pasteur)

### b. Fixation

*N.B.: Toutes les étapes doivent être réalisées sous la Sorbonne*

Resuspendre le culot dans 5 ml de tampon de fixation (pour chaque tube) puis transférer dans un bécher. Rincer les tubes avec 2,5 ml de tampon de fixation (pour chaque tube) et verser dans le bécher. Mettre un parafilm et placer à 4°C pendant 18 h, puis 4 heures à température ambiante. Le tampon de fixation doit être conservé à -20°C.

## 3. SECTION ET PERMEABILISATION DES NEMATODES

*N.B.: Toutes les étapes doivent être réalisées sous la Sorbonne*

- Culoter les nématodes dans 2 tubes en verre 2x2 min, 4100 rpm. Eliminer un maximum de surnageant
- Étaler une goutte de tampon de fixation avec les J2 en ruban sur une plaque de verre traitée à la RNase ZAP. Découper les J2 en 2 à 4 sections avec une lame de rasoir fixée sur une pompe d'aquarium (vibrante). Vérifier les sections à la loupe.
- Récupérer les sections dans un bécher en déposant quelques gouttes de tampon M9 à l'aide de pipettes pasteur (préalablement passées à 150°C pendant 3h).
- Culoter les sections dans 2 tubes eppendorf de 1,5 ml par centrifugations successives (12000 rpm, 2-3 min.).
- Laver 2 fois avec 1 ml de tampon M9 (centrifuger 1 min à 12000 rpm)
- Incuber les sections dans 500µl de protéinase K (0,5mg/ml) (26,4 µl de protéinase K (roche) 19mg/ml dans 1ml de tampon M9) dans du tampon M9, pendant 30 min à 22°C sous agitation (four à hybridation).
- Centrifuger 1 min à 12000 rpm puis retirer le surnageant.
- Laver avec 1 ml de tampon M9.
- Centrifuger 1 min à 12000 rpm. Eliminer le tampon M9.
- Congeler sur « dry ice » (glace bien tassée et passée à -80°C)
- Resuspendre le culot dans 1 ml de méthanol - 20°C, 30s à -20 °C
- Centrifuger 1 min à 12000rpm. Eliminer le méthanol.
- Resuspendre le culot dans 1 ml d'acétone à -20°C et placer 1 min à - 20 °C.
- Réhydrater les sections dans 1 ml d'eau RNase free pendant 5 min.

#### 4. HYBRIDATION

- Centrifuger les sections de J2 1 min à 12000 rpm et éliminer le surnageant.
- Laver le culot avec 500 µl de tampon d'hybridation. Centrifuger 1 min à 12000 rpm.
- Resuspendre dans 500 µl de tampon d'hybridation.
- Pré hybrider à 40 ou 45°C pendant 15 min sous agitation (four à hybridation).
- En même temps, dénaturer les sondes 10 min à 100°C, puis les placer 3 min sur la glace.
- Centrifuger 1 min à 12000 rpm.
- Resuspendre les nématodes dans un volume V de tampon d'hybridation ( $V = 100 \times \text{nombre de sondes testées}$ ).
- Aliquoter les nématodes resuspendus dans des tubes eppendorf neufs (100µl / tube / sonde testée).
- Ajouter un maximum de volume contenant la sonde soit environ 40 µl.
- Ajouter 110 µl de tampon d'hybridation pour obtenir un volume final de 250 µl.
- Hybrider sur la nuit à 40°C ou 45°C.
- Centrifuger le lendemain pendant 1 min à 12000 rpm, puis réaliser les lavages.
- 3 lavages 15 min. dans 500 µl SSC 4 x à 40°C ou 45°C sous agitation.
- 3 lavages 20 min. dans 500 µl SSC 0,1X / 0,1 % SDS à 40°C ou 45°C sous agitation.

#### 5. COLORATION

- Laver les sections dans 500 µl tampon acide maléique pendant 1 min.
- Incuber les nématodes dans 500 µl Boehringer blocking reagent 1% dans du tampon acide maléique, 30 min sous agitation lente.
- Incuber 2 heures avec 500 µl d'anticorps anti-Digoxygénine couplés à la phosphatase alcaline dilués au 1/1000 dans du Boehringer blocking reagent 1 % (lui même dilué dans du tampon acide maléique). Sous agitation lente.
- Laver 3 fois 15 min dans 500 µl tampon acide maléique + Tween 20 0.01 %
- Laver ensuite dans 500 µl de tampon de détection contenant de la phosphatase alcaline.
- Incuber les sections à 4°C sur la nuit (16 heures) dans 1 ml : tampon de détection de la phosphatase alcaline (178 µl de NBT/BCIP dans 10 ml tampon de détection).

#### 6. OBSERVATION DES LAMES

- Etaler une goutte contenant les sections de nématodes en cours de coloration sur une lame de verre et recouvrir d'une lamelle.
- Observer la coloration.
- Si le marquage est correct, la coloration est arrêtée : centrifugation 1min à 12000rpm, élimination du surnageant et ajout de 1mL d'eau distillée. Sinon, la coloration est poursuivie.



## TAMPONS POUR L'HYBRIDATION *IN SITU*

- **Tampon M9** : Pour 100 ml (RNase free) :  $\text{Na}_2\text{HPO}_4$  (0,6g),  $\text{KH}_2\text{PO}_4$  (0,3g), **NaCl (0,5g)** ,  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  (0,025g). Ajuster à **pH 7** avec du HCl.

- **Tampon de fixation** : Paraformaldéhyde dans du tampon M9

- Ajouter 1g de paraformaldehyde à 50 ml de tampon M9 (sous la sorbonne), dans un tube Falcon.

- Chauffer à 100°C, environ 20 minutes.

- Laisser refroidir à température ambiante.

- Ajuster le volume à 50ml avec du tampon M9 puis aliquoter.

- Stocker à -20°C plusieurs semaines

- **Protéinase K** :

Diluer 19mg/mL de solution stock dans du Tris-HCl 10 mM, pH 7.5 (Stockage à 4°C).

- **Tampon d'hybridation** : (Stocker à -20°C)

Pour 40 mL :

Formamide désionisée	20mL
Blocking reagent 10%	4mL
SDS 20%	4 ml
Denhardt's 100X	0,4 ml
EDTA 0,1 M, pH 8	0,4 ml
Sperme de saumon (10mg/mL)	0,8 ml
tRNA de levure (500U/mL)	0,25 ml
Eau RNase free	2,15 ml
SSC20X	8ml

- **Formamide désionisée** :

- Ajouter une cuillerée à soupe de résine d'Amberlite MD 150 (Sigma) à 100 ml de formamide.

- Agiter pendant 3 heures et filtrer sur papier filtre (N.B.: *la résine ne se dissout pas*). (Stocker à température ambiante)

- **Tampon acide maléique** : 0,1 M Acide maléique, 0,15 M NaCl. Ajuster le pH à 7,5 avec du NaOH et autoclaver (Stockage à -20°C).

- **Tampon de détection de la phosphatase alcaline** : 0,1 M Tris HCl, 0,1 M NaCl, 50mM  $\text{MgCl}_2$ , 6  $\text{H}_2\text{O}$ . Dissoudre dans de l'eau RNase free et ajuster le pH à 9,5. Autoclaver

- **Blocking Reagent 1%** : 1g de Blocking Reagent dans 10mL de tampon acide maléique Autoclaver et stocker à -20°C.

# AMPLIFICATION DES GENES PECTATE LYASE, CATHEPSINE L, FACTEUR D'ELONGATION 1 $\alpha$ ET RBP-1

## 1. CONDITIONS D'AMPLIFICATION DU GENE CATHEPSINE L

Les amplifications ont été réalisées sur des extractions d'ADN sur un kyste comme décrit précédemment. Pour chaque amplification, 5 $\mu$ l d'ADN ont été utilisés.

*N.B.: Les amplifications sont préférentiellement réalisées sur des extractions d'ADN fraîchement extraites.*

### Mix PCR :

Tampon PCR 5X	5 $\mu$ l
MgCl <sub>2</sub> 25mM	3 $\mu$ l GT/GR/H et 1,5 $\mu$ l GP/GM
dNTPs 10mM chaque	0,5 $\mu$ l
Amorce 5' 10 $\mu$ M	1 $\mu$ l
Amorce 3' 10 $\mu$ M	1 $\mu$ l
GoTaqFlexi (Promega) 5U/ $\mu$ l	0,25 $\mu$ l
DNA	5 $\mu$ l
Eau UP stérile	qsp 25 $\mu$ l

GT : *G. tabacum*, GR : *G. rostociensis*, GP: *G. pallida*, GM : *G. "mexicana"*, H : *Heterodera*.

Les amorces CathepFwd5 et CathepRev5 ont été utilisées pour amplifier le fragment cathepsine L chez les nématodes du genre *Globodera*. Pour les amplifications sur les espèces du genre *Heterodera*, les amorces utilisées sont : CathepFwd et CathepRev5.

### Programme :

95° C 1min	} x 35
95° C 30 sec	
54° C 30 sec	
72° C 3 min	
72° C 5min	
10° C $\infty$	

## 2. CONDITIONS D'AMPLIFICATION DU GENE PECTATE LYASE

De même que pour la cathepsine, les amplifications ont été réalisées sur 5µl d'ADN extrait d'un kyste.

### Mix PCR :

Tampon PCR 10X	2,5µl
MgCl <sub>2</sub> 25mM	2µl
dNTPs 10mM chaque	0,5µl
Amorce 5' 10µM	1,25µl
Amorce 3' 10µM	1,25µl
Taq (Qbiogen) 15U/µl	0,08µl
DNA	5µl
Eau UP stérile	qsp 25µl

Les amorces PLFwd.2 et PLRev ont été utilisées pour amplifier le fragment pectate lyase chez les nématodes du genre *Globodera*. Pour les amplifications sur les espèces du genre *Heterodera*, les amorces utilisées sont : PLFwd.2 et PLRev.2.

Les mêmes conditions PCR ont été utilisées pour amplifier les transcrits. Pour les espèces du genre *Globodera*, le couple d'amorces PLFwd4 and PLRev3 a été utilisé. Pour les nématodes du genre *Heterodera*, les amorces PLFwd5 et PLRev4.

### Programme :

96° C 1min	} x 35
96° C 20 sec	
55° C 20 sec	
72° C 1,5 min	
72° C 5min	
10° C ∞	

### 3. CONDITIONS D'AMPLIFICATION DU GENE FACTEUR D'ELONGATION 1 $\alpha$

Les amplifications ont été réalisées sur 5 $\mu$ l d'ADN extrait d'une larve.

#### Mix PCR :

Tampon PCR Qbiogen 10X (avec MgCl <sub>2</sub> )	2,5 $\mu$ l
dNTPs 25mM chaque	0,1 $\mu$ l
Amorce 5' 10 $\mu$ M	2,5 $\mu$ l
Amorce 3' 10 $\mu$ M	2,5 $\mu$ l
Taq (Qbiogen) 15U/ $\mu$ l	0,075 $\mu$ l
ADN	5 $\mu$ l
Eau UP stérile	qsp 25 $\mu$ l

Les amorces EF1Fwd et EF1Rev ont été utilisées pour amplifier le fragment du gène EF chez les Heteroderidae.

#### Programme :

94° C 4min  
94° C 30 sec  
58° C 30 sec  
72° C 1 min  
72° C 5min  
10° C  $\infty$

} x 30

#### 4. CONDITIONS D'AMPLIFICATION DU GENE *RBP-1*

Les amplifications du gène *rbp-1* ont été réalisées sur des extractions d'ADN génomique au CTAB (voir protocole précédemment décrit).

##### Mix PCR :

Tampon PCR 10X	2,5µl
MgCl <sub>2</sub> 25mM	1,5µl
dNTPs 10mM chaque	0,5µl
5'IC5.2 II 10µM	1,25µl
3'IC5.3 10µM	1,25µl
Taq (Promega) 5U/µl	0,25µl
DNA	~50 ng
Eau UP stérile	qsp 25µl

##### Programme :

96 °C 1min	
96 °C 20 sec	} x 35
59 °C 20 sec	
72 °C 3 min	
72 °C 5min	
10 °C ∞	

Les conditions d'amplification du transcrit *rbp-1* sont les suivantes :

##### Mix PCR :

Tampon PCR 5X	5µl
MgCl <sub>2</sub> 25mM	2µl
dNTPs 10mM chaque	0,5µl
5'IC5 10µM	1µl
3'IC5.3 10µM	1µl
Taq (Promega) 5U/µl	0,25µl
cDNA	~50 ng
Eau UP stérile	qsp 25µl

##### Programme :

96 °C 1min	
96 °C 20 sec	} x 35
57 °C 20 sec	
72 °C 30 sec + 1 sec/cycle	
72 °C 5min	
10 °C ∞	

## ANALYSE DES PRESSIONS DE SELECTION

Pour analyser les pressions de sélection subies par les gènes étudiés, différents logiciels ont été utilisés.

- Le logiciel MEGA (Kumar, Tamura et Nei, 2004) peut calculer le nombre de substitutions synonymes et non synonymes ainsi que le nombre potentiel de sites synonymes et non synonymes. Il permet ainsi d'obtenir des dN et dS qui correspondent à des p-distances corrigées en prenant en compte les substitutions multiples sur un même site (correction Jukes-Cantor). Ce logiciel ne peut calculer des ratios dN/dS que sur la totalité de la séquence. Ceci ne permet pas d'avoir une réelle idée des pressions de sélection puisqu'il a été montré à de nombreuses reprises que ces pressions ne concernent que des portions d'un gène.

- Nous avons alors cherché à déterminer les ratios de substitutions synonymes vs non synonymes le long des séquences par sliding window. Le logiciel DnaSP (Rozas and Rozas, 1999), dont l'accès est libre, a été utilisé comme décrit par Haag *et al.* (2005) pour détecter de la sélection positive le long de la séquence des gènes. Nous avons utilisé l'option "polymorphisme et divergence" en analysant la séquence entière par fenêtre de 10 nucléotides et en faisant glisser la fenêtre de 1 nucléotide à chaque fois pour obtenir une résolution relativement fine. Les gaps ont été exclus de l'analyse et les ratios Ka/Ks (= dN/dS) ont été calculés. Cette méthode nous a permis de mettre en évidence des ratios nettement plus élevés pour les gènes du pouvoir pathogène en comparaison avec le gène de ménage ainsi que des zones potentiellement soumises à des pressions de sélection diversificatrices. Si ce type d'analyse donne des indications sur les ratios Ka/Ks, elle ne permet ni d'utiliser plusieurs modèles de calculs, ni d'accorder un poids aux résultats en fonction de la vraisemblance. Il faut pour ce faire les analyses avec un autre logiciel. Le plus utilisé et celui qui semble le plus pertinent est PAML.

- Les différents modèles du module codeml (Goldman et Yang, 1994) du logiciel PAML (Yang, 1997) ont été utilisés pour calculer les ratios dN/dS le long des gènes cathepsine L et pectate lyase en calculant également un score de vraisemblance des résultats obtenus pour chaque modèle. Ce module permet de détecter la sélection positive le long de la séquence, sur les différents sites, mais également dans différentes lignées, sur les branches d'un arbre phylogénétique. Les paramètres qui ont été utilisés sont ceux recommandés par Yang. Les différents paramètres sont indiqués sur la figure 1. Les paramètres principaux sont détaillés.

Le modèle "sites" permet de calculer le ratio dN/dS sur les différents sites d'un gène, sur le jeu de données complet. Ce modèle comprend différents sous modèles de substitutions de codons M1 à M13. Les modèles M1, M2, M7 et M8 sont les modèles classiquement testés. Nous avons présentés dans cette étude les résultats du modèle M8 qui définissent un maximum de sites sous sélection positive. Les paramètres utilisés sont les suivants : runmode = 0, model = 0, NSsites = 1 à 8.

Le modèle "branche" autorise les variations des ratios dN/dS sur les différentes branches de l'arbre. Les paramètres utilisés pour ce modèles sont les suivants : runmode = 0, model = 1, NSsites = 0.

L'analyse combinant les deux modèles a été menée en intégrant les paramètres suivants : runmode = 0, model = 2 et NSsites = 2.

```

seqfile = lysin.nuc
treefile = lysin.trees

outfile = mlc          * main result file name
noisy = 9             * 0,1,2,3,9: how much rubbish on the screen
verbose = 1          * 1: detailed output, 0: concise output
runmode = 0          * 0: user tree; 1: semi-automatic; 2: automatic
                    * 3: StepwiseAddition; (4,5):PerturbationNNI; -2: pairwise

seqtype = 1          * 1:codons; 2:AAs; 3:codons-->AAs
CodonFreq = 2        * 0:1/61 each, 1:F1X4, 2:F3X4, 3:codon table
clock = 0            * 0: no clock, unrooted tree, 1: clock, rooted tree

model = 0             * models for codons:
                    * 0:one, 1:b, 2:2 or more dN/dS ratios for branches

NSsites = 8          * 0:one w;1:neutral;2:positive; 3:discrete;4:freqs;
                    * 5:gamma;6:2gamma;7:beta;8:beta&w;9:beta&gamma;
                    * 10:beta&l+gamma; 11:beta&l>normal; 12:0&2normal; 13:3normal
icode = 0            * 0:standard genetic code; 1:mammalian mt; 2-10:see below

fix_kappa = 0        * 1: kappa fixed, 0: kappa to be estimated
kappa = 1.6          * initial or fixed kappa
fix_omega = 0        * 1: omega or omega_1 fixed, 0: estimate
omega = .8           * initial or fixed omega, for codons or codon-based AAs

ncatG = 10           * # of categories in dG of NSsites models
getSE = 0            * 0: don't want them, 1: want S.E.s of estimates
RateAncestor = 0     * (0,1,2): rates (alpha>0) or ancestral states (1 or 2)

Small_Diff = 3e-7
cleandata = 0        * remove sites with ambiguity data (1:yes, 0:no)?
method = 0           * 0: simultaneous; 1: one branch at a time

```

Utilisé en runmode = 1 pour prendre en compte l'arbres de gène que nous avons généré pour chaque candidat.

Modél 0 : même ratio dN/dS sur toutes les branches.  
Modél 1 : un ratio pour chaque branche.  
Modél 2 : Nombre de ratios arbitraire.

Différents modèles autorisant les variations du ratio dN/dS sur les sites.  
M1 : modèle neutre  
M2 : sélection positive  
M7 et M8 : modèles prenant en compte les distributions des ratios.

Figure 1 : Structure du programme codeml du logiciel PAML.





*Annexe 3 :*  
*Alignements de*  
*séquences*

# ALIGNEMENT DES SEQUENCES GENOMIQUES PECTATE LYASE (EXONS ET INTRONS)

Name: PL-GPE1-HL7/8	Len: 1322	Check: 1474	Weight: 1.00
Name: PL-GPE3-MG25	Len: 1322	Check: 1715	Weight: 1.00
Name: PL-GPE3-MG13	Len: 1322	Check: 1871	Weight: 1.00
Name: PL-GPE3-MG21	Len: 1322	Check: 3160	Weight: 1.00
Name: PL-GPE3-HL11/1	Len: 1322	Check: 4131	Weight: 1.00
Name: PL-GPE3-MG24	Len: 1322	Check: 1908	Weight: 1.00
Name: PL-GPE1-IL1	Len: 1322	Check: 3806	Weight: 1.00
Name: PL-GPE2-IL2	Len: 1322	Check: 125	Weight: 1.00
Name: PL-GPS2-IL6	Len: 1322	Check: 9861	Weight: 1.00
Name: PL-GPE5-IL5	Len: 1322	Check: 8524	Weight: 1.00
Name: PL-GPS6-HX11	Len: 1322	Check: 496	Weight: 1.00
Name: PL-GPS6-IL10	Len: 1322	Check: 5311	Weight: 1.00
Name: PL-GM1-HL17/18	Len: 1322	Check: 5818	Weight: 1.00
Name: PL-GM1-IL7	Len: 1322	Check: 5166	Weight: 1.00
Name: PL-GPS5-IL9	Len: 1322	Check: 3766	Weight: 1.00
Name: PL-GPE2-HX2	Len: 1322	Check: 6331	Weight: 1.00
Name: PL-GPE3-MG11	Len: 1322	Check: 5316	Weight: 1.00
Name: PL-GPE3-MG12	Len: 1322	Check: 5316	Weight: 1.00
Name: PL-GPE3-MG14	Len: 1322	Check: 5316	Weight: 1.00
Name: PL-GPE4-HX4	Len: 1322	Check: 3509	Weight: 1.00
Name: PL-GPE3-IL3	Len: 1322	Check: 557	Weight: 1.00
Name: PL-GPE4-IL4	Len: 1322	Check: 4143	Weight: 1.00
Name: PL-GPS3-IL7	Len: 1322	Check: 7695	Weight: 1.00
Name: PL-GPS4-HX9	Len: 1322	Check: 1684	Weight: 1.00
Name: PL-GPS5-HX10	Len: 1322	Check: 8853	Weight: 1.00
Name: PL-GPS7-HX12	Len: 1322	Check: 2978	Weight: 1.00
Name: PL-GPS7-II21-2	Len: 1322	Check: 4224	Weight: 1.00
Name: PL-GPS8-ID3	Len: 1322	Check: 452	Weight: 1.00
Name: PL-GPS7-IJ21-2	Len: 1322	Check: 2877	Weight: 1.00
Name: PL-GM2-HX22	Len: 1322	Check: 6524	Weight: 1.00
Name: PL-GM2-II43/44	Len: 1322	Check: 4906	Weight: 1.00
Name: PL-GM2-IJ43/44	Len: 1322	Check: 3451	Weight: 1.00
Name: PL-GPS9-HX13	Len: 1322	Check: 4335	Weight: 1.00
Name: PL-GPS9-IJ25-2	Len: 1322	Check: 6108	Weight: 1.00
Name: PL-GPS10-HX14	Len: 1322	Check: 6920	Weight: 1.00
Name: PL-GPS10-II27-	Len: 1322	Check: 7061	Weight: 1.00
Name: PL-GPE5-HX5	Len: 1322	Check: 7619	Weight: 1.00
Name: PL-GPS2-HX6	Len: 1322	Check: 7380	Weight: 1.00
Name: PL-Gtv-HX25	Len: 1322	Check: 3344	Weight: 1.00
Name: PL-Gtt1-HX23	Len: 1322	Check: 3058	Weight: 1.00
Name: PL-Gtt1-IL18	Len: 1322	Check: 7055	Weight: 1.00
Name: PL-Gts-HX26	Len: 1322	Check: 7355	Weight: 1.00
Name: PL-Gts-IJ49-50	Len: 1322	Check: 6386	Weight: 1.00
Name: PL-Gts-II49-50	Len: 1322	Check: 4008	Weight: 1.00
Name: PL-Gta-MG3	Len: 1322	Check: 3459	Weight: 1.00
Name: PL-Gta-MG1	Len: 1322	Check: 3459	Weight: 1.00
Name: PL-Gta-MG2	Len: 1322	Check: 3459	Weight: 1.00
Name: PL-Gta-MG5	Len: 1322	Check: 3102	Weight: 1.00
Name: PL-Gta-MG28	Len: 1322	Check: 3102	Weight: 1.00
Name: PL-Gta-MG29	Len: 1322	Check: 3228	Weight: 1.00
Name: PL-Gta-MG30	Len: 1322	Check: 3228	Weight: 1.00
Name: PL-Gta-MG27	Len: 1322	Check: 2031	Weight: 1.00
Name: PL-Gta-II51-52	Len: 1322	Check: 3965	Weight: 1.00
Name: PL-Gta-IJ51-52	Len: 1322	Check: 9218	Weight: 1.00
Name: PL-GtaRev	Len: 1322	Check: 7997	Weight: 1.00
Name: PL-GRE1-HL21/2	Len: 1322	Check: 692	Weight: 1.00
Name: PL-GRE3-HX17	Len: 1322	Check: 5587	Weight: 1.00
Name: PL-GRS2-ID1	Len: 1322	Check: 4287	Weight: 1.00
Name: PL-GRS1-MO2	Len: 1322	Check: 2829	Weight: 1.00
Name: PL-GRS1-IL14	Len: 1322	Check: 6325	Weight: 1.00
Name: PL-GRE2-HX16	Len: 1322	Check: 7772	Weight: 1.00
Name: PL-GRE2-IL12	Len: 1322	Check: 2436	Weight: 1.00
Name: PL-GRE3-IL13	Len: 1322	Check: 7649	Weight: 1.00
Name: PL-GRE1-IL11	Len: 1322	Check: 4942	Weight: 1.00
Name: PL-GRS3-IL16	Len: 1322	Check: 9074	Weight: 1.00
Name: PL-HG	Len: 1322	Check: 8540	Weight: 1.00
Name: PL-HS41-IL21-2	Len: 1322	Check: 4020	Weight: 1.00
Name: Consensus	Len: 1322	Check: 7022	Weight: 0.00

//





Table with multiple columns containing sequence identifiers (e.g., PL-GPE1-HL), nucleotide sequences, and alignment markers. Includes a Consensus row at the bottom of each section. Two sections are visible, one starting at line 521 and another at line 651.





```

1301          1322
PL-GPE1-HL CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-HL CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE1-IL CAAAAT
PL-GPE2-IL CAAAATT
PL-GPE2-IL CAAAAT
PL-GPE5-IL CAAAATT
PL-GPE5-IL CAAAAT
PL-GPS6-HX CAAAAT
PL-GPS6-IL CAAAAT
PL-GM1-HL1 CAAAATTCCA ATTTCGGG
PL-GM1-IL1
PL-GPS5-IL CAAAATTCCA ATTTCGGG
PL-GPE2-HX CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE3-MG CAAAATTCCA ATTTCGGG
PL-GPE4-HX CAA
PL-GPE3-IL CAA
PL-GPE4-IL CAAAATTCAA ATTTCGGG
PL-GPS3-IL CAAAATT
PL-GPS4-HX CAAAATTCCG ATTTCGGG
PL-GPS5-HX CAAAATT
PL-GPS7-HX C
PL-GPS7-IL CAAAAT
PL-GPS8-ID CAAAATTTC
PL-GPS7-IJ CAAAATTCAA ATTTCGGG AA
PL-GM2-HX2 CAAAATTCCA ATTTCGGG A
PL-GM2-I14 CAAAATTCAA ATTTCGGG
PL-GM2-I14 CAAAAT
PL-GPS9-HX CAAAATTCCA ATTTCGGG
PL-GPS9-IJ CAAAAT
PL-GPS10-H CA
PL-GPS10-I CA
PL-GPE5-HX CAAAATTCCA ATTT
PL-GPS2-HX
PL-GU-HX2 CAAAAT
PL-Gt1-HX ACGGAATTCA ATTTCGGG
PL-Gt1-IL ACAAAAT
PL-Gt8-HX2 CAAAATT
PL-Gt8-I14 CAAAAT
PL-Gt8-I14 CAAAAT
PL-Gt8-MG3 CAAAATTCCA ATTTCGGG
PL-Gt8-MG1 CAAAATTCCA ATTTCGGG
PL-Gt8-MG2 CAAAATTCCA ATTTCGGG
PL-Gt8-MG5 CAAAATTCCA ATTTCGGG
PL-Gt8-MG2 CAAAATTCCA ATTTCGGG
PL-Gt8-MG2 CAAAATTCCA ATTTCGGG
PL-Gt8-MG3 CAAAATTCCA ATTTCGGG
PL-Gt8-MG2 CAAAATTCCA ATTTCGGG
PL-Gt8-I15 CAAAAT
PL-Gt8-I15 CAAAAT
PL-Gt8-REV
PL-GRE1-HL
PL-GRE3-HX
PL-GRS2-ID
PL-GRS1-WO
PL-GRS1-IL
PL-GRE2-HX
PL-GRE2-IL
PL-GRE3-IL
PL-GRE1-IL
PL-GRS3-IL
PL-HG
PL-HS41-IL C
Consensus caaaatt... ..

```





















2601		2730
Ca-GPE1	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE4	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE2	TCAgATGCcG	GGcAcCGcTtC
Ca-GM1	TCAgATGCcG	GGcAcCGcTtC
Ca-GM2	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE5	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE6	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE2	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE3	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE4	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE7	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE8	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE3	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE10	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE9	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE1	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE3	TCAgATGCcG	GGcAcCGcTtC
Ca-GPE3	TCAgATGCcG	GGcAcCGcTtC
Ca-GRE1	TCAgATGCcG	GGcAcCGcTtC
Ca-GRE2	TCAgATGCcG	GGcAcCGcTtC
Ca-Gt-a	TCAgATGCcG	GGcAcCGcTtC
Ca-Gt-b	TCAgATGCcG	GGcAcCGcTtC
Ca-Gt-c	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs1	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs4	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs11	TCAgATGCcG	GGcAcCGcTtC
Ca-Gt-v	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs4	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs41	TCAgATGCcG	GGcAcCGcTtC
Ca-Hs	TCAgATGCcG	GGcAcCGcTtC
Ca-HG-TM16	TCAgATGCcG	GGcAcCGcTtC
Ca-HG-TM19	TCAgATGCcG	GGcAcCGcTtC
Ca-HG-VL1	TCAgATGCcG	GGcAcCGcTtC
Consensus	TCAgATGCcG	GGcAcCGcTtC

2731		2764
Ca-GPE1	TAAAGcGcTt	CTtGcGcA
Ca-GPE4	TAAAGcGcTt	CTtGcGcG
Ca-GPE2	TAAAGcGcTt	CTtGcGcG
Ca-GM1	TAAAGcGcTt	CTtGcGcG
Ca-GM2	TAAAGcGcTt	CTtGcGcG
Ca-GPE5	TAAAGcGcTt	CTtGcGcG
Ca-GPE6	TAAAGcGcTt	CTtGcGcG
Ca-GPE2	TAAAGcGcTt	CTtGcGcG
Ca-GPE3	TAAAGcGcTt	CTtGcGcG
Ca-GPE4	TAAAGcGcTt	CTtGcGcG
Ca-GPE7	TAAAGcGcTt	CTtGcGcG
Ca-GPE8	TAAAGcGcTt	CTtGcGcG
Ca-GPE3	TAAAGcGcTt	CTtGcGcG
Ca-GPE10	TAAAGcGcTt	CTtGcGcG
Ca-GPE9	TAAAGcGcTt	CTtGcGcG
Ca-GPE1	TAAAGcGcTt	CTtGcGcG
Ca-GPE3	TAAAGcGcTt	CTtGcGcG
Ca-GPE3	TAAAGcGcTt	CTtGcGcG
Ca-GRE1	TAAAGcGcTt	CTtGcGcG
Ca-GRE2	TAAAGcGcTt	CTtGcGcG
Ca-Gt-a	TAAAGcGcTt	CTtGcGcG
Ca-Gt-b	TAAAGcGcTt	CTtGcGcG
Ca-Gt-c	TAAAGcGcTt	CTtGcGcG
Ca-Hs1	TAAAGcGcTt	CTtGcGcG
Ca-Hs4	TAAAGcGcTt	CTtGcGcG
Ca-Hs11	TAAAGcGcTt	CTtGcGcG
Ca-Gt-v	TAAAGcGcTt	CTtGcGcG
Ca-Hs4	TAAAGcGcTt	CTtGcGcG
Ca-Hs41	TAAAGcGcTt	CTtGcGcG
Ca-Hs	TAAAGcGcTt	CTtGcGcG
Ca-HG-TM16	TAAAGcGcTt	CTtGcGcG
Ca-HG-TM19	TAAAGcGcTt	CTtGcGcG
Ca-HG-VL1	TAAAGcGcTt	CTtGcGcG
Consensus	TAAAGcGcTt	CTtGcGcG

# ALIGNEMENT DES SEQUENCES GENOMIQUES DU FACTEUR D'ELONGATION (EXONS)

```
MSF: 498 Check: 0
Name: EF-GPE1-a Len: 498 Check: 4321 Weight: 1.00
Name: EF-GPE1-b Len: 498 Check: 2356 Weight: 1.00
Name: EF-GPE2-a Len: 498 Check: 3088 Weight: 1.00
Name: EF-GPE1-c Len: 498 Check: 7683 Weight: 1.00
Name: EF-GPE3-c Len: 498 Check: 6523 Weight: 1.00
Name: EF-GFS4-a Len: 498 Check: 3819 Weight: 1.00
Name: EF-GFS4-c Len: 498 Check: 9814 Weight: 1.00
Name: EF-GPE2-c Len: 498 Check: 5543 Weight: 1.00
Name: EF-GPE5-a Len: 498 Check: 4071 Weight: 1.00
Name: EF-GPE5-b Len: 498 Check: 4407 Weight: 1.00
Name: EF-GPE2-b Len: 498 Check: 3889 Weight: 1.00
Name: EF-GM2-b Len: 498 Check: 3903 Weight: 1.00
Name: EF-GPE3-a Len: 498 Check: 1749 Weight: 1.00
Name: EF-GPE3-b Len: 498 Check: 1400 Weight: 1.00
Name: EF-GPE3-c Len: 498 Check: 2034 Weight: 1.00
Name: EF-GPE3-a Len: 498 Check: 823 Weight: 1.00
Name: EF-GPE4-a Len: 498 Check: 2234 Weight: 1.00
Name: EF-GM1-f Len: 498 Check: 1214 Weight: 1.00
Name: EF-GM2-f Len: 498 Check: 321 Weight: 1.00
Name: EF-GM2-d Len: 498 Check: 1243 Weight: 1.00
Name: EF-GM2-e Len: 498 Check: 452 Weight: 1.00
Name: EF-GFS2-f Len: 498 Check: 8268 Weight: 1.00
Name: EF-GFS2-d Len: 498 Check: 9627 Weight: 1.00
Name: EF-GPE2-e Len: 498 Check: 9317 Weight: 1.00
Name: EF-GPE5-b Len: 498 Check: 9007 Weight: 1.00
Name: EF-GFS6-c Len: 498 Check: 8758 Weight: 1.00
Name: EF-GFS7-b Len: 498 Check: 7299 Weight: 1.00
Name: EF-GFS8-f Len: 498 Check: 8733 Weight: 1.00
Name: EF-GTV-b Len: 498 Check: 7293 Weight: 1.00
Name: EF-GTV-d Len: 498 Check: 8127 Weight: 1.00
Name: EF-GTV-c Len: 498 Check: 8098 Weight: 1.00
Name: EF-GTS-d Len: 498 Check: 9645 Weight: 1.00
Name: EF-GTS-h Len: 498 Check: 9645 Weight: 1.00
Name: EF-GTS-l Len: 498 Check: 9645 Weight: 1.00
Name: EF-GTT2-c Len: 498 Check: 8853 Weight: 1.00
Name: EF-GTT2-e Len: 498 Check: 8853 Weight: 1.00
Name: EF-GTT2-f Len: 498 Check: 8854 Weight: 1.00
Name: EF-GRE1-a Len: 498 Check: 9496 Weight: 1.00
Name: EF-GRE2-a Len: 498 Check: 486 Weight: 1.00
Name: EF-GTA-a Len: 498 Check: 8462 Weight: 1.00
Name: EF-GFS5-b Len: 498 Check: 1702 Weight: 1.00
Name: EF-GFS5-c Len: 498 Check: 3187 Weight: 1.00
Name: EF-GFS9-d Len: 498 Check: 7705 Weight: 1.00
Name: EF-GFS10-a Len: 498 Check: 7017 Weight: 1.00
Name: EF-GFS1-b Len: 498 Check: 7535 Weight: 1.00
Name: EF-GTA-b Len: 498 Check: 4253 Weight: 1.00
Name: EF-GTS-a Len: 498 Check: 4469 Weight: 1.00
Name: EF-GRE3-b Len: 498 Check: 4014 Weight: 1.00
Name: EF-GRE3-c Len: 498 Check: 4335 Weight: 1.00
Name: EF-GRS2-d Len: 498 Check: 7878 Weight: 1.00
Name: EF-GRS2-e Len: 498 Check: 7758 Weight: 1.00
Name: EF-GRS3-c Len: 498 Check: 9076 Weight: 1.00
Name: EF-GRS3-d Len: 498 Check: 8598 Weight: 1.00
Name: EF-GRS3-e Len: 498 Check: 8598 Weight: 1.00
Name: EF-GRS1-b Len: 498 Check: 4764 Weight: 1.00
Name: EF-GRS1-c Len: 498 Check: 4764 Weight: 1.00
Name: EF-GRE2-b Len: 498 Check: 1521 Weight: 1.00
Name: EF-GFS5-b Len: 498 Check: 3393 Weight: 1.00
Name: EF-GFS7-c Len: 498 Check: 3393 Weight: 1.00
Name: EF-GFS9-f Len: 498 Check: 5186 Weight: 1.00
Name: EF-GRE1-b Len: 498 Check: 6877 Weight: 1.00
Name: EF-GPE4-b Len: 498 Check: 8006 Weight: 1.00
Name: EF-GRE2-c Len: 498 Check: 7067 Weight: 1.00
Name: EF-GRE3-a Len: 498 Check: 9983 Weight: 1.00
Name: EF-GPE2-b Len: 498 Check: 6574 Weight: 1.00
Name: EF-GM1-a Len: 498 Check: 1586 Weight: 1.00
Name: EF-GM1-c Len: 498 Check: 2425 Weight: 1.00
Name: EF-GFS6-a Len: 498 Check: 2155 Weight: 1.00
Name: EF-GR-cDNA Len: 498 Check: 7830 Weight: 1.00
Name: EF-GTS-b Len: 498 Check: 4454 Weight: 1.00
Name: EF-GFS4-b Len: 498 Check: 6743 Weight: 1.00
Name: EF-GFS9-b Len: 498 Check: 3615 Weight: 1.00
Name: EF-GFS10-c Len: 498 Check: 2725 Weight: 1.00
Name: EF-GFS10-b Len: 498 Check: 3865 Weight: 1.00
Name: EF-GRE1-a Len: 498 Check: 2199 Weight: 1.00
Name: EF-GFS8-b Len: 498 Check: 8198 Weight: 1.00
Name: EF-GFS8-c Len: 498 Check: 9806 Weight: 1.00
Name: EF-GPE5-c Len: 498 Check: 4632 Weight: 1.00
Name: EF-GM1-b Len: 498 Check: 4129 Weight: 1.00
Name: EF-GM2-a Len: 498 Check: 2406 Weight: 1.00
Name: EF-GM2-c Len: 498 Check: 882 Weight: 1.00
Name: EF-GRE1-c Len: 498 Check: 916 Weight: 1.00
Name: EF-GTT1-b Len: 498 Check: 2135 Weight: 1.00
Name: EF-GTT1-c Len: 498 Check: 2714 Weight: 1.00
Name: EF-HS-cDNA Len: 498 Check: 3948 Weight: 1.00
Name: EF-HS41vir-a Len: 498 Check: 3185 Weight: 1.00
Name: EF-HS41-b Len: 498 Check: 3529 Weight: 1.00
Name: EF-HS41vir-b Len: 498 Check: 3412 Weight: 1.00
Name: EF-HS41-a Len: 498 Check: 3417 Weight: 1.00
Name: EF-HT-k Len: 498 Check: 1360 Weight: 1.00
Name: EF-HS41vir-c Len: 498 Check: 3214 Weight: 1.00
Name: EF-HS42-b Len: 498 Check: 4125 Weight: 1.00
Name: EF-HG1-a Len: 498 Check: 2024 Weight: 1.00
Name: EF-HG1-b Len: 498 Check: 1682 Weight: 1.00
Name: EF-HG16-a Len: 498 Check: 1040 Weight: 1.00
Name: EF-HS41-c Len: 498 Check: 1778 Weight: 1.00
Name: EF-HT-e Len: 498 Check: 3802 Weight: 1.00
Name: EF-HT-l Len: 498 Check: 215 Weight: 1.00
Name: EF-HC1C-b Len: 498 Check: 2574 Weight: 1.00
Name: EF-HC1C-a Len: 498 Check: 3483 Weight: 1.00
Name: EF-HC1C-c Len: 498 Check: 3509 Weight: 1.00
Name: EF-HG1-c Len: 498 Check: 943 Weight: 1.00
Name: EF-HH-a Len: 498 Check: 3870 Weight: 1.00
Name: EF-HCAJ-a Len: 498 Check: 4177 Weight: 1.00
Name: EF-HCAJ-b Len: 498 Check: 4177 Weight: 1.00
Name: EF-HCAJ-c Len: 498 Check: 5214 Weight: 1.00
Name: EF-HA-d Len: 498 Check: 5699 Weight: 1.00
Name: EF-HA-e Len: 498 Check: 7772 Weight: 1.00
Name: EF-HA-l Len: 498 Check: 7021 Weight: 1.00
Name: EF-HSAC-a Len: 498 Check: 3269 Weight: 1.00
Name: EF-HCAR-a Len: 498 Check: 7829 Weight: 1.00
Name: EF-HCAR-b Len: 498 Check: 9840 Weight: 1.00
Name: EF-HCAR-h Len: 498 Check: 8764 Weight: 1.00
Name: EF-GTT1-a Len: 498 Check: 6937 Weight: 1.00
Name: EF-GFS7-a Len: 498 Check: 9255 Weight: 1.00
Name: EF-GRE2-b Len: 498 Check: 9113 Weight: 1.00
Name: EF-HCAR-l Len: 498 Check: 4194 Weight: 1.00
Name: EF-HG-cDNA Len: 498 Check: 7760 Weight: 1.00
Name: EF-HS42-a Len: 498 Check: 678 Weight: 1.00
Name: Consensus Len: 498 Check: 1074 Weight: 0.00
```

//











*Annexe 4 :*  
*Communications*  
*affichées*



Poster  
N° B103

## Four new *Globodera* genes putatively involved in plant-nematode interaction

Alexandra BLANCHARD, Magali ESQUIBET, Didier FOUVILLE and Eric GRENIER

UMR INRA-AGROCAMPUS BIO3P. Domaine de la Motte. BP 35327, 35650 LeRheu. France.



**Nematode secretions have a key role** in plant-nematode interaction, particularly in cyst nematodes of the genus *Globodera*. Among them, *Globodera pallida* and *G. "mexicana"*, are two genetically closely related species that can be differentiated through their host range. A previous study by SSH has pointed out 4 TDFs\* differentially expressed between these 2 species.

\* TDF : Transcribed Differential Fragments

Here, we have located the transcripts of interest and obtained the full length cDNA in order to get an insight into their role in plant-nematode interaction.

	IB3	IVC12	IA7	IVG9
TDFs size	343bp	230bp	201bp	419bp
Over expression	<i>G. pallida</i>	<i>G. pallida</i>	<i>G. pallida</i>	<i>G. "mexicana"</i>

### RESULTS

#### Localization of the transcripts: is the product secreted ?

##### Method

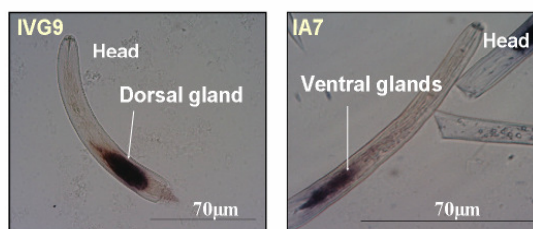
*In situ* hybridisation using Dig-labelled probes allowed us to locate the different mRNA in *G. pallida* and *G. "mexicana"* juveniles (J2).

##### Results

Two transcripts (IB3 and IVC12) had an intestinal location. The other two (IA7 and IVG9) were located in **oesophageal glands**: **IA7 in the ventral glands**, and **IVG9 in the dorsal gland**.

The oesophageal glands secretions are reported to be the key for plant-nematode interactions. So we focused further studies on IA7 and IVG9.

Because of their localization, IA7 should be involved in the earliest steps of parasitism (invasion and migration), whereas IVG9 could act later (syncytium initiation).



G. "mexicana" (Santa Ana)

G. pallida (Guiclan)

#### Full length cDNA: what is the role of corresponding proteins in parasitism ?

##### Method

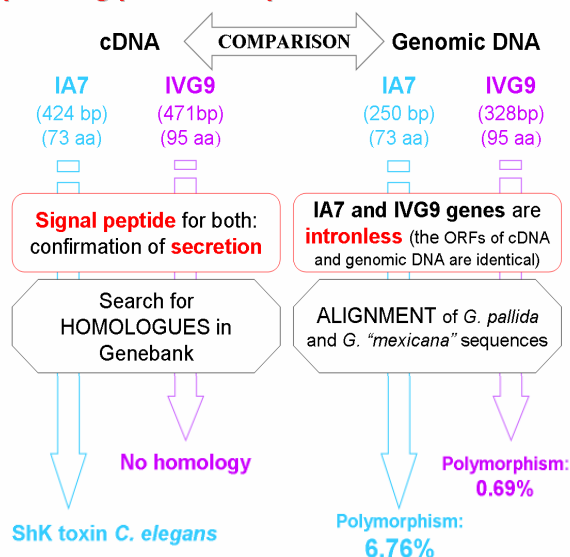
5' and 3'RACE allowed us to obtain the complete cDNAs. Open Reading Frames were amplified with new primers and each ORF was cloned and sequenced. These primers were also used to amplify genomic sequences.

##### Results

The **secretion of the protein** corresponding to the two transcripts was confirmed by the presence of a peptide signal.

**No homology** was obtained for IVG9 sequence in data bases, but IA7 matched with an **ShK-like** (*Stichodactyla helianthus* K channel) **toxin** identified in *C. elegans*, a gene putatively involved in parasitism.

IA7 sequences showed an important level of polymorphism between *G. pallida* and *G. "mexicana"*. **Is this polymorphism the origin of differential pathogenicity ?** As among the 15 substitutions found in the 222bp of the ORF, **14 are non-synonymous**, we can suspect that the proteins encoded by IA7 in *G. pallida* and *G. "mexicana"* are functionally different.



### CONCLUSION



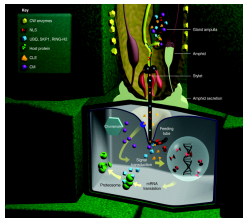
**This work revealed 2 genes** probably involved in plant-nematode interaction because of their localization which suggests a **secretion of corresponding proteins** and because of the homologues or lack of homologues obtained in data bases.

Further studies are in progress to determine **if the transcripts are stage-specific** and to further investigate the **polymorphism of these genes or gene families**.

## Identification and polymorphism of new potato cyst nematode genes putatively involved in plant-nematode interaction

Eric GRENIER, Alexandra BLANCHARD, Magali ESQUIBET and Didier FOUVILLE

UMR INRA-AGROCAMPUS BIO3P. Domaine de la Motte. BP 35327, 35650 Le Rheu. France.

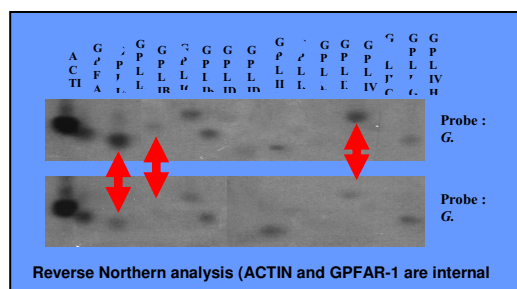


*Globodera pallida* and *G. mexicana* are closely related nematode species that can mate and form viable hybrids but usually develop on different Solanaceous plants. These plant parasites have evolved sophisticated relationships with host cells to sustain a sedentary parasitic habit through the elaboration and maintenance of a feeding site.

Host	<i>Solanum tuberosum</i>	<i>Solanum nigrum</i>	<i>Lycopersicon esculentum</i>	<i>Nicotiana tabacum</i>
<i>G. rostochiensis</i>	+	-	+	-
<i>G. pallida</i>	+	-	+	-
<i>G.</i>	-	+	+	-

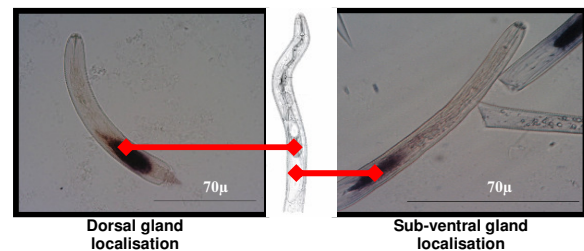
Identification of nematode genes involved in parasitism is important for elucidation of disease resistance mechanisms in plants. We have used suppression subtractive hybridization (SSH) to selectively amplify genes expressed or over-expressed in one of these nematode species while suppressing the amplification of genes common to both species.

Sequence analysis of the SSH clones revealed that nearly 85% of the sequences cloned were nematode specific and a high proportion were pioneer genes. Interestingly, homologues of a cellulase described in *G. rostochiensis* was also found. None appeared to be completely absent from the other transcriptome. Using reverse northern, clear differential expression was confirmed for some cDNA fragments. No signal was sometimes observed no matter which probe was used.



Pioneer genes which have no significant matches in *C. elegans* genome but display differential gene expression were considered of particular interest. Nematode oesophageal glands secretions released from the stylet are known to play a central role in plant parasitism. Transcripts of interest were therefore also studied by *in situ* hybridization.

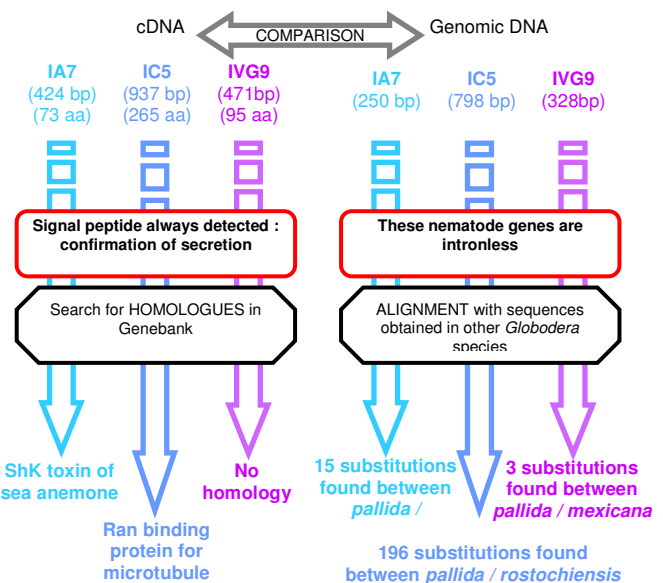
Because of their observed localisation, we assumed that transcript IA7 should be involved in the earliest steps of parasitism (invasion and migration), whereas transcripts IVG9 and IC5 should act later (feeding site initiation).



Detection of a peptide signal in the 5' end of the full length cDNA sequences of IA7, IC5 and IVG9 confirmed the putative secretion of the proteins corresponding to these three transcripts. No homology was obtained for IVG9 sequence in data bases, but IA7 matched with an ShK-like (*Stichodactyla helianthus* K channel) toxin and IC5 with a Ran binding protein in the microtubule organizing center.

Sequence comparisons with homologues sequenced in *G. mexicana* or found in data bases (*G. rostochiensis* species) revealed different levels of polymorphism and in the case of IC5 different members of a gene family. Compared to IVG9, IA7 sequences showed an important level of polymorphism between *G. pallida* and *G. mexicana*. Is this polymorphism the origin of differential pathogenicity? As among the 15 substitutions found in the open reading frame, 14 are non-synonymous, we can suspect that the proteins encoded by IA7 in *G. pallida* and *G. mexicana* are functionally different.

This work revealed three genes probably involved in plant nematode interaction. Further studies are in progress to investigate the sequence polymorphism of these genes or gene families in a large range of *Globodera* species and populations. Comparison of nonsynonymous / synonymous rate ratio in these protein coding genes should help us to determine if they are under a purifying or diversifying selection and provide us a first knowledge of the molecular evolution of the stylet secretions. These data are particularly of interest in the framework of developing artificial resistances against these plant parasites.

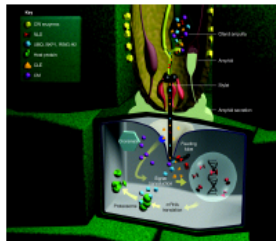


## EVOLUTION OF PARASITISM GENES IN THE HETERODERIDAE PLANT PARASITIC NEMATODES FAMILY

Alexandra BLANCHARD<sup>(1)</sup>, Olivier PLANTARD<sup>(1)</sup>, Philippe CASTAGNONE<sup>(2)</sup>, Marie-Noelle ROSSO<sup>(2)</sup> & Eric GRENIER<sup>(1)</sup>

(1) UMR INRA-AGROCAMPUS BIO3P, Domaine de la Motte, BP 35327, 35653 Le Rheu, France.

(2) UMR INRA/UNSA/CNRS Interactions Plantes-Microorganismes et Santé Végétale, 400, route des Chappes, 06903 Sophia Antipolis cedex, France



Cyst (*Globodera* and *Heterodera*) and root-knot (*Meloidogyne*) nematodes of the Heteroderidae family have evolved sophisticated relationships with plant hosts to sustain a sedentary parasitic habit. In this study we have evaluated the nucleotidic variations of two parasitism genes : (1) a cathepsin L product which was shown to be differentially expressed between virulent and avirulent populations of *Meloidogyne incognita*, (2) a pectate lyase product involved in the early events of parasitism. For each of these parasitism genes, polymorphism levels were compared with data obtained using a housekeeping gene, the elongation factor 1 $\alpha$ .

The dataset obtained for 37 populations of cyst nematodes

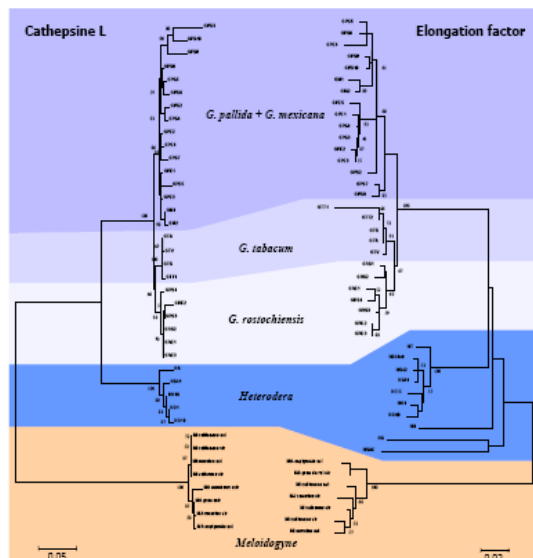
Pectate lyase	Size of the coding sequence (aa)	Number of variable sites	Number of polymorphic informative characters	%aT
Nucleotide cDNA sequence	690 bp	239	85	55.4
Amino acid sequence	230 aa	72	47	-
Cathepsin L	Size of the coding sequence (aa)	Number of variable sites	Number of polymorphic informative characters	%aT
Nucleotide gene sequence	861 bp	299	100	47.4
Amino acid sequence	287 aa	111	58	-
Elongation factor 1 $\alpha$	Size of the coding sequence (aa)	Number of variable sites	Number of polymorphic informative characters	%aT
Nucleotide gene sequence	371 bp	132	101	47.1
Amino acid sequence	122 aa	43	34	-

### RESULTS

Amplification products were obtained for 46 populations belonging to 14 *Globodera*, *Heterodera* and *Meloidogyne* species. Each gene was analysed in terms of polymorphism, phylogeny and rates of evolution.

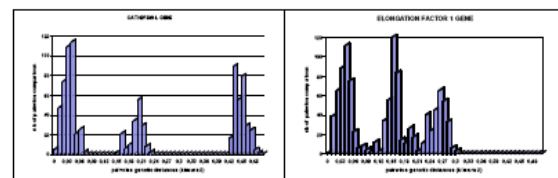
#### Parasitism genes tell the same history than housekeeping genes

Phylogenetic trees were generated in Neighbour Joining using the Kimura 2 parameters model. The clustering observed is in accordance with the species relationships admitted in the Heteroderidae family using ribosomal markers, except that *G. mexicana* appeared here to be part of the *G. pallida* species. According to branch length, root-knot nematodes appeared to be distantly related to cyst nematodes.

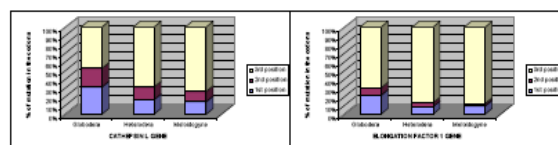


#### But, parasitism genes evolved differently than housekeeping genes

Analysis of the genetic distances revealed distinct features in the two studied parasitism genes compared to the housekeeping gene. In all cases, the distribution of the pairwise comparisons showed three groups corresponding to (from left to right) : intra genera pairwise comparisons, *Globodera/Heterodera* pairwise comparisons and cyst/root-knot nematodes pairwise comparisons, but the distribution appeared discontinuous with both the pectate lyase (data not shown) and cathepsin L genes suggesting an acceleration of evolution for parasitism genes, especially since the cyst *versus* root-knot divergence.



Study of the distribution of the mutations in the codons revealed that parasitism genes display a higher percentage of mutations in the 1st and 2nd positions of the codons (that mainly correspond to non synonymous mutations) than the housekeeping gene.



#### Positive selection involved in the cathepsin L parasitism gene ?

As a first analysis, dN/dS ratios were calculated using PALM on the total cathepsin L dataset. Results showed that this gene is under purifying selection (dN/dS < 1). Root-knot nematodes appeared to be again quite distantly related to cyst nematodes as a saturation effect was observed with dS values > 2 for all root-knot/cyst nematodes comparisons. More, a strong bias in the transition/transversion was observed on the 3<sup>rd</sup> positions of the root-knot nematode cathepsin L codons.

	<i>Globodera</i>	<i>Heterodera</i>	<i>Meloidogyne</i>
<i>Globodera</i>	0.157	0.114	-
<i>Heterodera</i>		0.124	-
<i>Meloidogyne</i>			0.046

### CONCLUSIONS

The two nematode parasitism genes investigated in this study showed both a high percentage of substitution on the 1<sup>st</sup> and 2<sup>nd</sup> positions of the codons and an acceleration of evolution, which suggest that they can be under diversifying selection. At the scale of the total dataset dN/dS ratio showed that these genes are under purifying selection. Investigations must be now carried out at a finer scale to determine sequence regions or nematode populations where positive selection may be detected.



**RESUME**

Les nématodes à kyste du genre *Globodera* sont des parasites telluriques inféodés aux Solanacées. Dans une interaction compatible, les seconds stades larvaires éclosent dans le sol et se dirigent vers les parties apicales des racines pour y pénétrer et migrer de façon intracellulaire jusqu'au cylindre central. Ils mettent en place une structure très élaborée, le syncytium, leur permettant de se nourrir et d'achever leur cycle de développement. Cette structure résulte d'une re-programmation génétique des cellules végétales. Les bases moléculaires de l'interaction plante/nématodes sont très peu connues mais pourraient grandement améliorer nos connaissances sur les nématodes phytoparasites. Nous avons ainsi caractérisé de nouveaux gènes du pouvoir pathogène et décrit leur variabilité afin de comprendre leur évolution. Cette approche nous a permis d'identifier les contraintes sélectives auxquelles ils sont soumis et qui définissent pour partie les capacités adaptatives de certaines espèces, populations ou genres.

Nous avons isolé trois nouveaux gènes du pouvoir pathogène chez *Globodera pallida* dont les transcrits sont localisés dans les glandes salivaires et qui correspondent ainsi à des protéines sécrétées par le nématode dans la plante. Parmi ces gènes, nous avons mis en évidence pour la première fois chez *G. pallida*, un homologue de RanBPM, *Gp-rbp-1*, spécifiquement exprimé aux stades pré-parasitaires et parasitaires qui conduisent à l'élaboration du syncytium. Les résultats obtenus en Southern blot suggèrent que ce gène appartient à une famille multigénique. L'étendue de la variabilité de cette famille de gènes a été étudiée par criblage d'une banque d'ADNc de *G. pallida* : 12 copies différentes de RanBPM ont été séquencées et présentent un taux de variation nucléotidique de 52,7%. L'analyse de la position des substitutions pour les gènes du parasitisme nouvellement identifiés montrent qu'ils présentent tous les trois un taux important de mutations en première et deuxième position de codon (~ 50%) suggérant une proportion très élevée de substitutions non synonymes. Toutefois les analyses devront être réalisées sur un jeu de données plus large à l'image de ce que nous avons réalisé sur deux autres gènes pour lesquels nous avons davantage d'indications sur leur rôle dans le parasitisme.

Parmi les nombreux gènes de la littérature, nous avons choisis un gène impliqué dans les phases précoces, la pectate lyase, et un autre dans les phases tardives, la cathepsine L. En parallèle, un gène de ménage, le facteur d'élongation  $1\alpha$ , *a priori* non lié au parasitisme a été choisi comme référent. Nous avons amplifié ces trois gènes dans 40 populations de nématodes à kyste. Après vérification de l'orthologie des séquences obtenues, nous avons montré que la variabilité des trois gènes est similaire, même si les distances génétiques calculées montrent une accélération du tempo d'évolution des gènes du parasitisme. Les substitutions observées sont beaucoup plus importantes sur les deux premières positions de codon des gènes de parasitisme (50%) que pour le gène de ménage (20%). Des analyses pointues des pressions de sélection ont permis de montrer que certaines populations et espèces présentent des taux d'évolution plus importants, qui pourraient être le reflet de contraintes exercées par les différentes plantes hôtes. De plus, nous avons montré que le gène cathepsine L possède une zone fortement soumise à la sélection positive, indiquant un fort potentiel évolutif de ce gène.

**MOTS CLES** : RanBPM, nématodes à kyste, pectate lyase, cathepsine L, sélection positive, sécrétions.

---

**ABSTRACT**

The cyst nematodes of the *Globodera* genus are telluric organisms that exclusively parasitize the Solanaceous botanic family. During compatible interaction, the second stage larvae hatch in the soil and move towards the apical parts of the roots where they penetrate. They then migrate intracellularly to the central cylinder. They elaborate an intimate relationship called syncytium which allows them to feed and complete their life cycle. This structure is the result of a genetic reprogramming of plant cells. The molecular bases of the plant/nematode interaction are relatively unknown but could largely improve our knowledge on these phytoparasitic nematodes. We first chose to characterise new pathogenicity genes, and then to study the variability of these genes in order to understand their evolution. This approach should make it possible to identify the selective constraints to which they are subjected and which define to some extent the adaptive capacities of some species, populations or genera.

We isolated three new pathogenicity genes in *Globodera pallida* for which transcripts were localised in salivary glands, corresponding to proteins secreted by the nematode in the plant. Among these genes, we highlighted for the first time in *G. pallida* a RanBPM homolog, *Gp-rbp-1*, specifically expressed at the pre parasitic and parasitic stages which lead to the development of the syncytium. Southern blot results suggest that this gene belongs to a multigenic family. The extent of the variability of this gene family was studied by screening a *G. pallida* cDNA library: 12 different RanBPM copies were sequenced and showed a 52.7% nucleotidic variability. Analysis of the substitutions positions observed for the three genes highlighted a high mutation rate on first and second codon position (~ 50%) suggesting a high proportion of non synonymous substitutions. However, the analyses will have to be carried out on a broader data set, as we did for two other genes for which we had more indications on their implication in parasitism.

Among many genes of the literature, we chose one gene involved in the early steps of parasitism, the pectate lyase gene, and one involved in the late stages, the cathepsin L gene. The elongation factor 1 $\alpha$ , *a priori* not related to the parasitism, was used as a comparison. We amplified these three genes in 40 populations of cyst nematodes. After checking of the orthology of the sequences, we have shown that the same variability was observed whatever the gene considered, even if the calculated genetic distances showed an increase of the evolutionary tempo on parasitism genes. Substitutions observed in the first two positions of the codon were much more important for parasitism genes (50%) than for the housekeeping gene (20%). Sharp analyses of the selection pressures made it possible to show that particular populations and species showed evolution rates more important than could be related in some instance to the influence of the host plant. Moreover, we showed that the cathepsin L gene possesses a domain subjected to positive selection, indicating a strong evolutionary potential of this gene.

**KEYWORDS:** RanBPM, cyst nematodes, pectate lyase, cathepsin L, positive selection, secretions.