



Analyse évolutive et fonctionnelle des Macrophage Migration Inhibitory Factors chez les eucaryotes

Claire Michelet

► To cite this version:

Claire Michelet. Analyse évolutive et fonctionnelle des Macrophage Migration Inhibitory Factors chez les eucaryotes. Biologie moléculaire. Université Côte d'Azur, 2018. Français. NNT : 2018AZUR4109 . tel-02426269

HAL Id: tel-02426269

<https://theses.hal.science/tel-02426269>

Submitted on 2 Jan 2020

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

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



THÈSE DE DOCTORAT

Analyse évolutive et fonctionnelle des Macrophage Migration Inhibitory Factors chez les eucaryotes

Claire Michelet

Institut Sophia Agrobiotech

Présentée en vue de l'obtention
du grade de docteur en SVS
au sein de l'Université Côte d'Azur
Dirigée par Christine Coustau
Soutenue le 30 novembre 2018
à l'Institut Sophia Agrobiotech

Devant le jury, composé de :

Présidente du jury

Dr Marie-Noëlle Rosso Directrice de recherche INRA

Rapporteurs

Dr Guillaume Mitta Professeur des Universités UPVH

Dr Eric Viscogliosi Directeur de recherche CNRS

Examineurs/Examinatrices

Dr Gabriel Markov Chargé de recherche CNRS

Dr Raquel Tavares Maître de conférence Univ. Lyon 1

Directrice de thèse

Dr Christine Coustau Directrice de recherche CNRS

Membre Invité

Dr Harald Keller Directeur de recherche INRA

Analyse évolutive et fonctionnelle des « Macrophage Migration Inhibitory Factor » chez les eucaryotes

Les cytokines MIFs (Macrophage Migration Inhibitory Factor) sont des protéines multifonctionnelles qui, chez les mammifères, interviennent dans plusieurs processus majeurs tels que le contrôle du cycle et de la mobilité cellulaire, l'activation de la réponse immunitaire et l'inhibition de l'apoptose. Des travaux récents montrent que les protéines MIFs peuvent également jouer un rôle majeur dans l'immunité des invertébrés, et être utilisées par des organismes parasites d'animaux ou de végétaux pour inhiber les défenses de leurs hôtes respectifs, ce qui soulève la question de leur diversité, de leur histoire évolutive et des potentielles différences fonctionnelles.

L'objectif général de ce travail de thèse était d'explorer la diversité et l'histoire évolutive des protéines MIFs à une échelle trans-règne, puis de rechercher leurs éventuelles différences fonctionnelles, en se focalisant sur les systèmes plantes-pathogènes.

Nous avons tout d'abord identifié les MIFs chez 803 espèces de plantes, champignons, protistes, et métazoaires, et analysé leur présence/absence et histoire évolutive en fonction des taxa, de l'écologie et du mode de vie (libre ou parasitaire) des espèces. Nous avons montré que l'histoire évolutive des MIFs, chez les eucaryotes, est complexe et implique des duplications ancestrales ainsi que des pertes multiples ou des re-duplications récentes. Les plantes (espèces libres autotrophes) et les parasites de plantes (autres que champignons) possèdent un nombre médian de trois MIFs, alors que les espèces hétérotrophes et les parasites d'animaux ont un nombre de MIF plus faible et/ou plus variable. De plus, les protéines MIFs semblent essentielles et fortement conservées, avec de nombreux résidus sous sélection purifiante, chez certains groupes comme les plantes, alors que dans d'autres groupes, elles semblent facultatives (e.g. champignons) ou présentes en plusieurs copies divergentes (e.g. nématodes, insectes), ce qui suggère de potentielles néofonctionnalisations.

Nous avons ensuite analysé l'effet des protéines MIFs de plusieurs espèces sur la mort cellulaire en système végétal. Tous les organismes testés (plantes oomycètes, protozoaires, insectes et nématodes), y compris ceux n'ayant pas d'interaction avec les plantes, possèdent au moins une protéine MIF capable d'inhiber cette mort cellulaire. Cela suggère que l'inhibition de la mort cellulaire en plante ne correspond pas à une néofonctionnalisation des MIFs de parasites de plantes, mais serait liée à des propriétés structurales et conservées des MIFs. Toutefois, aucun des paramètres étudiés (localisation subcellulaire) ou prédits *in silico* (présence de motifs, structures 3D, oligomérisation, modifications post-traductionnelles) ne semble lié à cette activité d'inhibition de la mort cellulaire. De futures études fonctionnelles poussées sont nécessaires à l'élucidation des relations structure/fonction de ces protéines complexes.

Mots clé: Macrophage Migration Inhibitory factor; reconstructions phylogénétiques, eucaryotes, Inhibition de mort cellulaire

Evolutionary and functional analysis of MIF in eukaryotes.

Macrophage migration inhibitory factors (MIF) are multifunctional proteins regulating major processes in mammals, including control of the cell cycle and migration, activation of innate immune responses, and prevention of p53-mediated apoptosis. MIF proteins also play a role in innate immunity of invertebrate organisms or serve as virulence factors in parasitic organisms, raising the question of their evolutionary history and of a putative differential evolution of structure/function relationships.

The general aim of this PhD was to explore the diversity and evolutionary history of MIF proteins across kingdoms, and to investigate their potential functional differences, with a special emphasis on host-parasite systems.

We first performed a broad survey of MIF presence or absence and evolutionary relationships across 803 species of plants, fungi, protists, and animals, and explored a potential relation with the taxonomic status, the ecology, and the lifestyle of individual species. We show that MIF evolutionary history in eukaryotes is complex, involving ancestral duplications, multiple gene losses and recent clade-specific re-duplications. Of note, plants and plant parasites (other than fungi) harbour a median number of three MIFs, while heterotrophic and animal parasite species harbour a lower or/and variable MIF number. Intriguingly, MIFs seem to be essential and highly conserved with many sites under purifying selection in some kingdoms (e.g. plants), while in other kingdoms they appear more dispensable (e.g. in fungi) or present in several diverged variants (e.g. insects, nematodes), suggesting potential neofunctionalizations within the protein superfamily.

We then analysed the effect of MIF proteins from selected species on plant cell death. All organisms tested (plant, oomycetes, protozoa, insects, and nematodes) including species that are not in interaction with plants, possess at least one MIF protein showing a significant cell death inhibitory effect. This suggests that plant cell death inhibition does not result from a neofunctionalization of MIF from plant-parasites, and is related to conserved structural features of MIF proteins. However, none of the parameters predicted *in silico* (sequence motifs, 3D structures, oligomerization, post-translational modifications) appeared to be related to the cell death inhibitory activity. Future extensive functional studies are necessary to unravel the structure-function relationship of these evolutionarily and functionally complex proteins.

Key words: Macrophage Migration Inhibitory factor; phylogenetic reconstructions, eukaryotes, cell death inhibition

« Tu sais que le corps humain est différent de ce qu'il était auparavant.
et qu'il s'est adapté à ce monde pollué. Il ne s'agit pas seulement des
humains. Les plantes et les arbres aussi ont dû évoluer »

Hayao Miyazaki – Nausicaä, la vallée du vent

Remerciements

Tout d'abord, un grand merci à **Christine** et **Harald**, pour leur aide et leur soutien durant l'ensemble de ma thèse. Je ne serais pas la même aujourd'hui si je ne vous avais pas rencontrés. J'ai adoré passer ces années avec vous. Je sais que je ne vous l'ai pas forcément beaucoup montré mais sincèrement merci pour votre aide, vos encouragements mais aussi pour vos critiques constructives. Merci de m'avoir appris tant de choses, de m'avoir donné confiance en moi et en mes capacités scientifiques...

Merci à **Maelle**, pour tes nombreux conseils, merci de m'avoir montré le chemin des vraies expérimentations, de m'avoir encouragée et ouvert la voie. J'aurais aimé que tu sois là 2 ans plus tôt. Merci pour ton pessimisme qui fait écho au mien et qui m'aide à toujours relativiser mes résultats.

Merci à **Thierry**, sans qui mes clonages seraient encore à l'état larvaire, merci pour toute l'aide que tu m'as apportée, pour les nombreux conseils avisés que tu as pu faire et la patience exceptionnelle dont tu as fait preuve à mon égard.

Merci à **Etienne**, pour tous tes conseils, tes coups de mains et surtout pour m'avoir montré le monde de MrBayes et de Phylml ... Merci de m'avoir fait découvrir le monde des nématodes et à quel point ils pouvaient être bizarres et passionnants à la fois.

Merci à tout le plateau de bio-informatique, sans qui mes analyses seraient restées dans des cartons. Merci de m'avoir fait découvrir Linux, et la programmation et de m'avoir donné tant envie de continuer sur cette voie.

Tous mes remerciements à **Pascaline** pour tes remarques avisées.

Un grand merci aussi à mes collègues de bureau, **Fanny, Bin et Luo**, pour m'avoir supportée pendant 3 ans dans leur bureau, pour m'avoir soutenue dans mes journées bio-informatiques, et comprise dans mes périodes de stress et bonne continuation à vous dans le monde merveilleux des drosophiles et de leurs parasitoïdes.

Merci à **Chami**, pour m'avoir écoutée et soutenue, comprise surtout au moment où je n'étais franchement plus sûre de moi du tout... Il est à la fois étonnant et incroyablement frustrant de voir que les gens ne te traite pas comme tu le mérites.

Merci à **Olivier**, pour les plantes, pour les phases de rire et de discussions dans les couloirs.

Merci à **Fred** et **Gaëlle**, bonne continuation au pays des P450. Merci pour vos remarques et idées qui m'ont permis d'avancer.

Merci à **Philippe**, puisses-tu continuer de t'envoler aussi souvent que possible et merci pour les recommandations littéraires

Merci à toute **l'équipe IPO** et à toute **l'équipe ID** pour leur soutien.

La Science ne fait pas tout donc merci donc aussi à tous ceux du milieu non scientifique, tous ceux qui n'ont simplement rien compris à mes analyses mais qui m'ont soutenue quand même.

Merci, par conséquent, à **mes parents**, qui m'auront bien amusée pendant ces 3 ans à parler de « mes pucerons – moucheron – truc mif », « mon ordinateur » et « mes plantes ». Je ne suis pas sûre que vous ayez compris quoi que ce soit à ce que j'expliquais mais merci de m'avoir écoutée quand même.

Merci à **Guillaume, Anaïs et Loïc**, j'espère ne pas trop vous avoir traumatisés à parler biologie. Prenez bien soin de votre petit et j'espère que nous pourrons continuer à nous voir.

Merci à **Fod, Merry et Marjo**, de m'avoir rappelé à quel point la Biologie peut être incompréhensible et foncièrement inintéressante si elle n'est pas bien expliquée. Merci de m'avoir donné envie de faire de la pédagogie.

Merci à **Boubou, Cyril et Paul**, pour les parties de jeu de rôle et avoir accepté que je me défoule régulièrement sur leurs personnages. Merci d'avoir accepté que mon humeur influe sur vos chances de survie, merci de m'avoir aidée à apprécier la vie telle qu'elle est. Un merci tout particulier à **Cyril** pour ses très nombreux messages d'encouragement tout au long de ma thèse, qui m'ont aidé à définir mes priorités musicales et littéraires des 10 prochaines années.

Merci à **Pauline**, qui me rappelle chaque jour qui passe combien la vie serait terne sans nos amis les animaux.

Merci à **Remi et Sam**, de m'avoir incitée à faire tellement de choses pendant ma thèse, merci de m'avoir pillé le temps qu'il me restait à faire des projets de fous... Merci de me rappeler chaque jour qui passe qu'il y a d'autres combats à mener que ceux que l'on affronte lors d'une thèse. Merci d'avoir essayé de comprendre, merci de m'avoir ouvert votre porte quand j'en avais besoin, merci pour tout !

Merci à **Cannelle**, pour ces nombreux « cadeaux », pour sa tentative d'assassinat de mon ordinateur en pleine rédaction, pour sa propension à transmettre sa flemme, et sa volonté de m'empêcher de dormir quand j'en ai le plus besoin. Merci de rendre ma vie plus drôle, plus caline et surtout beaucoup moins rangé.

Enfin pour finir en dernier dans les remerciements mais pas dans mon cœur, un très grand **MERCI** à **Pierre**, pour avoir supporté mes crises de nerfs, de larmes et de « je n'y arriverai jamais ». Merci d'être là, chaque fois que j'ai besoin de toi. Merci d'avoir compris (et accepté surtout) mon caractère qui, je le sais, n'est pas facile tous les jours. Merci de m'avoir soutenue dans mes pires moments et d'être heureux pour moi dans les meilleurs. Merci d'avoir su m'écouter et m'encourager. Merci d'être encore et toujours l'énergie qui me permet d'avancer, le sang dans mes veines et l'arbre sous lequel je me repose. Merci d'avoir compris, toi le premier, pourquoi j'aimais la Science.

Table des matières

Partie I : Introduction Générale	13
Partie II : Contexte bibliographique	17
Chapitre 1 : Les cytokines MIFs chez l'Homme	19
II.1.1- Généralités.....	19
II.1.2- Fonctions des MIFs chez l'être humain	23
Chapitre 2 : La mort cellulaire et les protéines MIFs	27
II.2.1- La PCD de type I ou l'apoptose	28
II.2.2- La PCD de type II ou l'autophagie.....	31
II.2.3- La PCD de type III ou la nécrose	32
Chapitre 3 : Les cytokines MIFs chez les autres espèces	34
II.3.1- Chez les espèces libres	34
II.3.2- Chez les parasites d'animaux	35
II.3.3- Les cytokines MIFs chez les pucerons	35
II.3.4- Les cytokines MIFs chez les plantes	37
Chapitre 4 : Diversité des MIFs	38
Partie III : Objectifs de la thèse	39
Partie IV : Matériels et Méthodes	43
Chapitre 1 : Analyses <i>in silico</i>	45
IV.1.1- Analyse phylogénétique des protéines MIFs.....	45
IV.1.2- Analyse de la sélection purificatrice ou diversifiante.....	48
IV.1.3- Analyses statistiques du nombre de protéines MIFs	49
IV.1.4- Analyses structurales des MIFs	51
IV.1.5- Prédiction des liaisons hydrogènes.....	52
IV.1.6- Prédiction des modifications post-traductionelles des protéines.....	52
Chapitre 2 : Analyses <i>in vivo</i>	53
IV.2.1- Expression des gènes <i>MIF</i> en cellule végétale	53
IV.2.2- Evaluation de la mort cellulaire.....	56
IV.2.3- Localisation subcellulaire	57
Partie V : Résultats et Discussions	59
Chapitre 1 : Histoire évolutive des MIFs.....	61
V.1.1- Présence des MIFs chez les eucaryotes.....	61
V.1.2- Reconstruction phylogénétique des MIFs d'eucaryotes.....	67
V.1.3- Analyse des sites sous sélection	79
V.1.4- Conclusion.....	82

Chapitre 2 : La modulation de la mort cellulaire par les MIFs	83
V.2.1- Introduction	83
V.2.2- Choix des modèles biologiques	84
V.2.3- Effet des MIFs sur la modulation de la mort cellulaire en modèle végétal.....	87
V.2.4- Localisation subcellulaire des MIFs.....	89
V.2.5- Recherche de relation structure-fonction	92
V.2.6- Analyse de la structure des MIFs	94
V.2.7- Analyse des modifications post-traductionnelles	99
V.2.8- Conclusions	107
Partie VI : Conclusion générale et perspectives	109
Partie V : Annexes.....	119
Article issus de mon travail de thèse	121
Figures supplémentaires	142
Bibliographie	178

Listes des abréviations

ABC	ATP Binding Cassette
ACKR3	Atypical Chemokine Receptor 3
ACM	Analyse multiple de correspondance
AFD	Analyse factorielle discriminante
AIC	Critère d'information d'Aikaide
AIF	Apoptosis inducing factor
aLRT	Test de ratio de probabilité approximatif
AtMDL	MIF d' <i>Arabidopsis thaliana</i>
Avrpto	Protéine d'avirulence du gène pto
BI	Inference Bayesienne
BxMIF	MIF de <i>Bursaphelenchus xylophilus</i>
CeMIF	MIF de <i>Caenorhabditis elegans</i>
CMH-II	Complexe Majeur d'Histocompatibilité de type II
DAMP	Damage-associated molecular patterns
D-DT	D-Dopachrome Tautomérase
DR	Death Receptor
FEL	Fixed Effects Likelihood
FUBAR	Fast Unconstrained Bayesian Approximation
GlcNAc	N-acetylglucosamine
HCA	Hydrophobic Cluster Analysis
HR	Réponse Hypersensible
HsMIF	MIF d' <i>Homo sapiens</i>
IL-1 β	Interleukine 1 bêta
IL8RB	Récepteur Bêta à l'Interleukine 8
LPS	Lipolysaccharides
MAP Kinase	Mitogen Activated Protein Kinase
MCP-1	Monocyte Chemoattractant Protein 1
MDL	MIF D-Dopachrome Tautomerase Like
MEME	Mixed-Effects Model of Evolution
MIF	Macrophage Migration Inhibitory Factor
MiMIF	MIF de <i>Meloidogyne incognita</i>
ML	Maximum Likelihood
MmMIF	MIF de <i>Mus musculus</i>
MOMP	Mitochondrial Outer Membrane Permeabilisation
NNI	Nearest Neighbor Interchanges
O-GlcNAc	O- (β -linked)- N-acetylglucosamine
P53	Tumor protein 53 (ou TP53)
PAMP	Pathogen Associated Molecular Patterns
PARP-1	Poly(ADP-ribose) polymerase-1
PCD	Programmed Cell Death
PFK-2	Enzyme 6-phosphofructo-2 kinase
PfMIF	MIF de <i>Plasmodium falciparum</i>
PGE ₂	Prostaglandine E2
PiMIF	MIF de <i>Phytophthora infestans</i>
PP	Posterior Probability
PRR	Pattern Recognition Receptor
PyMIF	MIF de <i>Plasmodium yoelii</i>
REL	Random Effect Likelihood
RIPK	Receptor interacting kinase
RMSD	Root mean square analysis

ROS	Reactive Oxygen Species
SINA	SILVA Incremental Aligner
SLAC	Single Likelihood Ancestor Counting
SIMIF	MIF de <i>Solanum lycopersicum</i>
SPR	Subtree Pruning and Regrafting
TLR4	Toll like Receptor 4
TNF- α	Tumor Necrosis Factor alpha
TPOR	Thiol-protein oxidoreductases
TXN	Thioredoxine
WbMIF	MIF de <i>Wuchereria bancrofti</i>
XIAP	X-linked inhibitor of apoptosis

Listes des figures et des tableaux

Figure 1 – Structures et Séquences des protéines MIFs et D-DT	20
Figure 2 – Schéma des voies de signalisation des protéines MIFs	22
Figure 3 – Les différents types de mort cellulaire programmée	27
Figure 4 – Schéma simplifié des voies de signalisation de l'apoptose	30
Figure 5 – Conservation des voies de signalisation de l'apoptose chez les vertébrés et les invertébrés	31
Figure 6 – Schéma simplifié de la voie de signalisation de la parthanatose	33
Figure 7 – Schéma récapitulatif des étapes de reconstructions des phylogénies	45
Figure 8 – Schéma récapitulatif des différentes analyses statistiques	50
Figure 9 – Protocole pour l'expression des gènes MIF en cellule végétale	54
Figure 10 – Arbre de vie simplifié des eucaryotes	62
Figure 11 – Analyse multiple des correspondances	64
Figure 12 – Analyse discriminante de la répartition du nombre de protéines MIFs	65
Figure 13 – Boxplot de la répartition du nombre de MIFs	66
Figure 14 – Reconstruction phylogénétique des MIFs	68
Figure 15 – Reconstruction phylogénétique des MIFs de Viridiplantae	70
Figure 16 – Reconstruction phylogénétique des MIFs de champignons	72
Figure 17 – Reconstruction phylogénétique des espèces de nématodes et des MIFs	74
Figure 18 – Reconstruction phylogénétique des espèces d'insectes et des MIFs	76
Figure 19 – Reconstruction phylogénétique des espèces et des MIFs d'Alvéolates et Straménopiles	78
Figure 20 – Analyse des sites sous sélection positive ou négative	81
Figure 21 – Phylogénie des protéines MIFs testées pour leur rôle dans la modulation de la mort cellulaire	86
Figure 22 – Mesure d'induction de la mort cellulaire par les protéines MIFs en cellule végétale	87
Figure 23 – Inhibition de la mort cellulaire des cellules végétales par les protéines MIFs	88
Figure 24 – Localisation subcellulaire des protéines MIFs en cellule végétale	90
Figure 25 – Co-localisation des protéines MIFs avec un marqueur du peroxysome	91
Figure 26 – Alignement des séquences protéiques MIFs	94
Figure 27 – Structure prédites des protéines MIFs	95
Figure 28 – Analyse de la zone C-terminale des protéines MIFs	97
Figure 29 – Liaisons hydrogènes prédites des protéines MIFs	99
Figure 30 – Modèle évolutif simplifié montrant les événements de perte et de duplications chez certains phyla d'eucaryotes	112
Figure 31 – Résumé des analyses de la mort cellulaire chez les MIFs	114
Figure 32 – Phylogénie (Maximum de vraisemblance) des Straménopiles et des Alvéolates	142
Figure 33 – Phylogénie (Maximum de vraisemblance) des nématodes (A) et des insectes (B)	143
Tableau 1 – Tableau des bases de données utilisées	47
Tableau 2 – Tableau des logiciels de prédiction de modifications post-traductionnelles	52
Tableau 3 – Tableau récapitulatif des extrémités Gateway®	55
Tableau 4 – Phosphorylations prédites chez les protéines MIFs	101
Tableau 5 – Analyse de la glycosylation des protéines MIFs	103
Tableau 6 – Tableau d'analyse des O-N-acétylglucosaminylation	105
Tableau 7 – Prédictions des acétylations des protéines MIFs	107
Tableau 8 – Listes des espèces eucaryotes utilisées pour l'étude	177

Partie I : Introduction Générale

Tous les organismes vivants sont en interaction fine avec leur milieu et ont développé des systèmes de défense efficaces pour lutter contre les infections par des pathogènes ou des parasites.

La première ligne de défense face aux agents pathogènes est constituée de barrières physiques et biochimiques telles que la peau, les muqueuses, ou les larmes chez l'homme, ou les cires épicuticulaires chez les plantes. Une fois ces barrières franchies par un pathogène, la réponse immunitaire se met en place. Elle débute généralement par une reconnaissance de motifs conservés présents à la surface des pathogènes (PAMP pour « Pathogen Associated Molecular Patterns »). Cette reconnaissance se fait grâce à des récepteurs (PRR pour « Pattern Recognition Receptors »), souvent exprimés par des cellules de défense circulantes et parfois solubles, pour les organismes animaux ou exprimés par toutes les cellules chez les plantes [1–4]. Cette reconnaissance de motifs du « non-soi » déclenche la réponse immunitaire innée.

L'immunité innée existe chez tous les organismes vivants et peut s'appuyer sur divers processus cellulaires et moléculaires, en fonction du règne ou du phylum de l'organisme [1, 5, 6].

L'immunité adaptative, elle, n'existe que chez les métazoaires vertébrés. Elle est caractérisée par des propriétés de spécificité (réponse très spécifique vis-à-vis de l'agent pathogène) et de mémoire, qui reposent sur l'existence de lignées cellulaires spécifiques (lymphocytes B, T) et la synthèse d'immunoglobulines adaptées [7].

Les cytokines sont des protéines immunitaires particulières qui jouent un rôle central dans « l'orchestration » à la fois de la réponse innée, et de la réponse adaptative chez les vertébrés. Elles sont sécrétées par certains types cellulaires de l'immunité et agissent sur plusieurs autres types cellulaires. Chez les vertébrés, elles sont indispensables à la réaction inflammatoire et ont des rôles très pléiotropes. Elles sont généralement séparées en trois catégories en fonction de leur action prépondérante : i) les cytokines proinflammatoires, ii) les cytokines immunorégulatrices et iii) les cytokines effectrices. Elles peuvent agir sur les cellules productrices (action autocrine) ou sur des cellules à distance (action endocrine). Les cytokines régulent de nombreux mécanismes tels que l'inflammation, la fièvre, ou la formation des cellules du système immunitaire. Certaines cytokines proinflammatoires sont aussi impliquées dans l'activation de l'immunité adaptative.

Les cytokines MIFs (Macrophage Migration Inhibitory Factor) font partie de ces cytokines participant à l'immunité innée et nécessaires à la mise en place de la réponse adaptative. Identifiées depuis les années 1960, elles ont tout d'abord été décrites comme ayant la capacité à inhiber la migration cellulaire des macrophages [8, 9]. Depuis, les découvertes sur leurs rôles dans l'immunité humaine et dans de nombreuses pathologies (cancers, maladies auto-immunes, maladies infectieuses, etc..) ne font qu'augmenter et le nombre de publications à leur sujet dépasse actuellement les 5000 (WOS, 2018).

Parallèlement à ces nombreux travaux sur les MIFs humaines ou de mammifères, seules quelques rares études ont été réalisées sur des MIFs d'autres espèces, et ont toutes porté sur des parasites de vertébrés. Ainsi, des MIFs ont été identifiées chez des nématodes tels que *Brugia malayi*, *Ancylostoma caninum*, ou *Wuchereria bancrofti* [10, 11] et chez des tiques comme *Dermacentor variabilis* [12, 13]. Chez ces parasites, bien que les modes d'action ne soient pas élucidés, il semble que les MIFs soient sécrétées et participent à la modulation de l'immunité de l'hôte vertébrés [10, 14, 15].

Le rôle de ces protéines MIFs chez d'autres organismes que l'homme ou les parasites humains n'est exploré que depuis peu. Ainsi, une première étude fonctionnelle a démontré que BgMIF1 du mollusque *Biomphalaria glabrata* présente certaines activités biologiques majeures du MIF humain, et participe à la réponse immunitaire après infection du mollusque par le parasite *Schistosoma mansoni* [16]. Différentes études ont ensuite identifié des MIFs chez d'autres invertébrés libres comme d'autres espèces de mollusques, des crustacés ou des échinodermes [17–19]. De façon inattendue, un travail de notre équipe a récemment montré que des parasites de plantes comme les pucerons du pois et du pêcher (*Acyrtosiphon pisum* et *Myzus persicae*) injectent également des MIFs dans leur plante hôte. Ces MIFs de pucerons modulent la réponse immunitaire de la plante et sont indispensables à l'établissement de l'infection [20]. Parallèlement, une étude a décrit la présence de cette protéine chez de très nombreuses plantes [21], mais sa fonction n'a pas été étudiée à ce jour.

Ainsi, les protéines MIFs peuvent non seulement jouer un rôle majeur dans l'immunité d'hôtes aussi différents que les plantes, les mammifères et les invertébrés mais également être utilisées par des organismes parasites pour inhiber les défenses de leurs hôtes respectifs.

Partant de ce constat, l'hypothèse principale de ce travail était que des pressions de sélection différentes aient pu s'exercer sur les MIFs impliquées uniquement dans l'immunité innée des organismes et les MIFs impliquées dans la modulation de l'immunité des hôtes. Ces pressions de sélection différentielles auraient pu conduire à une histoire évolutive particulière (différente de celle de l'arbre des espèces) entre les différentes espèces en fonction de leur mode de vie et à des différences fonctionnelles entre ces protéines MIFs.

L'objectif global de ce travail de thèse était donc d'explorer l'histoire évolutive des cytokines MIFs à une large échelle (trans-règnes), et leurs potentielles différences fonctionnelles, avec un intérêt particulier pour les systèmes hôtes-parasites.

La première partie de ce manuscrit présente le contexte bibliographique de l'étude, en se focalisant sur les fonctions protéiques décrites d'abord chez l'homme puis chez les autres espèces et en rappelant les différentes études sur la diversité des MIFs chez les eucaryotes. Après une présentation des objectifs (partie II), la troisième partie détaille l'ensemble des techniques utilisées pour la réalisation de cette étude. La quatrième partie expose les résultats et discussions et enfin, une conclusion générale et les perspectives issues de ces travaux sont présentées en cinquième partie.

Partie II : Contexte bibliographique

Chapitre 1 : Les cytokines MIFs chez l'Homme

II.1.1- Généralités

II.1.1.1- Historique

En 1966, la protéine MIF pour « Macrophage Migration Inhibitory Factor » est l'une des premières cytokines identifiées [22]. Elle a tout d'abord été décrite chez l'homme comme étant capable d'inhiber la migration des macrophages *in vitro* [8]. Depuis son clonage en 1989 [23], MIF est peu à peu apparue comme étant une cytokine/ chemokine jouant un rôle central en tant que régulateur de l'immunité innée mais aussi en tant que protéine capable de lier le système endocrinien et le système immunitaire [24] au point d'être nommée « Most Interesting Factor » en 2000 [25].

La protéine MIF a longtemps été considérée comme la seule protéine de cette famille présente chez l'humain. En 1998, la D-dopachrome tautomérase (D-DT) aussi appelée « MIF 2 » a été décrite comme ayant une structure et une séquence similaire à celle des MIFs [26]. Celle-ci a dans un premier temps (1993) été décrite comme une enzyme qui convertit le D-dopachrome en acide 5,6-dihydroxyindole-2-carboxylique dans les mélanomes humains [27]. Elle est maintenant décrite comme une protéine capable de réaliser la majorité des actions du MIF humain. Dans ce manuscrit de thèse, le terme MIF sera utilisé pour désigner l'ensemble des membres de cette famille multigénique, incluant les MIFs et les D-DT.

II.1.1.2- Structure des gènes et des protéines

Chez l'être humain, les deux gènes codant pour les protéines MIFs et D-DT sont localisés sur le chromosome 22 à seulement 80 kb l'un de l'autre. Les deux gènes ont une structure très similaire et sont composés de 3 exons de taille quasi identique. De plus, le taux de conservation de leur ARNm est plutôt élevé (environ 40%) [28]. Les deux ARNm codent pour des protéines de petite taille (115 acides aminés pour 12.5 kDa).

Ces deux gènes sont aussi présents chez la souris (*Mus musculus*) sur le chromosome 10. Leur localisation est très conservée, ce qui indique une conservation de la synténie chez l'humain et les souris [29, 30]. La structure, et le taux d'identité de leur ARNm sont similaires à celles observées chez l'homme, ce qui suggère que ces deux gènes sont issus d'une ancienne duplication [24].

Les protéines MIFs et D-DT sont peu conservées : ces deux protéines ne présentent que 34% d'identité de séquence chez l'homme et 27% chez la souris. Toutefois, la structure de ces deux protéines reste similaire : leurs structures (MIF et D-DT) ont été déterminées par cristallographie respectivement en 1996 et 1998 [26, 31] puis confirmées expérimentalement comme étant préférentiellement trimériques chez l'homme quelques années plus tard [32]. La structure quaternaire des MIF consiste donc en 3 monomères composés chacun de 6 feuillets β et de 2 hélices α formant le motif $\beta\alpha\beta\beta\alpha\beta$.

Les protéines MIFs ne disposent pas de domaine intra-protéique particulier. Seuls certains motifs et des acides aminés ont été décrits comme étant caractéristiques d’une des fonctions des MIFs. Ceux – ci seront détaillés dans le paragraphe correspondant à chaque fonction.

Les protéines MIFs ont été découvertes chez un grand nombre d’espèces eucaryotes (voir Chapitre 3 :). La cristallographie des MIFs de protistes (*Plasmodium falciparum* et *berghei* [33], *Giardia lamblia* [34], ...) et des nématodes (*Ancylostoma ceylanicum* [14], *Trichinella spiralis* [35], *Brugia malayi* [36]) révèle que les protéines possèdent toutes la même structure, malgré un faible taux de conservation en acides aminés entre les séquences. Cette protéine se retrouve aussi chez des bactéries telles que la cyanobactérie *Prochlorococcus marinus* [37] et présente une structure similaire chez cette espèce.

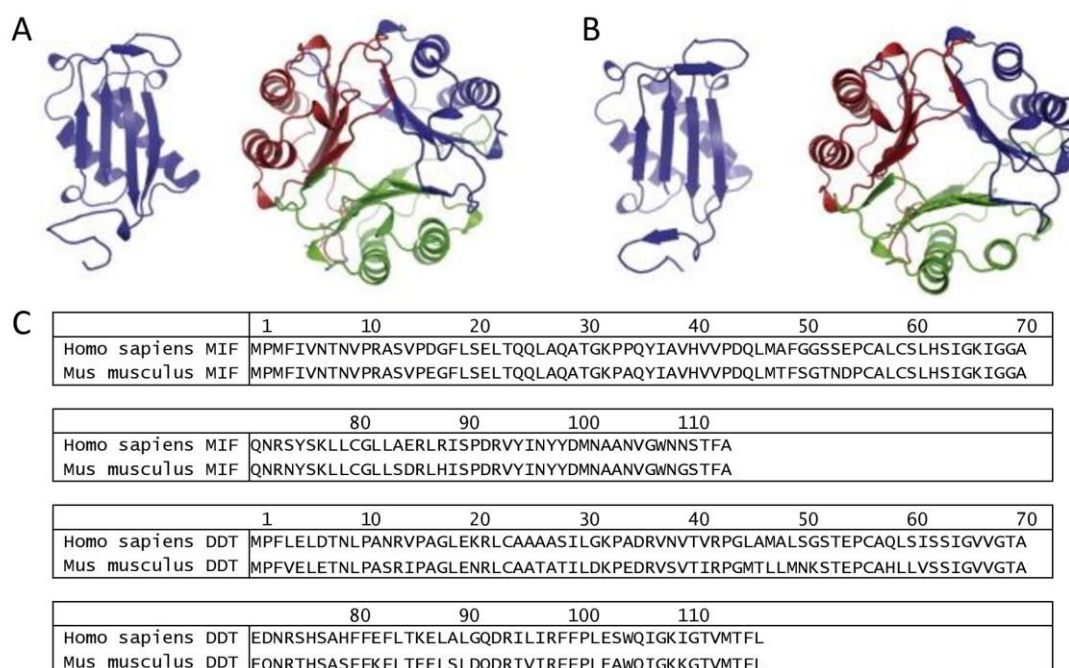


Figure 1 - Structures et Séquences des protéines MIFs et D-DT

Les protéines MIFs (A) et D-DT (B) forment des homotrimères. Les protéines MIFs chez l’homme et la souris (C) ont des séquences très proches (90% d’identité pour les protéines MIF et 73% pour les protéines D-DT) – adaptée de Merk et Mitchell [38].

II.1.1.3- Localisation et expression des MIFs

Les MIFs sont majoritairement localisées dans le cytoplasme des cellules mais elles peuvent aussi être relocalisées dans le noyau. Bien qu’elles ne possèdent pas de peptide signal, ces protéines peuvent être sécrétées grâce à un mécanisme de sécrétion non – classique assez méconnu utilisant vraisemblablement les transporteurs ABC [39]. Elles peuvent agir à l’intérieur de leur cellule de production (action autocrine), sur les cellules proches (action paracrine), ou encore sur des cellules très éloignées (action endocrine) [24]. MIF est l’une des premières protéines à avoir été décrite comme capable d’agir à la fois sur les cellules de l’organisme de production mais aussi sur les cellules d’un autre organisme, notamment en cas d’infection ou de parasitisme [15, 40].

Les MIFs sont produites et sécrétées par un large panel de cellules impliquées dans la réponse immunitaire (les lymphocytes T, les monocytes, les macrophages, ...) ainsi que par un grand nombre d'organes et de tissus en contact avec le milieu extérieur tels que la gorge, le tractus génital, les yeux, le foie, les reins ou encore la paroi intestinale [41]. De plus, les MIFs sont fortement exprimées par la majorité des tissus du système endocrine, notamment l'hypothalamus et les glandes surrénales et pituitaires. Les tissus exprimant la protéine MIF produisent également la protéine D-DT [38].

Les protéines MIFs sont produites et sécrétées par les cellules de façon constitutive. Contrairement à l'immense majorité des cytokines, ces protéines sont aussi stockées dans les cellules sous la forme de pools de protéines préformées. Ainsi, pour une sécrétion rapide des protéines MIFs par les cellules, la synthèse de protéines *de novo* n'est pas nécessaire [24].

II.1.1.4- Les récepteurs des MIFs

Les MIFs peuvent agir sur les cellules en se fixant à différents récepteurs membranaires mais aussi en agissant directement sur leur cible suite à une endocytose. Le principal récepteur des protéines MIFs est le récepteur CD74 [24, 42]. Ce récepteur est aussi connu sous le nom de la chaîne gamma (ou fragment li) du complexe majeur d'histocompatibilité de classe II (CMH-II). Cette chaîne du CMH-II est considérée comme la chaîne invariante de ce complexe. Le CMH de classe II est nécessaire à la reconnaissance des antigènes lors de l'activation des lymphocytes T helper. Ce processus est l'un des processus majeurs de l'immunité acquise chez les êtres humains. Le récepteur CD74 module la présentation des antigènes au CMH-II [43].

Les protéines MIFs peuvent aussi se fixer sur les récepteurs CXCR2 [44, 45], CXCR4 [46] et CXCR7 [41, 47] (Figure 2).

Le récepteur CXCR2, aussi appelé IL8RB, pour récepteur bêta à l'interleukine 8, se lie à l'interleukine 8 avec une forte affinité. Ce récepteur est aussi capable de se lier aux chimiokines de type C-X-C 1, 2, 3 et 5 (CXCL1, 2, 3 et 5). La fixation de ces différentes chimiokines aux récepteurs CXCR2 est un processus de régulation de la fixation des protéines MIFs à leurs récepteurs [41].

Le récepteur CXCR7, aussi appelé ACKR3 (atypical chemokine receptor 3), se fixe aux chimiokines CXCL11 et CXCL12 [48]. La surexpression de CXCR7 a été démontré comme étant corrélée avec une augmentation de la progression des tumeurs. Le récepteur CXCR7 peut former des hétéro - oligomères avec le récepteur CXCR4 [49]. Les voies activées par la fixation des protéines MIFs sur les hétéro - oligomères CXCR7 et CXCR4 sont différentes des voies activées par la fixation des protéines MIFs sur les homo-oligomères CXCR4 et CXCR7 [41].

II.1.1.5- Les voies de signalisation des MIFs

Les protéines MIFs agissent sur de nombreuses voies de signalisation, notamment sur l'activation des MAP kinases ERK1 et 2 [50]. Elles agissent aussi sur la régulation de voies d'activation de protéines pro inflammatoires (TNF- α , IL-1 β , ...), la surexpression de TLR4 [51, 52], la surexpression de l'activité des p53 [53, 54] ou encore la régulation du cycle cellulaire. Au total, en 2016, les protéines MIFs ont été considérées comme interagissant directement avec 68 protéines différentes et induisant la sur - expression de plus de 50 gènes différents [41].

Les protéines MIFs agissent dans le milieu extracellulaire des cellules, dans le cytoplasme et dans le noyau des cellules humaines. Elles peuvent passer d'un compartiment cellulaire à l'autre. Via la fixation à la thioredoxine (TXN), les protéines MIFs peuvent être endocytées et ainsi passer directement du milieu extracellulaire au cytoplasme. Ce processus est l'un des principaux processus de régulation des protéines MIFs. La compétition des protéines MIFs pour la fixation aux récepteurs jouent aussi un rôle dans la régulation des actions des protéines.

Les protéines MIFs agissent sur deux régulateurs clés du fonctionnement cellulaire : la protéine CD44 et les MAP kinases 1 et 2. Ces protéines sont impliquées dans la régulation de la glycolyse, l'angiogénèse, la production de prostaglandine PG2 ou encore l'ostéoclastogénèse.

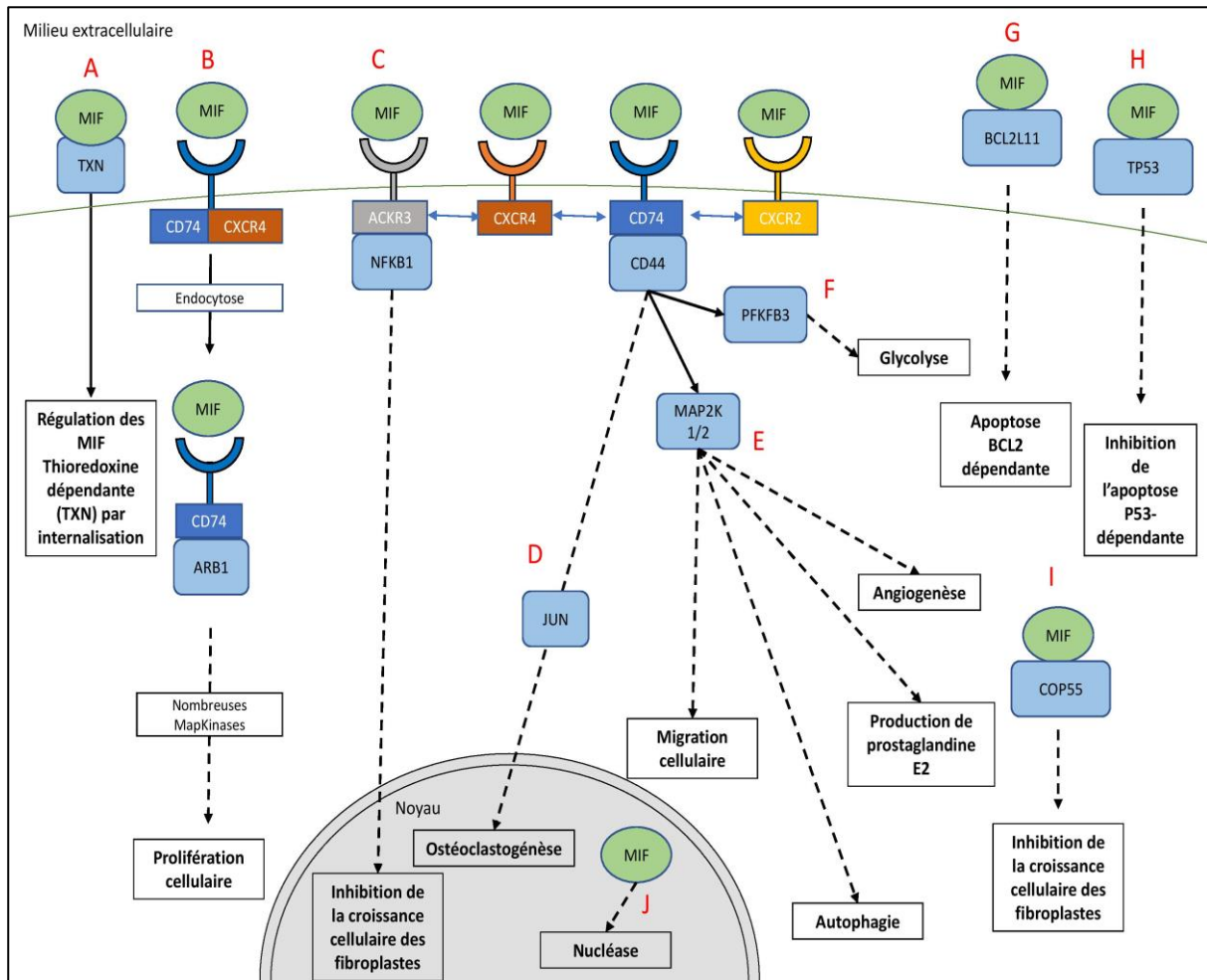


Figure 2 – Schéma des voies de signalisation des protéines MIFs

Schéma simplifié des voies d'actions des protéines MIFs, adapté de Subbannayya et al. [41]. A) phénomène d'endocytose des protéines MIFs par la thioredoxine. B) et C) phénomène d'hétéro-oligomérisation des récepteurs CXCR4, CD74 et CXCR2 ainsi qu'une partie des différentes voies activées par la fixation des protéines MIFs sur ces récepteurs. D), E), F), une partie des voies activées par les régulateurs clés CD44 et les MAPK1 et 2. G), H), I) et J) Autres actions des protéines MIFs.

II.1.2- Fonctions des MIFs chez l'être humain

Chez l'homme, les MIFs sont considérées à la fois comme des cytokines majeures de l'immunité et comme des hormones. En tant que cytokines, elles jouent un rôle majeur à la fois dans la régulation de la prolifération cellulaire et dans l'inhibition de l'apoptose par différentes voies de signalisation (figure 2). Via leur rôle dans la production de prostaglandine E2, elles jouent aussi un rôle dans de nombreux processus d'inflammation des tissus suite à une infection. Elles possèdent un rôle majeur dans l'immunité et l'inflammation.

De nombreuses pathologies humaines sont associées à une augmentation de l'expression des MIFs comme par exemple le diabète de type 2 [55], certains cancers [56] mais aussi de nombreuses maladies inflammatoires comme l'arthrite rhumatoïde [57] ou la maladie de Crohn [58, 59].

Pour des raisons de simplicité, je ne détaillerai par la suite que quelques-unes des fonctions des protéines MIFs en détaillant particulièrement les fonctions liées à l'immunité. Ces fonctions ont pour la plupart uniquement été décrites pour MIF1 chez l'humain et seules quelques-unes d'entre elles ont aussi été vérifiées pour la protéine D-DT (ou MIF2).

II.1.2.1- Fonctions des MIFs dans l'immunité

a) Rôles des MIF dans l'interaction avec les macrophages

La première fonction qui a été décrite pour les MIFs est l'inhibition de la migration aléatoire des macrophages [8, 9]. Elles sont impliquées dans la production de protéines chemo - attractantes telles que MCP-1 (monocyte chemoattractant protein 1) qui permet le recrutement de macrophages sur le site de l'infection [60].

Depuis, l'importance des MIFs pour de nombreuses fonctions des macrophages a été démontrée. L'expression des MIFs est nécessaire à de nombreuses fonctions de ces cellules telles que l'adhérence, la phagocytose, ou la réduction de l'induction de l'apoptose *via* la voie de p53. L'ensemble de ces fonctions passent par la voie des MAP kinase ERK1/2 [50, 61]. Cette voie dépend entièrement de la présence des récepteurs CD74/CD44.

b) Rôles des MIFs dans l'inflammation

Les MIFs ont un rôle prépondérant dans la régulation de l'inflammation des tissus face à une infection. Les protéines MIFs sont essentielles à la production des cytokines pro - inflammatoires telles que le tumor necrosis factor (TNF- α) ou l'interleukine 1 (IL-1 β) par les macrophages. Elles sont aussi impliquées dans la production de molécules liées à l'inflammation telles que la prostaglandine E2 (PGE₂) ou le monoxyde d'azote [54]. Enfin, les protéines MIFs ont la capacité de réguler et d'inhiber les actions immunosuppressives des glucocorticoïdes telle que la cortisone [62].

c) Rôles des MIFs dans l'induction de cancers

Les protéines MIFs sont impliquées dans la propagation et le développement de nombreux cancers. En effet, l'inhibition des MIFs ou leur délétion sont associées à une réduction de la taille de nombreuses tumeurs. Cette réduction est due au fait que les protéines MIFs agissent sur l'angiogenèse, sur la voie de l'apoptose P53 – dépendante et sur l'inhibition de cycle cellulaire. Les protéines MIFs et DDT ont toutes deux un effet similaire sur les cellules malignes. Les MIFs agissent sur les cellules tumorales *via* la fixation au récepteur CD74 et dans une moindre mesure la fixation au récepteur CXCR4.

d) Rôles des MIF dans l'interaction avec les bactéries

Les protéines MIFs ont un rôle important dans la réaction de l'organisme face à une infection bactérienne, que ce soit des bactéries Gram positives ou Gram négatives [24].

Pour les infections par des bactéries Gram négatives, les protéines MIFs agissent notamment *via* la voie des TLR4. Ce récepteur permet la reconnaissance des lipopolysaccharides bactériens (LPS). Les cellules humaines sous-exprimant les MIFs ont, par conséquent, une réponse très faible au LPS et aux bactéries Gram négatives. Ainsi, les souris déficientes en MIF ne peuvent pas endiguer une infection induite par *Salmonella typhimurium* [63]. Par contre, les mêmes souris sont plus résistantes au choc septique, induit par une réaction excessive au LPS présent à la surface des bactéries.

Pour les infections par des bactéries Gram positives, de la même manière, les protéines MIFs sont nécessaires à la reconnaissance des bactéries. Les anticorps anti-MIF sont ainsi associés à une augmentation de la survie suite à un choc septique ayant lieu lors d'une infection induite par *Streptococcus pneumoniae*, un des agents pathogènes responsables de la pneumonie [64, 65].

e) Rôle des MIFs dans l'immunité adaptative

Le rôle des protéines MIFs dans l'immunité adaptative est pour l'instant beaucoup moins connu et étudié que celui des protéines MIFs dans l'immunité innée. Il a, toutefois, été démontré que les protéines MIFs jouent un rôle essentiel dans l'activation et la prolifération des lymphocytes T et dans la production d'interleukine 2 [66]. Elles sont, de plus, impliquées dans l'inhibition des lymphocytes CD8⁺ cytotoxiques et dans la régulation du trafic des lymphocytes [67].

II.1.2.2- Les activités enzymatiques des MIFs

En plus de leurs nombreuses fonctions, les MIFs possèdent 3 activités enzymatiques.

a) L'activité tautomérase

La protéine MIF possède une activité kéto-énol tautomérase [27, 68]. Il s'agit d'une des premières fonctions découvertes chez la protéine D-DT [27]. L'activité kéto-énol tautomérase permet la transformation de certains substrats artificiels tels que le D-dopachrome, l'hydroxyphénylpyruvate, ou encore le phénylpyruvate. L'activité tautomérase du L-dopachrome est connue comme étant un des mécanismes de défense chez les insectes. Il semblerait notamment être lié à l'encapsulation mélanique. Toutefois, les protéines MIFs ne tautomérisent que le substrat artificiel D-dopachrome (stéréoisomère) et n'agissent pas sur le substrat naturel présent chez les insectes, le L-dopachrome. Son rôle exact, particulièrement chez les vertébrés, reste donc inconnu. Il a été démontré que les acides aminés Pro-2, Lys-33, Ile-65, Tyr-96 et Trp-109 sont nécessaires à la protéine MIF pour réaliser l'activité tautomérase [69].

b) L'activité oxido-réductase

Les protéines MIFs ont une activité oxido-réductase via un motif Cys – Xaa – Xaa - Cys (aussi appelé CXXC) [70–72]. Ce site est caractéristique de la superfamille des thioredoxines, plus particulièrement des « thiol-protein oxidoreductase » (TPOR), c'est-à-dire des protéines thiols capables d'activité oxydoréductase. Cette activité n'est pas présente chez les protéines D-DT due à l'absence de la seconde cystéine [72]. Toutes les fonctions liées à cette activité enzymatique sont donc absentes chez les protéines D-DT. L'activité oxido-réductase joue un rôle majeur dans l'immunité acquise et innée ainsi que dans la régulation de la balance réduction – oxydation cellulaire, de l'inhibition de l'apoptose, l'activation des macrophages et enfin un rôle modulateur dans la fixation des MIFs à leurs récepteurs [72].

c) L'activité nucléase

Il a récemment été démontré que les MIFs possèdent une activité 3' exonucléase et endonucléase [73]. Les domaines majeurs nécessaires à cette activité sont les domaines PD-D/E(X)K présents en position 17 à 22, 45 à 55 et 93 à 101 chez le MIF humain. La protéine MIF possède de plus un domaine zinc finger (CxxCxxHx(n)C), qui est couramment décrit comme important chez les protéines impliquées dans la réponse aux dommages à l'ADN. Cette activité des protéines MIFs joue un rôle majeur dans la mise en place de la parthanatose, une forme de mort cellulaire (voir chapitre 2).

II.1.2.3- Autres rôles des MIFs

L'immunité n'est qu'une des facettes des activités des MIF. Directement ou indirectement, les protéines MIFs sont aussi impliquées dans de nombreuses maladies liés à une dérégulation de l'homéostasie du glucose (diabète de type 1 et 2) ou une dérégulation de la croissance cellulaire (cancers). Ici, je ne présenterai que quelques unes des fonctions des MIFs qui ne sont pas liées directement à l'immunité.

a) Rôles des MIFs dans l'homéostasie du glucose

Les protéines MIFs jouent un rôle majeur dans la régulation de l'homéostasie du glucose [74]. Elles sont produites par les cellules productrices d'insuline (cellules Béta du pancréas) et par les cellules réceptrices de l'insuline (adipocytes, myocytes et cardiomyocytes) [75–80]. Ainsi, elles sont impliquées dans la production d'insuline par les cellules Béta du pancréas mais peuvent aussi diminuer la transduction des signaux induits par l'insuline dans les adipocytes. Elles sont aussi impliquées dans la modulation de l'absorption du glucose, dans la glycolyse et dans les phénomènes de résistance à l'insuline dans les adipocytes.

Les différentes actions liées à l'homéostasie du glucose passent toutes par la fixation des protéines MIFs aux complexes formés du récepteur CD74 et du récepteur CD44. L'activation de ce complexe de récepteurs entraîne notamment l'activation de l'enzyme 6-phosphofructo-2-kinase (PFK-2) (Figure 2), qui est associée à la synthèse de fructose 2,6 biphosphate, un des régulateurs majeurs de la glycolyse [80].

Il y a un lien très marqué entre le rôle des protéines MIFs dans la régulation de l'immunité et leur rôle dans la régulation du métabolisme du glucose. Ainsi, les diabètes de type 1 et 2 sont couramment associés à des inflammations chroniques des tissus et des réactions immunitaires excessives [75]. Les multiples actions des protéines MIFs expliquent pour partie ce lien entre la régulation de l'immunité et le métabolisme du glucose.

b) Rôles des MIFs dans l'angiogénèse

Les protéines MIFs sont capables d'induire la production de nouveaux vaisseaux sanguins dans certaines situations de stress ou d'hypoxie. Cette activité joue un rôle primordial dans le développement de certains cancers [81]. En effet, le contournement de l'hypoxie lié à l'excès de cellules est un mécanisme nécessaire pour permettre l'expansion d'un cancer [82].

Chapitre 2 : La mort cellulaire et les protéines MIFs

Les protéines MIFs sont décrites chez l'homme comme jouant un rôle dans la modulation de la mort cellulaire. Ce chapitre présente tout d'abord les différents types de mort cellulaire décrites à ce jour (en particuliers ceux affectés par les MIFs). Il évoque ensuite le rôle des protéines MIFs dans la modulation de ces différents processus.

Chez les eucaryotes, il existe plusieurs types de mort cellulaire programmée (Programmed Cell Death, PCD) [83]. Un schéma représentatif des caractéristiques majeures de ces différentes formes de mort cellulaire est présenté ci-dessous (Figure 3).

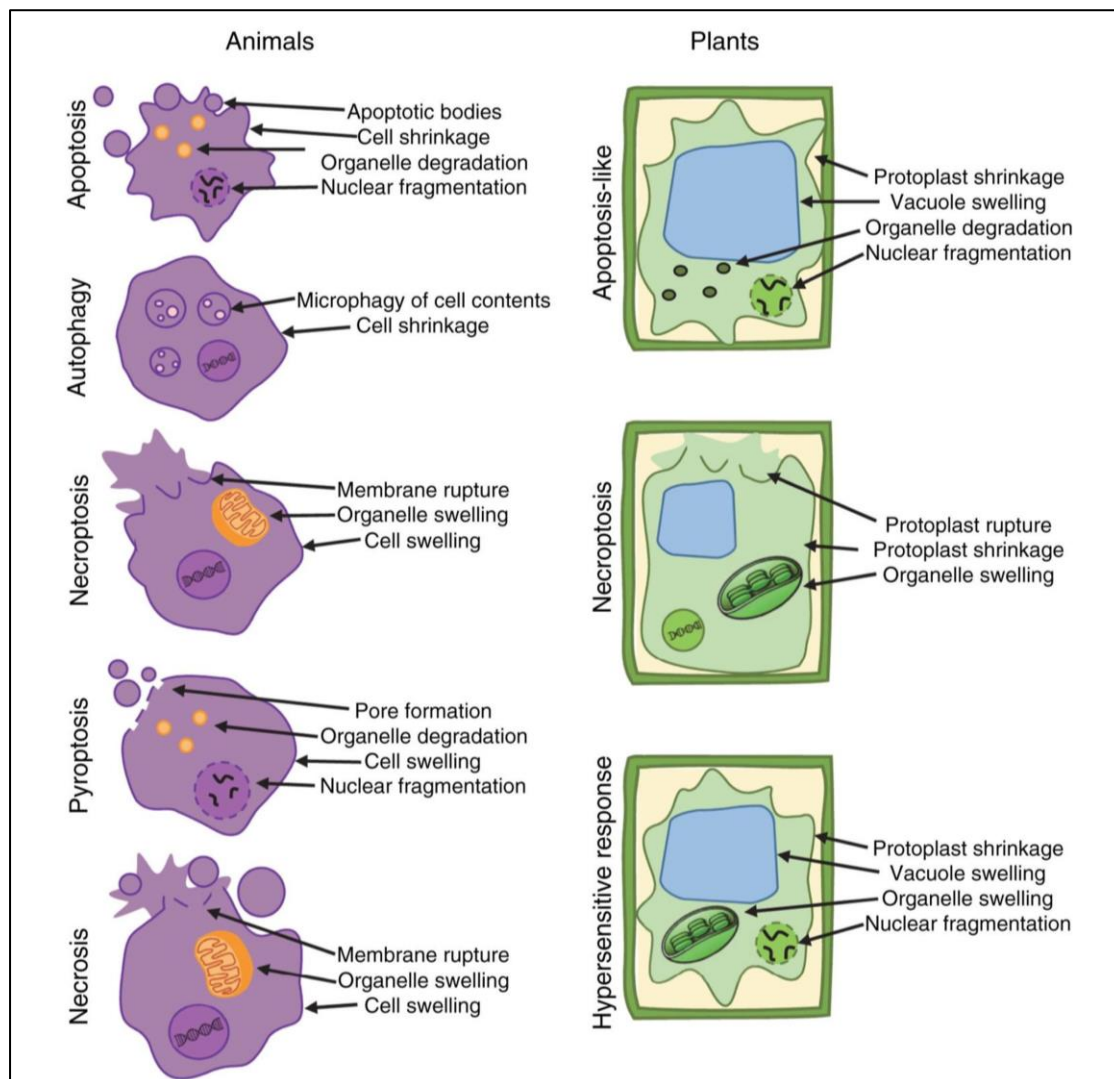


Figure 3 – Les différents types de mort cellulaire programmée

Schéma récapitulatif des différents types de mort cellulaire programmée connus chez les vertébrés et les plantes (adaptée de Mukhtar et al, [84]).

Contrairement aux processus de mort cellulaire des animaux, ceux des plantes restent, à ce jour, largement inconnus. La suite de ce chapitre se concentrera donc sur la mort cellulaire dans les cellules animales.

II.2.1- La PCD de type I ou l'apoptose

II.2.1.1- Rappel sur l'apoptose

La PCD de type I, couramment appelée l'apoptose, est la forme la plus documentée. Elle est essentielle pour permettre le développement normal des animaux et le maintien de l'homéostasie des tissus. Il est estimé qu'environ 10 millions de cellules mourraient chaque jour par apoptose [85] lors du développement normal d'un être humain. L'apoptose est aussi nécessaire pour permettre l'élimination des cellules infectées par un pathogène (virus) et des cellules agressives envers les cellules du soi. La dérégulation de l'apoptose est une composante majeure de nombreuses maladies telles que des cancers, la prolifération de lymphocytes lors d'une maladie auto-immune, la maladie de Parkinson, de Huntington ou d'Alzheimer.

L'apoptose se caractérise par un ensemble de transformations ayant lieu dans la cellule, comprenant une réduction du volume cellulaire (ou pycnose), une condensation de la chromatine, la fragmentation du noyau et la formation de corps apoptotiques à partir de la membrane plasmique des cellules (membrane blebbing)[86]. Cette forme de mort cellulaire implique l'activation de nombreuses caspases [87], qui sont responsables de l'ensemble des changements morphologiques caractéristiques de l'apoptose.

Chez les êtres humains, deux voies de signalisation majeures permettent l'activation de l'apoptose dans les cellules : une signalisation externe passant par des récepteurs dit « death receptor » (ou DR) et une voie de signalisation intrinsèque, dite mitochondriale. Les récepteurs DR – comme, par exemple, le récepteur Fas (ou CD95), le récepteur TRAIL ou le récepteur TNFR-1 induisent l'apoptose par le recrutement direct des caspases. Ces différents récepteurs recrutent principalement la caspase 8. La voie de signalisation intrasèque est, elle, contrôlée par les protéines de la famille des Bcl2. Ces deux voies de signalisation se terminent toutes deux par l'activation de nombreuses caspases. Ces caspases activent de nombreuses protéines qui vont notamment perturber le cytosquelette de la cellule, permettre la fragmentation de l'ADN et inhiber le bon fonctionnement des mitochondries. Ces différentes activités conduisent à l'apoptose [88].

Pour l'apoptose mitochondriale, sous l'action de différents stimulus (dégradation de l'ADN, stress du réticulum endoplasmique, ...), les protéines Bcl2 inductrices de l'apoptose sont activées (Bax, Bak, ...). Elles activent alors la perméabilisation de la membrane extérieure des mitochondries (mitochondrial outer membrane permeabilisation, MOMP). La MOMP permet la libération, à l'extérieur des mitochondries, des protéines SMAC, OMI et des cytochromes c. Les cytochromes c vont se fixer à la caspase 9 et à la protéine APAF-1 formant des apoptosomes. Les apoptosomes vont ensuite activer les différentes caspases nécessaires à la mise en place de l'apoptose. Pendant ce temps, les protéines OMI et SMAC vont inhiber les caspases inhibitrices de l'apoptosome (X-linked inhibitor of apoptosis, XIAP). Les protéines de la famille Bcl2 jouent un rôle crucial dans la régulation de l'apoptose : certaines induisent l'apoptose (Bax et Bak), d'autres bloquent la MOMP (Bcl2, Mcl1, Bcl-xL), et enfin certaines

vont soit activer les protéines inductrices de l'apoptose, soit neutraliser les effets antiapoptiques des autres protéines de la famille des Bcl2 (Figure 4).

II.2.1.2- Rôle des protéines MIFs dans l'apoptose

Les protéines MIFs agissent à plusieurs niveaux sur la modulation de l'apoptose (Figure 4). Elles se fixent à la protéine TP53 qui par la suite active une voie de signalisation, dont les détails exacts ne sont pas encore connus, qui conduira à terme à la dégradation du protéosome et l'inhibition de la voie inductrice de l'apoptose dépendante de p53. La protéine p53 favorise la transcription de gènes dont ceux codant la protéine pro-apoptotique Bax.

Les protéines MIFs peuvent aussi agir plus directement sur l'apoptose. En effet, elles sont capables d'inhiber la formation du complexe protéique formé des protéines BID et BAX. L'inhibition de ce complexe induit la régulation négative de l'expression des gènes codant pour le complexe enzymatique formé par le cytochrome c et la protéine SMAC. Ce complexe ne peut donc pas réguler négativement la production de caspase 3.

Enfin, les protéines MIFs peuvent se fixer aux protéines BCL2L11, aussi appelées BIM. Ces protéines permettent l'induction de l'apoptose en inhibant l'action des protéines anti-apoptotiques telles que MCL-1 ou BCL2.

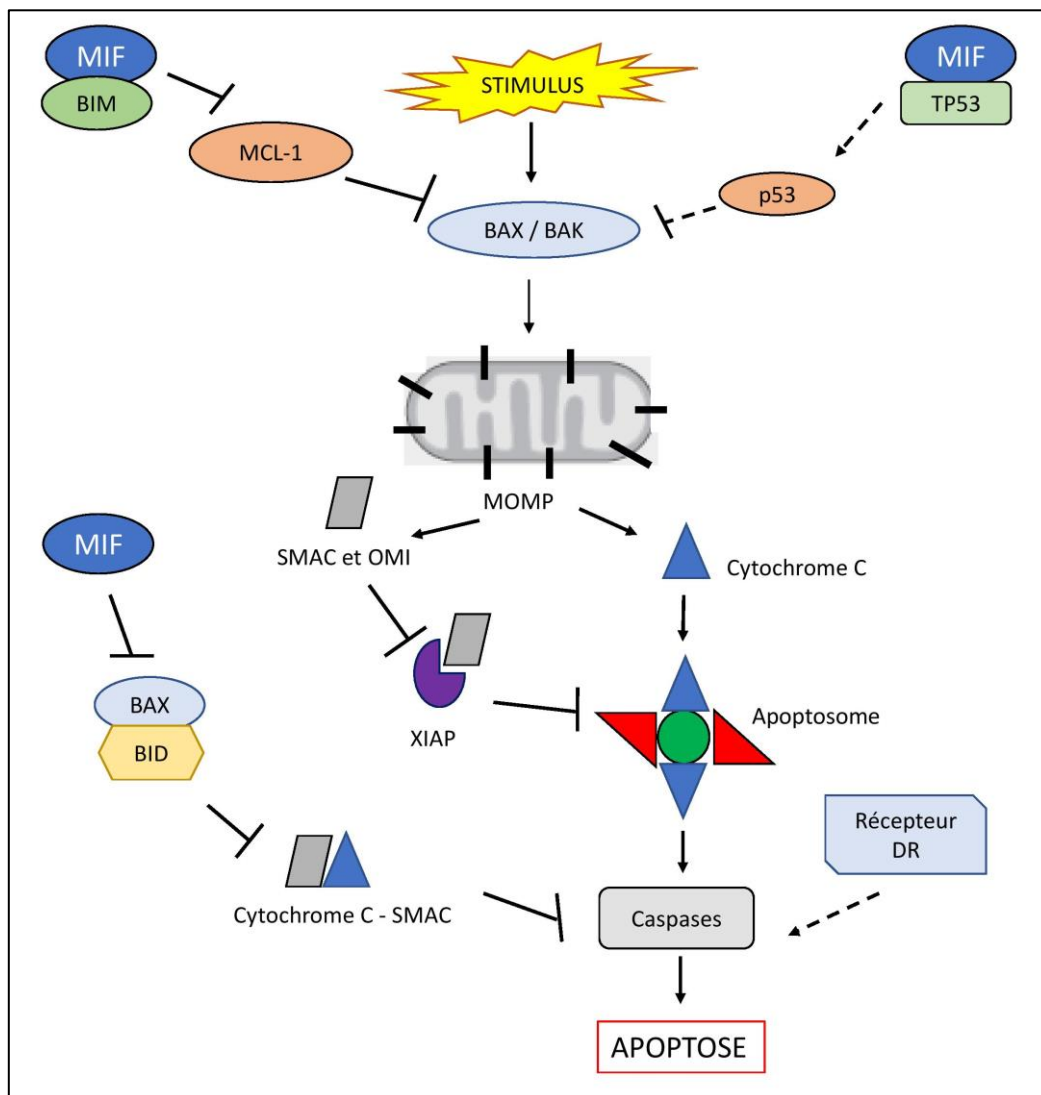


Figure 4 – Schéma simplifié des voies de signalisation de l'apoptose

Schéma des voies de signalisation de l'apoptose – focalisé sur l'apoptose induite par des signaux intrasèques et comprenant le rôle des protéines MIFs (adaptée de Green et Llambi [88] et Subbannayya et al. [41]). Les flèches pleines correspondent à des activations. Les flèches en pointillés correspondent à des activités inductrices ou inhibitrices dont les mécanismes sont encore inconnus. Les flèches se terminant par le symbole -| correspondent à des activités inhibitrices.

II.2.1.3- L'apoptose et les protéines MIFs chez les autres espèces

L'apoptose est un phénomène conservé chez de nombreuses espèces eucaryotes. Toutefois, les voies d'activation et les protéines impliquées diffèrent d'une espèce à l'autre (Figure 5). Les protéines permettant la mise en place de l'apoptose étant différentes, le rôle des protéines MIFs sur l'apoptose pour d'autres espèces que l'être humain reste inconnu.

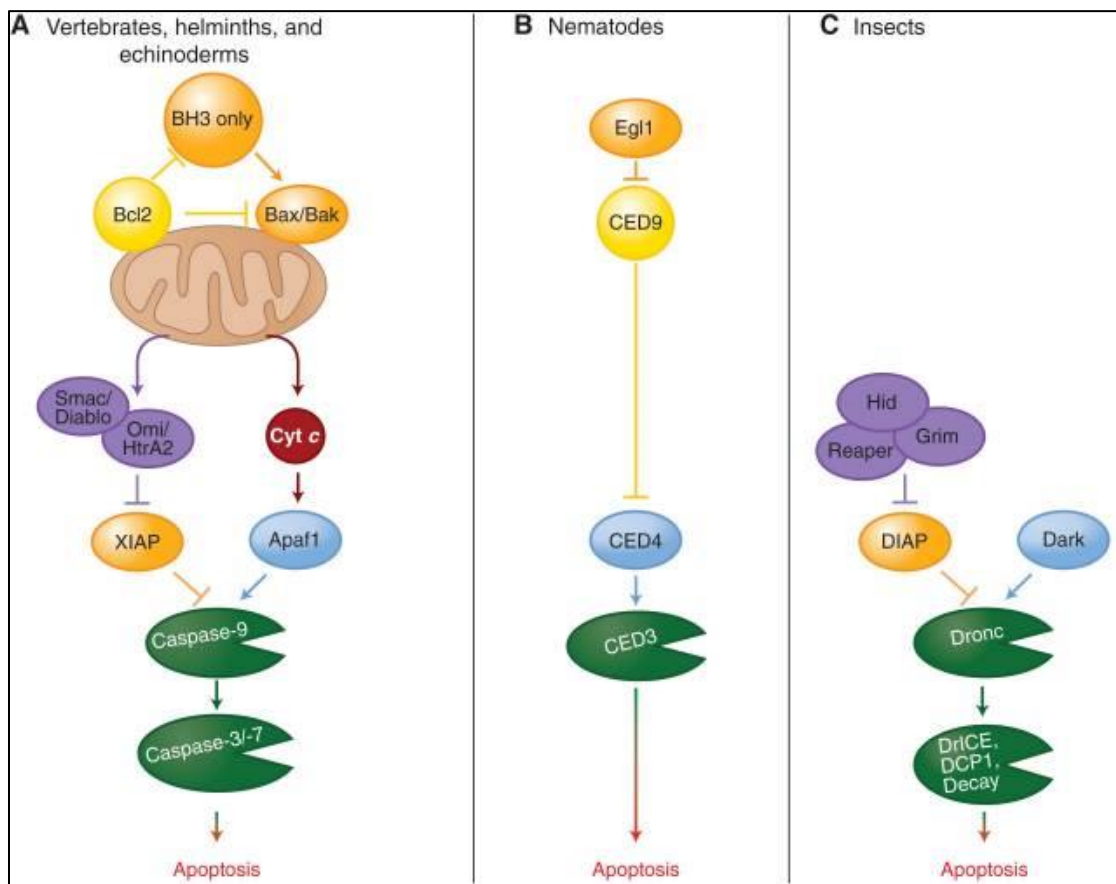


Figure 5 – Conservation des voies de signalisation de l'apoptose chez les vertébrés et les invertébrés

Cette figure, adaptée de Green et Llambi [88], représente l'ensemble des protéines connues comme étant impliquées dans l'activation de l'apoptose chez les vertébrés, les nématodes et les insectes. Les protéines jouant des rôles équivalents d'une espèce à l'autre sont représentées de la même couleur. Le rôle et l'implication des protéines MIFs dans ces voies chez les invertébrés reste encore inconnu.

II.2.2- La PCD de type II ou l'autophagie

II.2.2.1- Rappel sur l'autophagie

La macroautophagie, ou plus simplement l'autophagie est un processus cellulaire très conservé chez l'ensemble des eucaryotes. Ce processus catabolique se met en place en présence de stress métabolique tel que la carence en nutriments. L'autophagie peut aussi être induite en cas d'excès de dérivés réactifs de l'oxygène (ROS, pour reactive oxygen species). Il permet aussi la dégradation des organites endommagés (comme les mitochondries), des agrégats protéiques et des organismes infectieux [89–91]. L'autophagie joue un rôle crucial dans l'immunité innée et acquise, à la fois en permettant l'élimination des bactéries intracellulaires et en permettant la présentation des antigènes à la surface des cellules à l'aide du complexe d'histocompatibilité

de type I ou II (CMH I ou II) [92]. L'autophagie est considérée comme un processus de survie cellulaire.

L'autophagie est caractérisée par la formation d'une structure particulière entourée d'une double membrane : l'autophagosome. L'autophagosome enveloppe le contenu intracellulaire destiné à la dégradation puis fusionne avec les lysosomes, formant l'autolysosome. Cette fusion conduira à la dégradation du contenu de l'autophagosome [89, 93].

II.2.2.2- Rôle des protéines MIFs sur l'autophagie

Des études récentes ont montré que les protéines MIFs sont impliquées à plusieurs niveaux dans la régulation de l'autophagie. Les mécanismes exacts restent toutefois inconnus. Tout d'abord, certaines études montrent que les MIFs sont capables d'inhiber l'autophagie en activant la voie de signalisation PI3K/AKT[94]. *A contrario*, d'autres études ont montré que les protéines MIFs induisent ou facilitent l'autophagie en favorisant la production de ROS.

En parallèle, l'autophagie est un processus connu pour réguler la production et la sécrétion de nombreuses cytokines pro-inflammatoires dont les protéines MIFs [95] en réponse à la présence de pathogènes viraux ou bactériens. De plus, la carence en nutriments est un facteur favorisant la sécrétion de protéines MIFs à l'extérieur de la cellule [96, 97].

Les protéines MIFs sont donc des protéines qui peuvent à la fois agir comme des protéines inhibitrices ou activatrices de l'autophagie en fonction des types cellulaires, mais qui sont aussi produites et sécrétées en cas d'autophagie.

II.2.3- La PCD de type III ou la nécrose

La nécrose regroupe un grand nombre de processus biologiques ayant pour point commun la perte de l'intégrité de la membrane plasmique et un relargage du contenu cytoplasmique dans le milieu extracellulaire [98, 99]. Ce processus peut être provoqué par des blessures ou des infections et ne requiert alors l'activation d'aucune voie de signalisation spécifique. Contrairement à l'apoptose, ce processus active une réponse inflammatoire. Il existe, par ailleurs, de nombreux types de nécrose activée, et seuls ceux pour lesquels les protéines MIFs jouent un rôle sont présentés ici.

II.2.3.1- La nécroptose

Parmi les PCD de type III, la nécroptose, est activée volontairement. Il s'agit de la forme la plus caractérisée de nécrose programmée. La nécroptose nécessite l'activation de récepteurs spécifiques RIPK1 et 3 (receptor-interacting kinases). L'activation de ces récepteurs induit une réponse inflammatoire passant par la sécrétion de protéines pro-inflammatoires tel que les protéines MIFs.

A l'heure actuelle, les protéines MIFs, ont été décrites comme impliquées dans la nécroptose dans le cadre d'une réponse inflammatoire uniquement. Le processus exacte permettant l'activation de la réponse inflammatoire n'est pas encore connu mais passerait par le relargage

de protéines disposant de motifs moléculaires associés aux dégâts dus à la nécroptose (DAMP, damage-associated molecular patterns) [100].

II.2.3.2- La parthanatose

La parthanatose est une forme de nécrose programmée, due, par exemple, à une excitation excessive des cellules neuronales (excitotoxicité) ou à l'action d'un agent mutagène [101]. Cette forme de mort cellulaire est aussi impliquée dans la mise en place de nombreuses maladies neurodégénératives. Des découvertes récentes ont pu démontrer le rôle prépondérant des protéines MIFs dans la mise en place de cette forme de mort cellulaire [73].

La parthanatose a lieu en présence de nombreux dégâts à l'ADN. En présence d'un ADN dégradé, la poly(ADP-ribose) polymérase -1 (PARP-1), une enzyme nucléaire, est activée (Figure 6). Celle-ci sort alors du noyau de la cellule et permet la sortie des protéines AIF (apoptosis-inducing factor) des mitochondries. Les protéines MIFs vont alors se fixer aux protéines AIF formant un complexe. Ce complexe AIF-MIF est ensuite transloqué dans le noyau de la cellule puis fixé à l'ADN. La fixation du complexe à l'ADN entraîne sa fragmentation via la fonction nucléase des MIFs. Cette fragmentation conduit à la mort de la cellule par parthanatose [101].

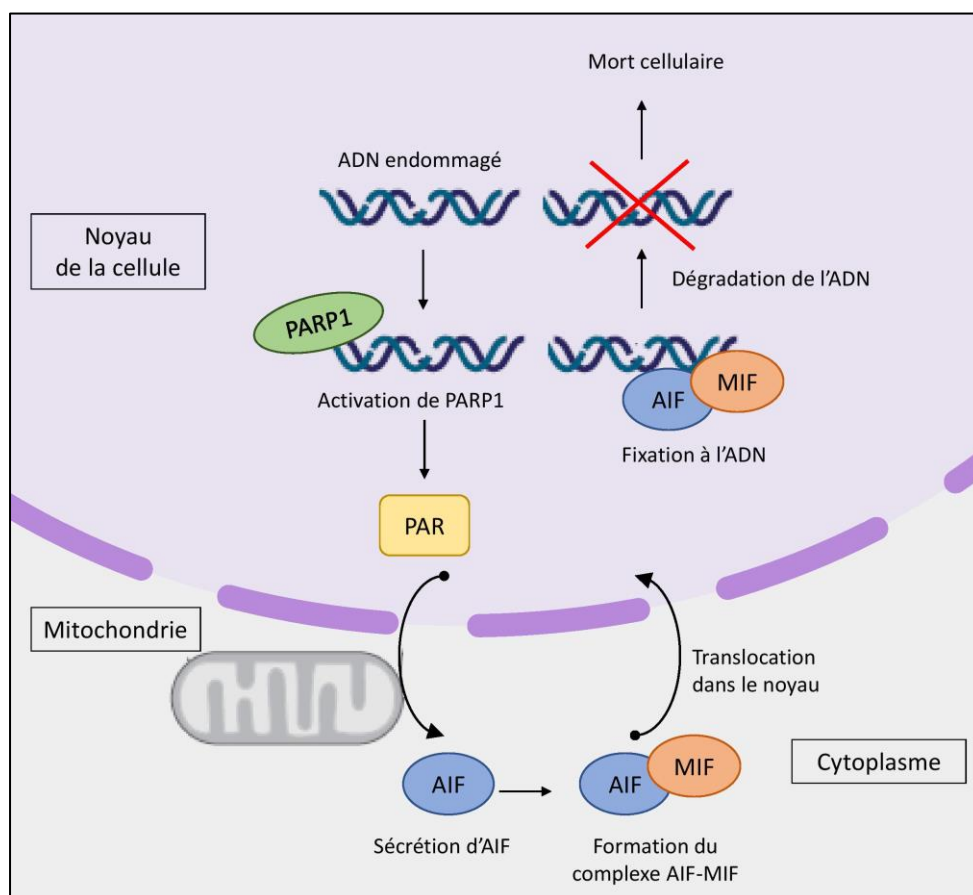


Figure 6 – Schéma simplifié de la voie de signalisation de la parthanatose

Schéma simplifié de la voie de signalisation de la parthanatose, adapté de Wang et al. [73]. De nombreux intermédiaires ne figurent pas sur ce schéma.

Chapitre 3 : Les cytokines MIFs chez les autres espèces

Les MIFs semblent être de très anciennes protéines. En dehors des vertébrés, on en retrouve chez de nombreux métazoaires tels que des mollusques comme *Biomphalaria glabrata* ou chez des espèces plus éloignées telles que les plantes ou des protistes. De plus, certaines études indiqueraient la présence de cette protéine chez certains procaryotes tel que *Prochlorococcus marinus*, une cyanobactérie [37]. Cependant, les travaux sur les fonctions des protéines MIF en dehors des mammifères restent extrêmement rares.

Ce chapitre présente brièvement les fonctions des MIFs décrites chez d'autres espèces que l'homme.

II.3.1- Chez les espèces libres

II.3.1.1- Chez les autres vertébrés

Les MIFs ont en grande majorité été étudiées chez les vertébrés notamment chez quelques espèces de poissons (*Danio rerio*, *Tetraodon nigroviridis*, *Branchiostoma belcheri*, ...), chez quelques oiseaux (*Gallus gallus*) et chez quelques amphibiens (*Xenopus laevis*) [102]. Les MIFs semblent toutes avoir un rôle dans l'immunité innée et acquise de ces espèces. En effet, elles jouent un rôle similaire à celui observé chez l'être humain quant à la différenciation cellulaire liée à l'immunité, la modification de la migration cellulaire ou encore la mort cellulaire [103–110].

Chez quelques espèces dont *Danio rerio* (poisson) [103], seules les fonctions des MIFs liées à la morphogénèse ont pu être confirmées.

Les fonctions enzymatiques des MIFs (tautomérase, oxidoréductase et nucléase) ont été très peu étudiées chez les autres espèces. La fonction nucléase, n'a par exemple pas du tout été étudiée, sa découverte étant assez récente. Toutefois, pour tous les vertébrés ayant fait l'objet d'une étude, les MIFs ont toutes conservé leurs fonctions oxydoréductase et tautomérase.

II.3.1.2- Chez les invertébrés

Une des premières études fonctionnelles sur les MIFs chez des invertébrés a été réalisée chez le gastéropode *Biomphalaria glabrata*, [16]. Chez cette espèce, les fonctions majeures des MIFs humains ont pu être confirmées, en particulier l'activité tautomérase, l'induction de la prolifération cellulaire liée à un challenge immunitaire et l'inhibition de la voie de l'apoptose induite par p53 [16].

Depuis, différentes études ont été réalisées sur d'autres espèces d'invertébrés comme des échinodermes (*Patiria pectinifera* [17]), des arthropodes, des mollusques et des helminthes [11, 18, 19, 35, 36].

Chez l'ensemble des espèces étudiées, les MIFs jouent un rôle dans l'immunité de l'organisme. Elles contribuent notamment au recrutement de cellules nécessaires à la mise en place d'une

réponse immunitaire chez *Haliotis diversicolor* [111] (couramment appelé l'ormeau) et chez *Patiria pectinifera* [17], une étoile de mer.

Cette apparente conservation des fonctions liées à l'immunité est particulièrement intrigante car les invertébrés ne possèdent pas les récepteurs spécifiques connus chez les vertébrés (CD74, CXCR2, 4 et 7). Les mécanismes exacts permettant la régulation de la migration cellulaire, de la prolifération cellulaire et des autres fonctions ne sont donc pas encore décrits chez ces espèces et ne passent pas par les récepteurs ou les voies métaboliques décrites chez l'homme.

II.3.2- Chez les parasites d'animaux

Les cytokines MIFs ont été décrites chez de nombreuses espèces de parasites de vertébrés. On les retrouve notamment chez de nombreux nématodes tels que *Wuchereria bancrofti*, *Anisakis simplex* ou encore *Brugia malayi* ainsi que chez de nombreux protistes tels que *Plasmodium falciparum* ou *Leishmania major*, responsables respectivement du paludisme (aussi appelé malaria) et de la leishmaniose [10, 11, 15, 33, 35, 36, 108, 112–120]. Pour les espèces pluricellulaires (*Brugia malayi*), les protéines MIFs sont liées à la régulation de la migration des cellules de l'immunité [10, 36].

De plus, chez ces espèces parasite, les MIFs jouent aussi un rôle majeur dans l'interaction avec leur hôtes. En effet, les MIFs sont injectées par les parasites dans leurs hôtes respectifs et agissent sur leurs défenses. Ainsi, pour ne donner qu'un exemple, durant une infection par *Plasmodium falciparum*, les protéines MIFs sont sécrétées par le parasite pendant la longue phase dite « sanguine » de propagation du parasite à l'intérieur du corps de l'hôte mammifère. Elles sont directement liées à l'inhibition de la migration des macrophages de l'hôte, ainsi qu'à la sous-expression à la surface des monocytes des récepteurs de TLR2, TLR4 et CD86 [15, 33, 117]. Ces mécanismes permettent une meilleure propagation du parasite en diminuant fortement les défenses de l'hôte. Chez d'autres espèces, telle que *Eimeria acervulina*, responsable de la coccidiose, les MIFs ont été démontrées comme induisant la sur-expression de cytokines proinflammatoires, renforçant le développement de la maladie [113].

Les MIFs chez les espèces parasites de vertébrés sont donc capables d'inhiber ou de perturber l'immunité de leurs hôtes.

II.3.3- Les cytokines MIFs chez les pucerons

II.3.3.1- Rappel sur les pucerons

Les pucerons sont des insectes de l'ordre des Hémiptères et de la super famille des *Aphidoidea*. Cette super famille contient plus de 4300 espèces, toutes parasites de plantes. Ravageurs majeurs des cultures, les pucerons sont des insectes piqueurs - suceurs de petite taille (entre 1.5 et 3.5 mm) [121]. Ils se nourrissent exclusivement de la sève élaborée contenue dans le phloème des plantes. Pour se nourrir, ils disposent de pièces buccales spécifiques, le stylet, permettant à la fois l'absorption de sève élaborée par l'insecte mais aussi l'excrétion de salive et son injection dans les cellules végétales. Dès le premier contact du stylet avec les cellules de l'épiderme de

la plante et tout au long de son interaction avec celle-ci, le puceron injecte un cocktail de sécrétions salivaires dans les cellules végétales [122].

Si l'effet de cette salive sur la plante et sa composition exacte reste encore à élucider, il est maintenant démontré qu'elle est nécessaire à l'inhibition des défenses de la plante, différente d'une espèce de puceron à l'autre et expliquerait en partie le panel d'hôtes de certains pucerons.

Les premières études sur le sujet ont démontré que la salive des pucerons est composée de très nombreux éléments. Elle est composée en partie d'enzymes, telle que les pectinases ou les cellulases qui permettent de faciliter la progression du stylet des pucerons dans les cellules végétales. De plus, elle a été démontré comme contenant des effecteurs permettant la régulation de l'immunité de la plante [123]. Chez *Myzus persicae*, le puceron vert du pècher, la protéine Mp55 est ainsi connue comme nécessaire à l'alimentation du puceron et capable d'inhiber les défenses de la plante [124]. Une autre protéine, Mp10, altère les défenses de la plante en agissant sur les voies de signalisation dépendantes de l'acide jasmonique et de l'acide salicylique. Ces deux voies permettent l'induction des mécanismes de défense de la plante, notamment l'induction de mort cellulaire sur le site d'infection [123, 125, 126].

II.3.3.2- Les MIFs chez les pucerons

En 2010, le premier génome de puceron fut séquencé et publié [127]. L'analyse de ce génome a permis de réaliser la première description de MIF chez un parasite de plantes. Chez *Acyrtosiphon pisum*, le puceron du pois, il existe quatre gènes codant des protéines MIFs différentes [128]. Les transcrits de ces gènes ont été retrouvés dans les cellules immunitaires des pucerons, les hémocytes, où ils sont induits après un challenge immunitaire (traitement avec des parasitoïdes, ...) [128]. En 2015, notre équipe a démontré qu'au moins une des différentes protéines MIFs est présente dans les glandes salivaires du puceron *Acyrtosiphon pisum*. Les protéines MIFs sont injectées dans les tissus de la plante lors de la nutrition du puceron. Les analyses ont pu démontrer que ces protéines sont nécessaires à la survie et à la nutrition des pucerons sur leur plante hôte [20].

Les protéines MIFs ont aussi été décrites comme étant capables d'inhiber un des mécanismes de défense de la plante : l'induction de la mort cellulaire [20, 128]. En effet, après la perception d'un agent pathogène, la plante peut déclencher la mort cellulaire programmée des cellules végétales, afin de limiter la propagation de cet agent [129]. On appelle ce mécanisme la réponse hypersensible ou HR. La cryptogéine, un éliciteur de plante de 10 kDa, issue d'un oomycète pathogène de plante, *Phytophthora cryptogea*, est capable d'induire l'ensemble des voies de signalisation qui conduisent à la HR. L'étude réalisée au sein de l'équipe a permis de démontrer que les protéines MIFs des pucerons sont capables d'inhiber la réponse hypersensible induite par la cryptogéine dans les cellules végétales [20].

Les protéines MIFs sont donc utilisées par des parasites d'animaux et par des parasites de plantes pour faciliter l'infection de l'hôte.

Le mécanisme qui permet l'inhibition des défenses de la plante est en cours d'étude dans notre équipe et reste encore inconnu. Toutefois, notons qu'aucun des récepteurs des MIFs connus chez l'humain n'existe chez le puceron ou chez la plante. De la même manière, seules certaines voies de signalisation affectées par les MIFs chez l'homme existent chez l'ensemble des eucaryotes tel que les voies impliquant les MAP kinases. Le mécanisme d'action des protéines MIFs chez les plantes ou les pucerons restent donc à élucider.

II.3.4- Les cytokines MIFs chez les plantes

Seule une étude *in silico* a pour l'instant été réalisée sur les MIFs chez les plantes. Selon cette étude, les MIFs sont présents chez l'ensemble des *Archeoplastida*. Chez *Arabidopsis thaliana*, la plante modèle, 3 gènes *MIF* ont été détectés *in silico*. Du fait de leur ressemblance avec les protéines MIF et D-DT, les protéines MIFs de plante ont été nommées MDL pour MIF Dopachrome Like. Aucune fonction n'a encore été décrite pour les différents MIFs de plantes [21].

Les premières analyses réalisées par l'équipe semblent indiquer que les protéines MIFs chez *Arabidopsis thaliana* sont nécessaires à la mise en place d'une réaction immunitaire face à un pathogène. En effet, en absence d'une de leurs MIFs, les mutants d'*Arabidopsis* sont plus sensibles que les plantes sauvages à des infections par des parasites aussi différents que champignons, oomycètes et pucerons (résultats de l'équipe, non-publiés).

Chapitre 4 : Diversité des MIFs

Comme indiqué dans le chapitre précédent, les MIFs ont été décrites chez de très nombreuses espèces allant des procaryotes (*Prochlorococcus marinus* [37]) à l'ensemble des eucaryotes (plantes, vertébrés, invertébrés, ...).

La majorité des études précédemment réalisées sur les MIFs d'organismes autres que mammifères sont des études fonctionnelles focalisées sur un petit nombre d'espèces. Aucune étude globale portant sur la présence des MIFs ou leur histoire évolutive dans l'ensemble des eucaryotes n'a pour l'instant été réalisée. Seules des phylogénies partielles ne comprenant qu'une partie des eucaryotes ont été réalisées.

Ces études ont toutefois pu mettre en évidence la forte conservation des protéines MIFs chez l'ensemble des eucaryotes [38]. Ainsi, le pourcentage d'identité des séquences protéiques des MIFs humains et de la protéine MIF d'*Eimeria acervukina* (un protiste) est de 53%. Il n'est en revanche que de 30% en moyenne pour les MIF d'*Arabidopsis thaliana* (AtMDL), la plante modèle.

Une publication récente résume les études réalisées sur les protéines MIFs chez les espèces non mammifères [130] et propose une phylogénie basée sur un petit nombre d'espèces. Elle ne comprend que les protéines MIFs pour lesquelles une analyse fonctionnelle a été réalisée et ne permet donc pas d'analyser l'évolution des protéines MIFs. Elle met cependant en évidence la grande diversité des fonctions des protéines MIFs et leur présence chez un grand nombre d'espèces.

Deux études se distinguent dans l'analyse de l'histoire évolutive des MIFs. Toutes deux se sont uniquement focalisées sur un petit groupe de métazoaires, respectivement les insectes et les nématodes.

L'étude sur les insectes, réalisée par notre équipe [128], a mis en avant des phénomènes évolutifs intéressants. Chez deux grands groupes d'insectes, les Diptères et les Hyménoptères (comme respectivement les mouches et les abeilles) aucune espèce séquencée à l'heure actuelle ne possède de MIF. *A contrario*, les MIFs ont subi des duplications multiples chez les Aphididées, ce qui semble indiquer une importance évolutive des MIFs chez ces espèces [128].

Chez les nématodes, une phylogénie partielle a été réalisée [131]. Elle ne comprend que les protéines MIF de nématodes et ne permet donc pas d'analyser l'évolution des protéines MIFs de nématodes vis-à-vis des autres espèces. Elle ne permet notamment pas d'analyser l'évolution des protéines MIFs des nématodes sous l'angle du parasitisme ou d'une éventuelle co – évolution des protéines MIFs des nématodes et des plantes. Dans cette étude, les protéines MIFs de chaque espèce forment deux groupes bien séparés qui pourraient correspondre aux deux variants MIFs et D-DT décrits chez l'homme [38, 131]. Toutefois, en absence des protéines MIFs et D-DT humaines dans cette analyse, cette observation n'a jusque là pas pu être confirmée et l'origine évolutive des MIFs et des D-DT n'a pas pu être étudiée en détail.

Au sein des eucaryotes, aucune étude exhaustive n'a été réalisée pour déterminer la présence ou l'absence des MIFs dans l'ensemble des espèces. Seules des phylogénies partielles, ne comprenant pas, soit les plantes, soit les espèces parasites de plantes ont été réalisées.

Partie III : Objectifs de la thèse

Objectifs de la thèse

L'état actuel de nos connaissances sur les MIFs chez les eukaryotes pose différentes questions évolutives et fonctionnelles.

Une première question est celle de l'importance fonctionnelle et de la diversification potentielle des MIFs dans les règnes vivants. En effet, il semble que certaines espèces comme les pucerons possèdent 3 à 5 copies de MIFs, alors que la drosophile, par exemple n'en possède aucune [128].

Une seconde question est celle de l'histoire évolutive des MIFs et des relations structure/fonction au sein de cette famille multigénique.

En effet, les analyses fonctionnelles indiquent que les MIFs peuvent soit (i) être indispensables à la mise en place de la réaction immunitaire chez les organismes libres et/ou (ii) être nécessaires au contournement ou à l'inhibition de ces mêmes défenses chez les espèces parasites, suggérant des fonctions potentiellement opposées. En particulier, notons le cas de l'effet sur la mort cellulaire. Les protéines MIFs humaines sont connues pour interagir sur des voies de signalisation capables d'inhiber plusieurs types de morts cellulaires (apoptose, autophagie, parthanatose). *A contrario*, des MIFs de parasites de l'homme ou de mammifères induisent différents types de mort cellulaire [40, 41, 73]. Enfin, les MIFs de pucerons inhibent également la mort cellulaire, qui, chez les plantes, est un mécanisme majeur de défense. Cette opposition apparente de fonction, à la fois entre hôtes et parasites, et entre parasites de mammifères et parasites de plantes, pose la question de l'origine et de l'évolution de ces fonctions au sein de l'ensemble des eucaryotes.

*** Le premier objectif de cette thèse était donc de répertorier et analyser la répartition du nombre de MIF chez les eucaryotes, et de rechercher un éventuel lien entre nombre de MIF et écologie ou mode de vie des espèces. Cette analyse a porté sur plus de 800 espèces eucaryotes appartenant aux règnes des plantes, des champignons, des protistes et des métazoaires.**

*** Le second objectif de cette thèse était d'étudier l'histoire évolutive des MIFs grâce à des reconstructions phylogénétiques et d'analyser les sites sous sélection positive ou négative, sur un sous-ensemble de séquences. Une attention particulière a été donnée aux phyla comprenant à la fois des organismes libres et des parasites de plantes ou d'animaux.**

*** Le troisième objectif était d'explorer l'origine et la diversité de l'activité des MIFs sur la mort cellulaire. Pour cela, nous avons comparé l'effet des MIFs de différentes espèces (végétales ou animales) sur la mort cellulaire en cellule végétale.**

Partie IV : Matériels et Méthodes

Chapitre 1 : Analyses *in silico*

IV.1.1- Analyse phylogénétique des protéines MIFs

L'ensemble des étapes permettant les reconstructions des différentes phylogénies est résumé dans la Figure 7.

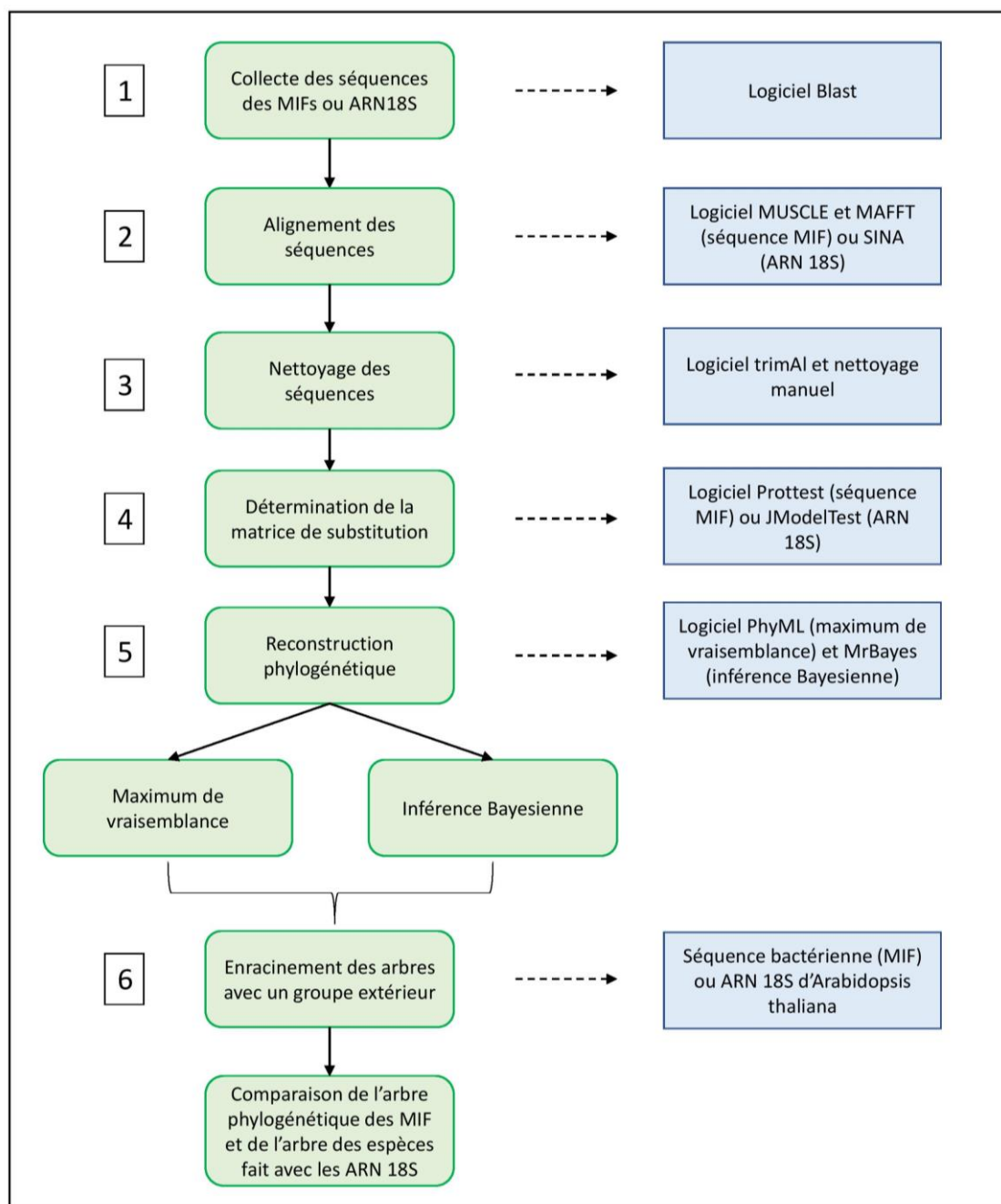


Figure 7 – Schéma récapitulatif des étapes de reconstructions des phylogénies

IV.1.1.1- Identification des protéines MIFs

La première étape d'identification de séquences MIFs d'espèces eucaryotes s'est focalisée sur des phyla comprenant à la fois des espèces libres et parasites de plantes ou d'animaux (**Erreur ! Source du renvoi introuvable.**, étape 1). Plusieurs séquences de MIFs bien caractérisées ont systématiquement été utilisées comme « queries » : les séquences MIF et D-DT humaines (Genbank accession numbers: NP_002406.1 and NP_001077861), celles de *Plasmodium berghei* [15] (Alvéolates ; Ensembl accession number : CDS52161), de *Leishmania major* [112, 132] (Excavates ; Ensembl accession numbers: LmjF.33.1740 and LmjF.33.1750), d'*Arabidopsis thaliana* [21] (Archaeplastida; Genbank accession numbers : NP_200527.1, NP_195785.1 and NP_566955.1), de *Biomphalaria glabrata* [16] (gastéropode ; Genbank accession numbers : XP_013094629.1 et ACR81565.1), et *Ancylostoma caninum* [10, 14] (nématode ; wormbase parasite accession numbers : ANCCAN_11528 and ANCCAN_25101).

Ces séquences ont été utilisées pour faire des recherches de similarités de séquences (BLAST) dans les différentes bases de données génomiques ou transcriptomiques listées dans le Tableau 1.

Seules les séquences s'alignant avec une E-value inférieure à 10^{-10} ont été conservées pour l'analyse.

Les séquences incomplètes, les erreurs de séquençages, les séquences peu ou mal alignées avec les autres ont été retirées de l'analyse.

La présence de la signature MIF dans l'ensemble des séquences a été vérifiée en utilisant le logiciel Interproscan [133].

Afin de pouvoir comparer au mieux l'évolution des espèces et celle des protéines MIFs, une phylogénie des espèces a été réalisée. Les séquences d'ARN 18S de l'ensemble des espèces utilisées pour l'analyse ont été recherchées dans la base de données SILVA [134].

Type de base de donnée	Nom de la base de donnée	Adresse du site	Référence
Généraliste	Ensembl	http://www.ensembl.org/index.html	[135]
Généraliste	NCBI	https://www.ncbi.nlm.nih.gov/	[136]
Champignons	Joint Genome Institute	http://jgi.doe.gov/	[137]
Champignons	FungiDB	http://fungidb.org/fungidb/	[138]
Diptères	Flybase	http://beta.flybase.org/	[139]
Pucerons	Aphidbase	http://bipaa.genouest.org/is/aphidbase/	[140]
Pathogènes	GeneDB	http://www.genedb.org/Homepage	[141]
Hyménoptères	HymenopteraMine	http://hymenopteragenome.org/hymenopteramine/begin.do	[142]
Nématodes	Wormbase	http://www.wormbase.org/	[143]
Nématodes parasites	Wormbase parasite	http://parasite.wormbase.org/	[144]
Cyanophora	Génome	http://cyanophora.rutgers.edu/cyanophora/home.php	[145]
Cryptomonas	Génome	http://cryptodb.org/cryptodb/	[146]
Globodera pallida	Génome	http://www.sanger.ac.uk/cgi-bin/blast/submitblast/g_pallida	
Pristionchus	Génome	http://www.pristionchus.org/blast/	
Meloidogyne	Génome	http://meloidogyne.inra.fr	[147]

Tableau 1 – Tableau des bases de données utilisées

Liste des bases de données et des génomes utilisés pour la recherche des séquences MIFs et informations sur le type de données et l'URL du site correspondant.

IV.1.1.2- Analyses phylogénétiques

Les séquences protéiques ont été alignées en utilisant les outils MUSCLE [148] et MAFFT [149] disponibles dans le logiciel Jalview [150] version 2.9. Les alignements issus de cette double analyse ont ensuite été comparés entre eux afin de s'assurer de leur similarité. Les séquences d'ARN 18S ont été alignées en utilisant le logiciel SINA [151] (SILVA Incremental Aligner), spécifiquement dédié aux séquences 18S (<https://www.arb-silva.de/>) (Etape 2 de la Figure 7).

Les séquences protéiques ont été nettoyées en utilisant le logiciel trimAl (mode gappyout, resoverlap et seqoverlap). Les sites ne présentant pas assez d'information phylogénétique et les séquences s'alignant sur moins de 85% de l'alignement ont été retirées (Etape 3 de la Figure 7). A l'issue du nettoyage, 113 positions ont été conservées pour l'analyse.

Le meilleur modèle de substitution des acides aminés pour l'analyse des séquences des protéines MIFs a été déterminé à l'aide du logiciel ProtTest version 3. Le modèle WAG a été estimé comme le plus approprié pour l'analyse des séquences des MIFs sans la distribution gamma ni la proportion des sites invariables. Pour les séquences des ARN 18S, le logiciel

JModelTest a été utilisé pour estimer le meilleur modèle de substitution. Le modèle GTR a été estimé comme le plus approprié pour l'analyse des séquences des ARN 18S (Etape 4 de la Figure 7).

Les phylogénies présentées dans ce manuscrit ont été élaborées par deux méthodes de reconstruction différentes : le maximum de vraisemblance et l'inférence Bayésienne (Etape 5 de la Figure 7).

Le logiciel PhyML a été utilisé pour la reconstruction phylogénétique basée sur le maximum de vraisemblance. Il est implémenté dans le logiciel SeaView, et a été utilisé avec les paramètres suivants: fréquences d'acides aminés (optimisés), la proportion des sites invariables et le nombre de substitution ont été estimés à partir de l'alignement. L'arbre de départ a été lancé via BIONJ et la méthode de reconstitution de l'arbre a été améliorée en utilisant à la fois la méthode NNI (Nearest Neighbor Interchanges) et SPR (Subtree Pruning and Regrafting). Les valeurs de support de branches ont été calculées en utilisant un test de ratio de probabilité approximatif (aLRT).

Pour l'analyse phylogénétique par inférence bayésienne, le logiciel MrBayes version 3.2.6 a été utilisé. Quatre analyses indépendantes ont été lancées avec différents arbres aléatoires au départ et les paramètres par défaut.

Pour les analyses de l'ARN 18S, la séquence de l'ARN 18S d'*Arabidopsis thaliana* a été utilisée en groupe extérieur (Etape 6 de la Figure 7). Pour les analyses des séquences des protéines MIFs, les séquences de 5 espèces de bactéries ont été utilisées : *Synechococcus* sp., *Thioalkalivibrio nitratreducens*, *Thioalkalivibrio paradoxus*, *Oscillatoria* sp. and *Microcoleus* sp. [Numéro d'accèsion GenBank respectif: WP_011360252.1, WP_015260126.1, WP_006746032.1, WP_017717428.1 and WP_015181735.1].

IV.1.2- Analyse de la sélection purificatrice ou diversifiante

L'analyse des sites sous sélection diversifiante (aussi appelée positive) ou purificatrice (aussi appelée négative) a été menée sur les séquences de 5 phyla différents contenant des espèces libres et parasites : (i) les Alvéolates (ii) les nématodes, (iii) les Straménopiles, (iv) les Aphididées, et (v) les plantes.

Pour l'analyse des Alvéolates, cinq espèces d'*Eimeria* et dix espèces de *Plasmodium* ont été sélectionnées, ainsi que *Toxoplasma gondii*, *Hammondi hammondi* et *Neospora caninum*. Pour les straménopiles, les séquences de *Phytophthora sojae*, *P. parasitica* et *P. infestans* ont été sélectionnées pour l'analyse. Pour l'analyse des nématodes parasites d'animaux, les séquences d'*Ostertagia ostertagi*, *Ascaris suum*, *Onchocerca volvulus*, *Brugia malayi*, *Loa loa*, *Trichinella spiralis* et *Ancylostoma ceylanicum* ont été sélectionnées. Pour les nématodes parasites de plantes, les séquences des protéines MIFs de *Bursaphelenchus xylophilus*, *Meloidogyne hapla* et *M. incognita* ont été sélectionnées. Pour l'analyse des Aphididées, les séquences MIFs d'*Acyrtosiphon pisum*, de *Myzus persicae*, de *Rhopalosiphum maidis* et *Rhopalosiphum padi* ont été sélectionnées. Enfin, pour l'analyse des plantes, les séquences de sept moniliophytes (fougères), de cinq chlorophytes (algues vertes), de deux bryophytes (mousses), d'un marchantiophyte (mousses de la famille des marchantiales), d'un spermatophyte, de cinq espèces monocotylédones et de cinq espèces eucotylédones ont été sélectionnées.

Les séquences MIFs choisies ont été alignées via les logiciels MUSCLE et MAFFT. Les séquences nucléotidiques correspondantes ont ensuite été alignées via le logiciel PAL2NAL.

La présence de sites sous sélection purificatrice ou diversifiante a été analysée à l'aide d'algorithmes différents implémentés dans le serveur Datamonkey [152] et de l'algorithme présent dans le logiciel Selecton [153].

Parmi les modèles disponibles dans le serveur Datamonkey, cinq modèles ont été sélectionnés :

- Le **modèle SLAC** (Single Likelihood Ancestor Counting) utilise une combinaison de maximum de vraisemblance et une approche excluante pour inférer le taux de substitution non synonyme (dN) et synonyme (dS) des différentes séquences sur les phylogénies correspondantes.
- Le **modèle FEL** (Fixed Effects Likelihood) [154] utilise une approche de maximum de vraisemblance pour déterminer le rapport entre les substitutions synonymes et non synonymes pour chaque site des séquences étudiées.
- Le **modèle REL** (Random Effect Likelihood) utilise une approche similaire au modèle REL mais en faisant varier l'impact des mutations aléatoires dans le calcul des rapports de substitutions synonymes et non synonymes pour chaque site des séquences étudiées.
- Le **modèle MEME** (Mixed-Effects Model of Evolution) [155] combine les effets fixes et les effets aléatoires pour identifier des phénomènes épisodiques de pression de sélection diversifiantes et des phénomènes de pression de sélection spécifique à une seule branche.
- Le **modèle FUBAR** (Fast Unconstrained Bayesian Approximation) utilise une méthode bayésienne de hiérarchisation approximative en utilisant une routine Monte Carlo de chaîne de Markov.

Le **logiciel Selecton** calcule le rapport entre les substitutions synonymes et non synonymes (la valeur ω) pour chaque codon des séquences choisies. Le logiciel a été utilisé avec le modèle de sélection positive (M8, $\beta + \omega \geq 1$). Le modèle M8 est le modèle par défaut dans le logiciel Selecton, il permet de détecter à la fois les sites sous sélection positive et négative. L'hypothèse nulle a été testée en utilisant à la fois le modèle M8a ($\beta + \omega = 1$) et le modèle M7. La comparaison de ces deux modèles a été faite en comparant les valeurs de LRT (likelihood ratio) et les valeurs d'AIC (Akaike Information Criterion). Pour l'ensemble des analyses réalisées, le modèle sous sélection comprend la valeur d'AIC la plus basse. Le niveau de précision de l'analyse est directement lié au moment où l'analyse bayésienne est considérée comme ayant convergé. Pour détecter au mieux les sites soumis à des pressions de sélection, le cut-off (ϵ) a été placé à 0.01, soit le niveau le plus élevé de précision défini par le logiciel Selecton.

IV.1.3- Analyses statistiques du nombre de protéines MIFs

Les différentes étapes des analyses statistiques du nombre de protéines MIFs par espèces sont résumées dans la Figure 8.

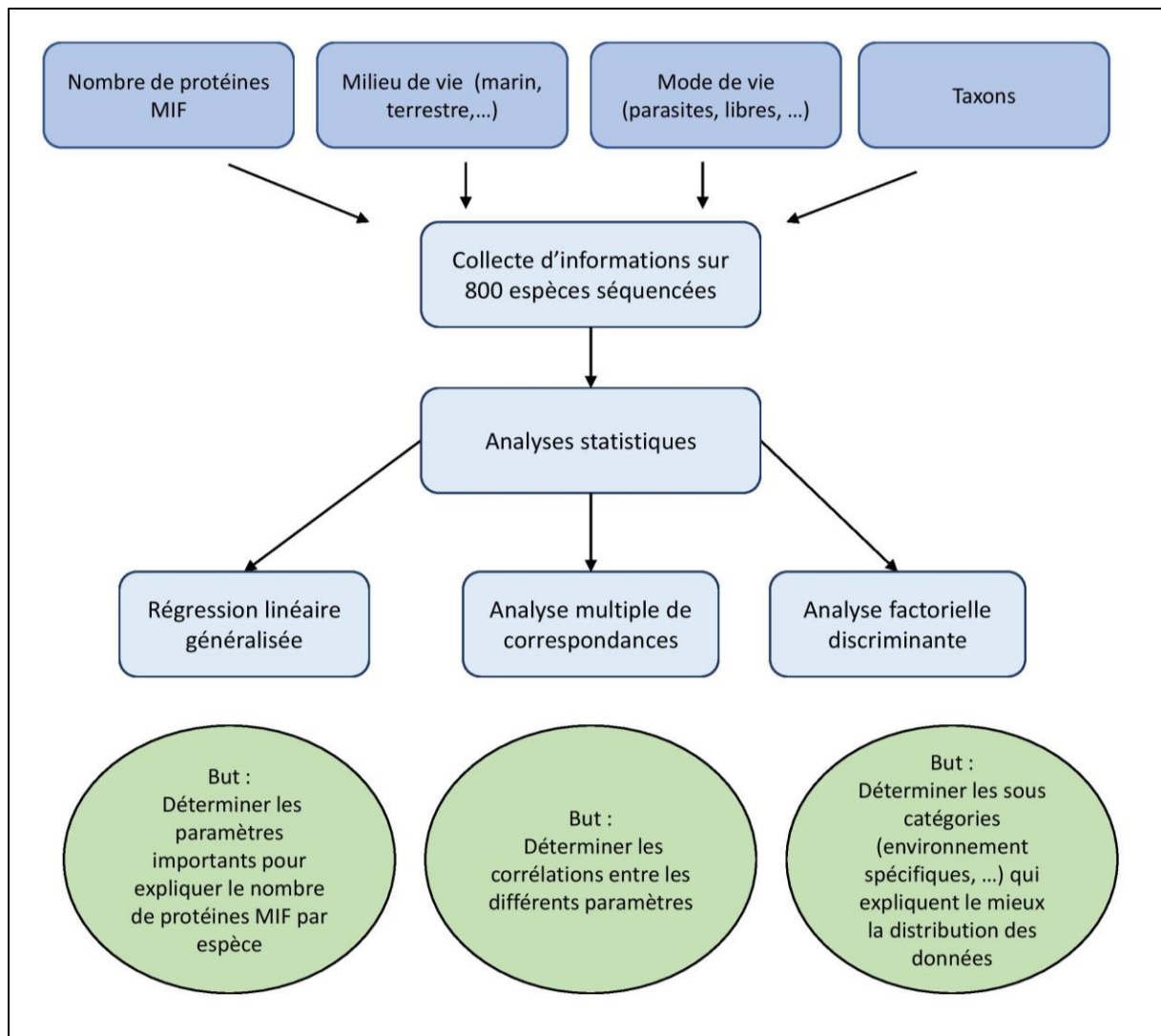


Figure 8 – Schéma récapitulatif des différentes analyses statistiques

Pour chacune des espèces eucaryotes étudiées, le nombre de protéines MIFs, le mode de vie de l'espèce (parasite d'animaux, parasite de plantes, espèces libres autotrophes ou espèces libres hétérotrophes), son taxon et son environnement (marin, terrestre ou eau douce) ont été compilés. Le tableau regroupant l'ensemble de ces informations est disponible en annexe (Tableau 8, page 177). Tout au long de l'analyse, le terme « parasite » a été utilisé lorsque l'espèce est en interaction durable avec son hôte. Ainsi, des symbiontes mutualistes, ou des organismes classiquement considérés comme des « pathogènes » (e.g. champignons phytopathogènes) ont été classés en tant que « parasite » dans ce travail.

Le lien potentiel entre le nombre de protéines MIFs présentes dans chaque espèce et le mode de vie, le taxon et l'environnement des espèces a été analysé par trois analyses statistiques différentes : (i) une régression linéaire généralisée, avec une distribution de Poisson, (ii) une analyse multiple des correspondances (ACM) et (iii) une analyse factorielle discriminante (AFD).

La régression linéaire généralisée permet de discriminer quels sont les facteurs importants (environnement, taxons ou mode de vie) pour déterminer le nombre de protéines MIFs par espèce. Le modèle permettant d'expliquer les données a été choisi en utilisant un algorithme progressif (stepwise). Cette analyse statistique a été réalisée à l'aide du logiciel R version 3.1.3.

Une analyse multiple des correspondances (ACM) a été réalisée afin de déterminer les éventuels liens de corrélation entre les taxons, les modes de vie, les environnements et le nombre de MIF par espèce. Elle a été effectuée *via* le package « ade4 » (Analysis of Ecological Data : Exploratory and Euclidean Methods in Environmental Sciences) présent dans le logiciel R. La fonction « scatter » a été utilisée afin d'obtenir une représentation graphique de cette analyse.

Une analyse factorielle discriminante (AFD) a été réalisée *via* le logiciel XLSTAT pour déterminer quels sont les sous catégories (taxons spécifiques, mode de vie spécifiques, ...) qui expliquent le mieux la distribution des données.

IV.1.4- Analyses structurales des MIFs

Les structures des protéines MIFs étudiées ont été prédites en utilisant deux logiciels : Phyre2 [156] et SWISS-MODEL [157–161]. Leur algorithme de calcul de prédiction de la structure est différent. Le logiciel SWISS-MODEL prédit aussi les éventuels assemblages des monomères de la même protéine. Pour l'ensemble des protéines étudiées, les structures obtenues par ces deux analyses sont quasiment identiques. De ce fait, seule la version obtenue avec le logiciel SWISS-MODEL a été utilisée par la suite. La prédiction des structures quaternaires des protéines a été confirmée à l'aide du logiciel ITasser [162–164].

La qualité de la prédiction de la structure des protéines a été validée par différentes méthodes :

- Le diagramme de Ramachandran a été réalisé avec ses paramètres par défaut *via* le logiciel « Ramachandran plot » inclus dans le logiciel de visualisation Swiss pdb viewer. Il permet de déterminer si les rapports entre les angles des différents atomes de la protéine restent dans des valeurs cohérentes avec la structure prédite.
- Le profil HCA (Hydrophobic Cluster Analysis) a été réalisé avec le logiciel en ligne HCA version 1.0.2 présent sur le serveur « RPBS Web Portal » (à l'adresse <http://mobyale.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=HCA#forms::HCA>). Il permet de déterminer les profils hydrophobiques de la protéine, et d'en déduire une structure potentielle de la protéine basée sur ses charges.

Les structures prédites ont été comparées manuellement et avec le logiciel MatchMaker inclus dans le logiciel de visualisation de protéine UCSF Chimera version 1.10.2.. La somme des différences qu'il y a entre chaque protéine MIF et la protéine MIF humaine a été estimée à l'aide de la valeur du RMSD (Root-mean-square deviation of atomic positions). Cette valeur correspond à la moyenne des distances qu'il y a entre les atomes de la première structure et celui des atomes correspondant sur la seconde structure. Le RMSD a été calculé à l'aide du logiciel « calculate RMS » présent dans le logiciel de visualisation Swiss pdb viewer.

IV.1.5- Prédiction des liaisons hydrogènes

A partir des structures protéiques prédites par le logiciel SWISS-MODEL, les liaisons hydrogènes présentes entre les monomères des protéines ont été déterminées. Pour cela, le logiciel « AddH » implémenté dans l'interface de visualisation Chimera a été utilisé pour permettre la visualisation de l'ensemble des hydrogènes de la protéine. Puis, les liaisons hydrogènes entre les monomères des protéines ont été déterminées à l'aide du logiciel FindHbond [165]. Les contraintes de détermination des liaisons hydrogènes ont été fixées à 0.4 Angströms de distance et 20 degrés d'angle maximum.

IV.1.6- Prédiction des modifications post-traductionnelles des protéines

Les séquences des protéines MIFs ont été étudiées pour leurs différentes modifications post-traductionnelles à l'aide de différents logiciels présents en ligne (Tableau 2). Tous les logiciels ont été utilisés avec leurs paramétrages par défaut.

Modification post traductionnelle	Nom du logiciel	Site web
Phosphorylation	NetPhos version 3.1	http://www.cbs.dtu.dk/services/NetPhos/
O-linked β -N-acétylglucosamine glycosylation	Yinoyang version 1.2	http://www.cbs.dtu.dk/services/YinOYang/
O-glycosylation	NetOGlyc version 4.0	http://www.cbs.dtu.dk/services/NetOGlyc/
N-glycosylation	NetNGlyc version 1.0	http://www.cbs.dtu.dk/services/NetNGlyc/
C-glycosylation	NetCGlyc version 1.0	http://www.cbs.dtu.dk/services/NetCGlyc/
Acétylation	NetAcet version 1.0	http://www.cbs.dtu.dk/services/NetAcet/

Tableau 2 – Tableau des logiciels de prédiction de modifications post-traductionnelles

Chapitre 2 : Analyses *in vivo*

Les MIFs d'une dizaine d'espèces ont été sélectionnées pour tester leur impact sur la modulation de la mort cellulaire en cellule végétale: *Bursaphelenchus xylophilus*, *Caenorhabditis elegans*, *Lingulodinium polyedrum*, *Meloidogyne incognita*, *Plasmodium falciparum*, *Solanum lycopersicae*, *Phytophthora infestans*, *Wuchereria bancrofti*, *Homo sapiens*, *Arabidopsis thaliana* et *Myzus persicae* (voir Partie V : Chapitre 2 :).

IV.2.1- Expression des gènes *MIF* en cellule végétale

Pour l'expression transitoire des gènes *MIF* en cellule végétale, le modèle de transfection des cellules d'épiderme de tabac par les bactéries *Agrobacterium* a été choisi. Les différentes étapes nécessaires à l'expression des gènes *MIF* en cellule végétale sont présentées dans la Figure 9.

IV.2.1.1- Conditions de cultures

a) Cultures des plantes

Les tabacs (*Nicotiana benthamiana*) ont été cultivés en pièce climatisée à hygrométrie et température constante (23°C +/- 1°C, 60% d'hygrométrie) sous photopériode (16h de jour). Les plants ont été repiqués après 2 semaines de culture et infiltrés 4 semaines après le semis. Les plants sont issus d'un stock de graines originaires du Sainsbury laboratory (Norwich-Royaume Uni).

b) Multiplication des cultures bactériennes

Les bactéries *Escherichia coli* (DH5 α , Invitrogen®) et *Agrobacterium tumefaciens* (GV3101) ont été multipliées respectivement à 37°C et 28°C dans du milieu Luria-Bertani (LB), additionné des antibiotiques adéquats (gentamycine, spectinomycine, voir plus loin).

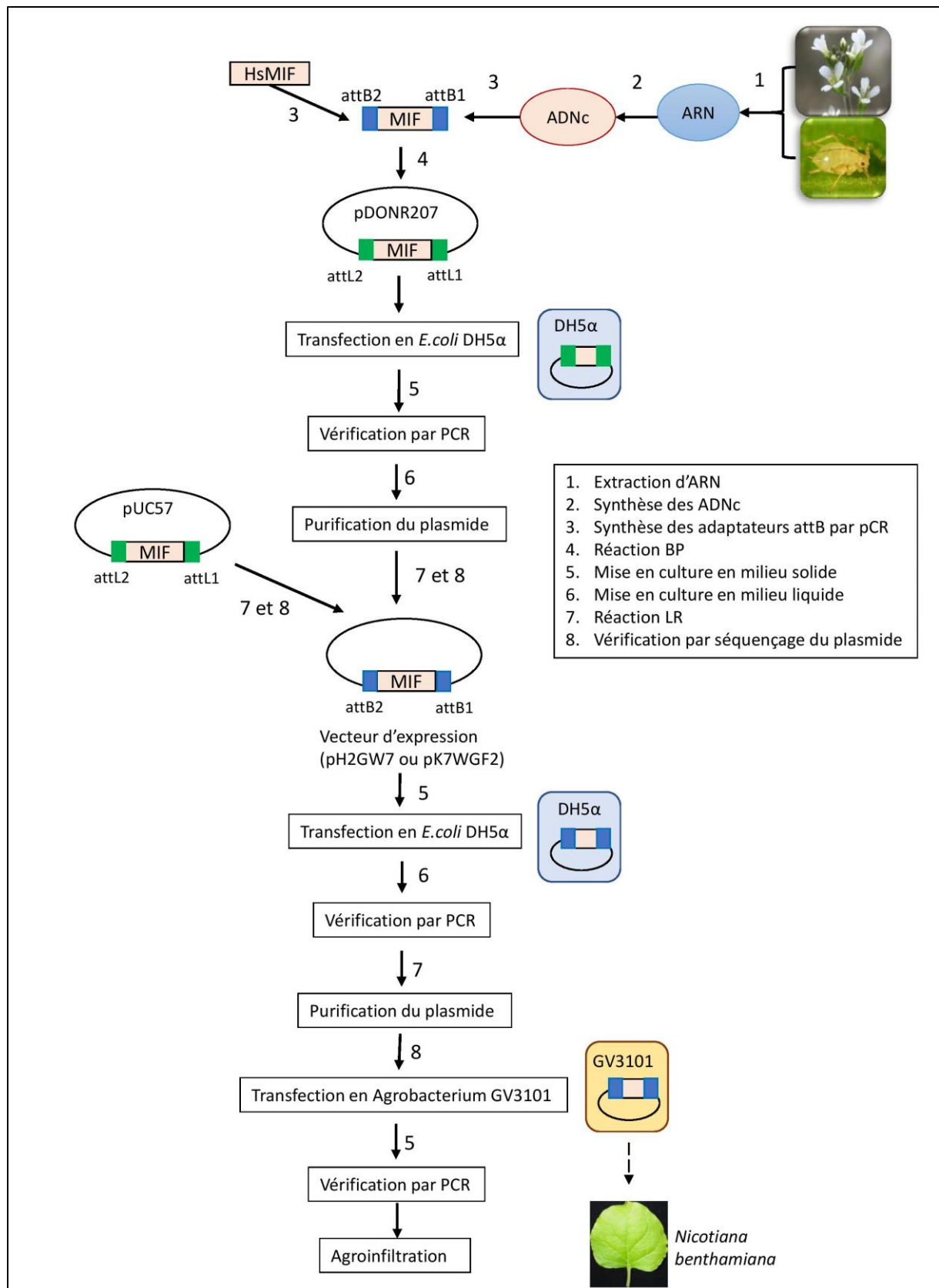


Figure 9 – Protocole pour l'expression des gènes *MIF* en cellule végétale

IV.2.1.2- Production des séquences d'ADN

Les gènes *MIF* ont été synthétisés par la société Genscript et directement fournis dans le vecteur pUC57. Les adaptateurs Gateway® (Invitrogen) attL1 et attL2 (Tableau 3) ont été ajoutés lors de la synthèse afin de faire du vecteur pUC57 un vecteur donneur.

Le clonage dans les vecteurs d'expression (ou plasmides de destination) a été obtenu par la recombinaison des extrémités attL et attR (réaction LR). La réaction de LR a été réalisée à 25°C durant la nuit en suivant les indications du fournisseur (Etape 7 de la Figure 9). Les plasmides de destination choisis sont : pH2GW7 (surexpression sous contrôle du promoteur fort p35S) et pK7WG2 (fusion GFP N-terminale sous contrôle du promoteur 35S).

Les bactéries *Escherichia coli* DH5α (Invitrogen®) ont ensuite été transformées, par choc thermique, avec la totalité de la réaction LR, puis mises en culture sur boîte de pétri contenant du LB agar additionné de 100µg/ml de spectinomycine. Une PCR a été réalisée sur chacune des colonies résistantes à l'antibiotique afin de vérifier l'intégration du gène dans le vecteur de destination (Etape 5 de la Figure 9). Dans le même temps, les colonies ont été mises en culture en milieu LB liquide. Après multiplication bactérienne, les plasmides ont été purifiés à l'aide du kit Genelute plasmid miniprep (Sigma®) (Etape 6 de la Figure 9). Les constructions géniques ont ensuite été séquencées par la société GATC (Etape 8 de la Figure 9). Le plasmide a été intégré dans les bactéries *Agrobacterium tumefaciens* (souche GV3101) par électroporation. La présence du plasmide a ensuite été confirmée par PCR sur colonies.

Pour *Arabidopsis thaliana* et *Myzus persicae*, les constructions géniques ont été faites à partir d'ARN. Les ARNs totaux ont été extraits via le protocole TRIzol™ Reagent (Etape 1 de la Figure 9). Via le kit iScript™ CDNA Synthesis Kit (Bio RAD), les ADNc correspondants ont été synthétisés (Etape 2 de la Figure 9). A l'aide d'amorces spécifiques (voir Tableau 3), par PCR, les séquences des gènes *MIF* ciblés ont été récupérés et les extrémités attB1 et attB2 ont été rajoutées (Etape 3 de la Figure 9).

Pour le gène *MIF* humain (HsMIF), la construction génique provient du professeur J. Bernhagen (Ludwig-Maximilians-University, Munich).

Les constructions obtenues (HsMIF, AtMDL et MpMIF) ont été clonées dans un plasmide pDON207 par une réaction BP grâce à la technologie Gateway® [166] (Invitrogen). La réaction BP a été réalisée à 25°C durant la nuit dans des quantités respectant les indications du fournisseur (Etape 4 de la Figure 9).

Après cette étape, le protocole est le même que pour les constructions issues du vecteur pUC57.

Extrémités	Séquences
Extrémités attL1	AAATAATGATTTTATTTTGAAGTATGACCTGTTCTGCAACAAATTGATGAG CAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCT
Extrémités attL2	CACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTG TTGCAACGAACAGGTCATCAGTCAAAATAAAATCATTATT

Tableau 3 – Tableau récapitulatif des extrémités Gateway®

IV.2.1.3- Agroinfiltration

Les différents vecteurs de destination (pH2GW7 et pK7WG2) ont été introduits dans les bactéries *Agrobacterium* (souche GV3101) par électroporation.

Les *Agrobacterium* possédant le vecteur d'expression ont été mis en culture dans du milieu LB additionné d'antibiotiques de sélection (50µg/mL de Rifampicine, 25µg/mL de Gentamycine et 100µg/ml de Spectinomycine) à 28°C et 200 rpm pendant toute la nuit. Le lendemain, une culture fraîche a été lancée jusqu'à atteindre une densité optique comprise entre 1 et 1.5. Après centrifugation, le culot bactérien a ensuite été resuspendu pour une DO₆₀₀ finale de 0,8 dans de l'Agromix (100 mM de milieu MES pH 5.7, 100 mM de MgCl₂ et 100 µM d'acétosyringone) pendant 1h. Un mélange volume à volume a été réalisé avec l'inducteur de mort cellulaire (AvrPto).

Les feuilles de *N. benthamiana*, âgées de quatre semaines, ont été infiltrées au niveau de l'épiderme inférieur. 100 µL de culture ont ensuite été infiltrés.

IV.2.2- Evaluation de la mort cellulaire

IV.2.2.1- Design expérimental

Trois réplicats biologiques ont été réalisés, chacun contenant au minimum dix points d'infiltration répartis sur six plantes différentes.

Afin d'obtenir des résultats comparables, les différentes constructions - correspondant à un même organisme vivant - ont été testées sur la même feuille et comparées à un contrôle surexprimant la GFP. Elles ont ensuite été analysées via une quantification du taux de chlorophylle restant dans les tissus.

Afin de tester un éventuel rôle inhibiteur des MIFs sur la mort cellulaire, la réponse hypersensible (HR) a été induite chez *N. benthamiana* par l'agroinfiltration d'AvrPto, un effecteur de la bactérie phytopathogène, *Pseudomonas syringae*. AvrPto est reconnu par la protéine de résistance Pto de la plante et induit une HR, visible par une mort cellulaire confluent dans la zone d'infiltration des Agrobactéries. Pour analyser l'activité des MIFs sur la mort cellulaire induite par AvrPto, l'effecteur a été co-exprimé avec les différents variants des MIFs. La co-expression d'AvrPto avec la GFP a été utilisée en tant que contrôle pour le développement de la HR en absence d'inhibiteur potentiel. Toutes les constructions protéiques MIF testées ont été comparées avec ce contrôle.

Un test T de Student a été réalisé pour évaluer le degré de significativité de l'ensemble des réplicats biologiques.

IV.2.2.2- Mesure de la quantité de chlorophylle

La mort cellulaire a été mesurée de façon quantitative grâce à la mesure de la concentration en chlorophylle. Les parties infiltrées ont été prélevées 3 jours après l'infiltration et découpées à l'aide d'un emporte pièce de taille standardisée. Les échantillons ont ensuite été disposés dans une plaque de 96 puits (plaque ELISA) contenant 300 µL d'alcool à 99%. Après 24h à 4°C, 150 µL de ces échantillons ont été transférés dans une nouvelle plaque. L'absorbance des échantillons a ensuite été mesurée aux longueurs d'onde de 750nm, 647nm et 664nm au spectrophotomètre.

A partir de cette lecture, la concentration en chlorophylle totale a été calculée à l'aide des équations suivantes :

Chlorophylle a (µg/mL) = 12.25 x (Abs. 664 nm – Abs. 750 nm) – 2.55 x (Abs. 647 nm – Abs. 750 nm)

Chlorophylle b (µg/mL) = 20.31 x (Abs. 647 nm – Abs 750 nm) – 4.91 x (Abs. 664 nm – Abs. 750 nm)

Chlorophylles totales = 17.76 x (Abs. 647 nm – Abs 750 nm) + 7.34 x (Abs. 664 nm – Abs 750 nm)

IV.2.3- Localisation subcellulaire

La localisation subcellulaire des différentes protéines a été étudiée en épiderme de tabac 48h après infiltration. Les feuilles ont été infiltrées avec les constructions présentant une fusion GFP en N-terminal de la protéine. Un marqueur du peroxisome a été exprimé concomitamment avec les constructions géniques à une densité optique d'environ 0.4.

Les zones d'infiltration ont été prélevées puis observées au microscope confocal (Microscope confocal LSM880 - Zeiss®) au grossissement x63. La GFP a été excitée à 488 nm avec un laser argon. La RFP a été excitée à 543 nm avec un laser hélium/néon.

Partie V : Résultats et Discussions

Chapitre 1 : Histoire évolutive des MIFs

Cette partie de la thèse fait l'objet d'une publication soumise à « Scientific Reports » le 20/09/2018. Celle - ci est présentée en annexe.

Dans l'ensemble de cette étude, les espèces considérées comme « parasites », au sens large, sont celles qui établissent une interaction durable avec un hôte. Ces espèces partagent des contraintes, importantes sur le plan fonctionnel et évolutif, comme la nécessité d'établir un dialogue moléculaire durable avec l'hôte et de contourner, moduler ou inhiber sa réponse immunitaire. Ainsi, les moustiques par exemple, ne sont pas considérés comme parasites dans cette étude, dans la mesure où l'interaction est de courte durée. *A contrario*, un champignon mycorhizien sera considéré « parasite » bien que son interaction avec la plante soit bénéfique pour l'hôte.

V.1.1- Présence des MIFs chez les eucaryotes

Pour associer la présence et le nombre de MIF par espèce avec des paramètres de mode de vie (autotrophie, hétérotrophie, parasitisme de plantes ou d'animaux) et l'environnement (terrestre, marin, eau douce), les données correspondantes ont été récoltées pour 803 espèces eucaryotes. Il faut noter que la détermination du nombre de gènes *MIF* par espèce n'a pas toujours été possible, en fonction de l'accessibilité et de la qualité de l'assemblage des génomes. De ce fait, dans ce travail nous n'analysons pas le nombre de copies de gènes *MIF*, mais le nombre de séquences protéiques différentes identifiées à l'issue d'une recherche *in silico*. L'ensemble des données est représenté dans le tableau en annexe (Tableau 8 – Listes des espèces eucaryotes utilisées pour l'étude).

Ce premier criblage montre que le nombre de MIF varie globalement de 0 à 5 sur l'ensemble des organismes eucaryotes (Figure 10). Des espèces possédant des protéines MIFs se trouvent dans l'ensemble des clades phylogénétiques, à l'exception des *Rhizaria*. Cependant, l'absence de MIF dans ce groupe taxonomique est incertaine car seules les séquences de génomes de cinq espèces sont actuellement disponibles.

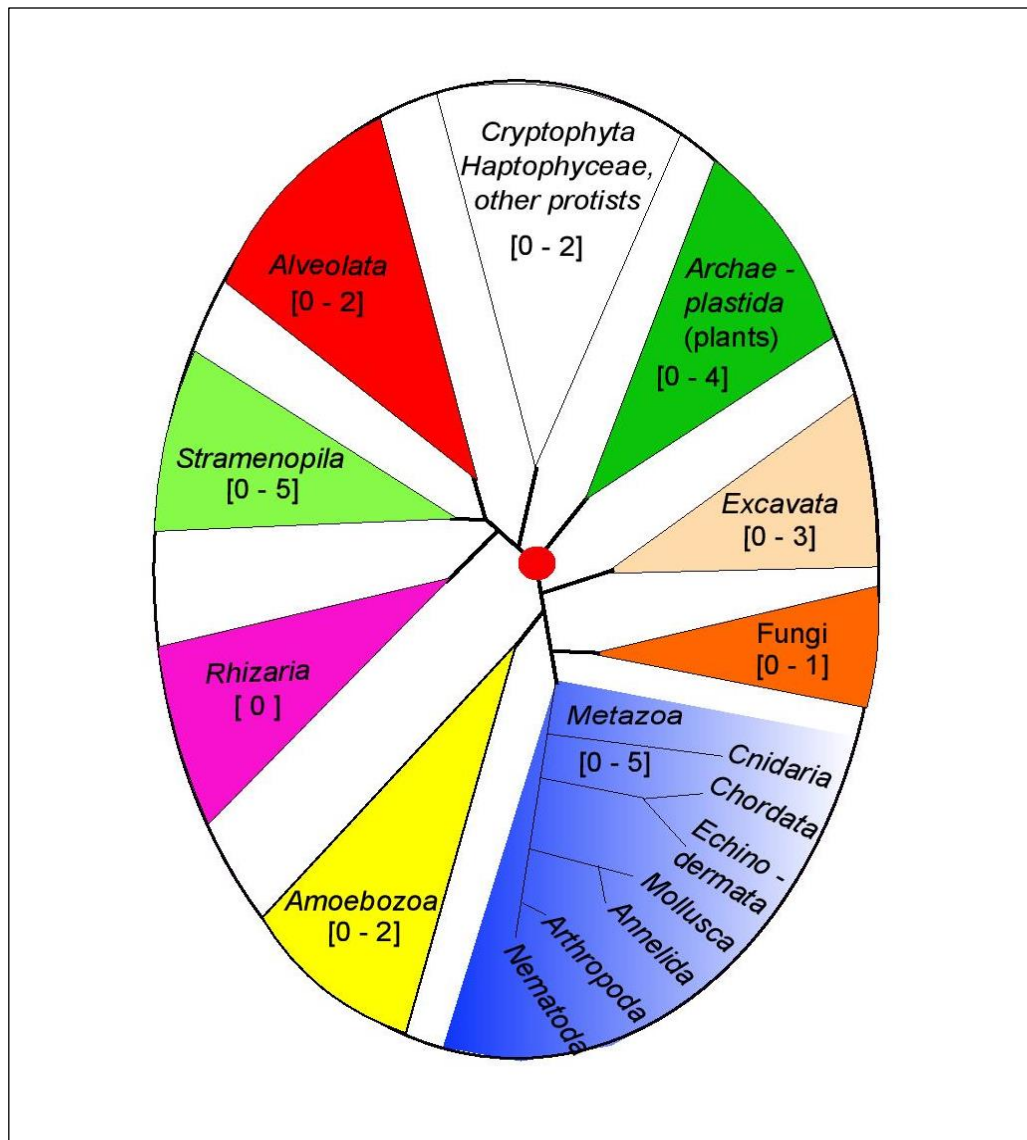


Figure 10 – Arbre de vie simplifié des eucaryotes

Arbre des eucaryotes basé sur Burki et al., 2014 [167], montrant, entre crochets le nombre de protéines MIFs présentes dans les espèces de chacun des clades. Les différents clades ont été distingués par des couleurs qui seront reprises dans l'ensemble des phylogénies présentées dans cette étude.

La suite de l'analyse a été réalisée uniquement sur les séquences de bonne qualité, et les groupes suffisamment importants. Ainsi, les groupes ne comportant qu'un faible nombre d'espèces séquencées tel que les Haptophytes, les *Parabasalia* et les *Apusozoa* pour lesquels il n'y a qu'un seul génome décrit ont été éliminés. Au total, 677 espèces sur les 803 espèces criblées ont été retenues pour la suite du travail. Trois analyses complémentaires ont été réalisées : une analyse « pas à pas » (stepwise) par régression linéaire généralisée, une analyse multiple de correspondance et une analyse discriminante.

L'analyse « pas à pas » des données par régression linéaire généralisée a été réalisée dans le but de déterminer les paramètres (mode de vie, environnement, taxon) qui expliquent le mieux la répartition du nombre de MIF par espèce. Cette analyse indique que l'environnement ne constitue pas un paramètre déterminant pour prédire le nombre de MIF par espèce. A

l'inverse, les paramètres « taxons » et « mode de vie » expliquent mieux les données que ne le ferait le hasard selon le critère d'information d'Aikaide, AIC.

L'analyse multiple de correspondance (ACM) (Figure 11) permet d'analyser les éventuels regroupements entre un mode de vie spécifique et un nombre de MIF particulier. Cette analyse a pu mettre en évidence la séparation distincte entre les différents modes de vie : les espèces autotrophes, hétérotrophes et parasites de plantes et d'animaux sont clairement regroupées en 3 grands groupes distincts. Cette forte séparation peut être expliquée par deux phénomènes :

- (i) Le fait que les espèces autotrophes sont en grande majorité représentées par le groupe des *Archaeplastida*. Or, dans ce groupe, la majorité des espèces possèdent trois MIFs.
- (ii) Le fait que 68% des parasites de plantes sont des champignons et que la majorité des champignons ne possède pas de MIF.

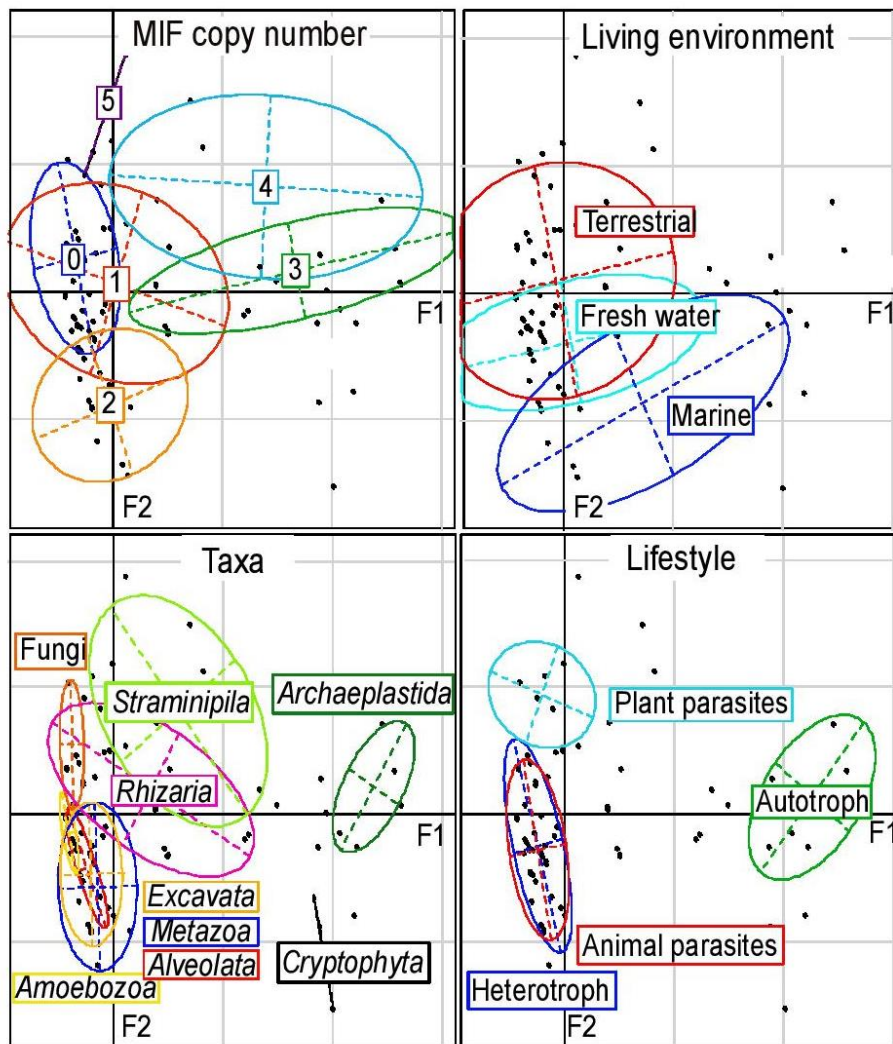


Figure 11 – Analyse multiple des correspondances

Représentation graphique de l'analyse multiple des correspondances (ACM), montrant séparément la distribution du nombre de protéines MIFs par espèces, les environnements, les taxa et les modes de vie des espèces (chaque point représente une espèce). Les ellipses délimitent le centroïde autour de chaque variable. Les axes F1 et F2 présentent des inerties ajustées de 52.54% et 15.64%, respectivement.

L'analyse discriminante (Figure 12) met en avant les groupes (taxons spécifiques, modes de vie spécifiques) ayant le plus d'influence dans la biais de répartition du nombre de MIF par espèce que l'on a pu observer dans les deux analyses précédentes. Cette analyse nous indique le rôle majeur des champignons, des *Archaeplastida* et des espèces autotrophes dans cette répartition. Les observations réalisées au sein des parasites de plantes et des espèces autotrophes sont donc potentiellement influencées par l'impact des champignons et des *Archeoplastida*.

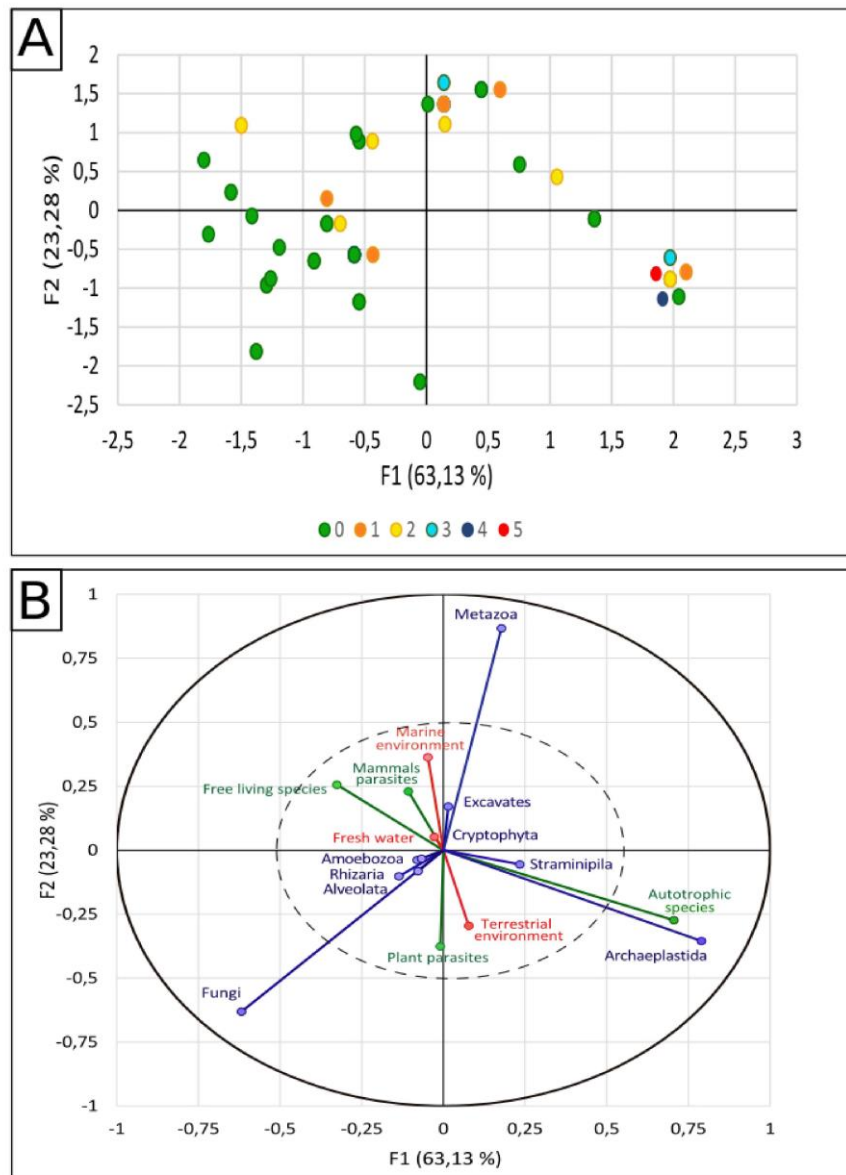


Figure 12 – Analyse discriminante de la répartition du nombre de protéines MIFs

Analyse discriminante montrant en (A) la répartition des différents nombres de protéines MIFs en fonction des taxons, des environnements et des modes de vie de chaque espèce et en (B) l'importance des différents paramètres (taxons, mode de vie et environnement). Les taxons sont représentés en bleu, les modes de vie en vert et les environnements en rouge. L'importance de chaque variable est notée par la proximité de l'extrémité de l'axe avec le cercle noir. Les critères compris dans le cercle en pointillés sont considérés comme ayant un rôle mineur dans la détermination du nombre de protéines MIFs par espèces.

Afin de limiter l'impact et le biais créé par les *Archeoplastida* et les champignons sur l'analyse, nous avons poursuivi cette étude en nous focalisant sur des groupes taxonomiques autres que *Archeoplastida* et champignons, et comprenant à la fois des espèces libres et parasites. Il s'agit des Alvéolates, Straménopiles, Nématodes et insectes (Figure 13).

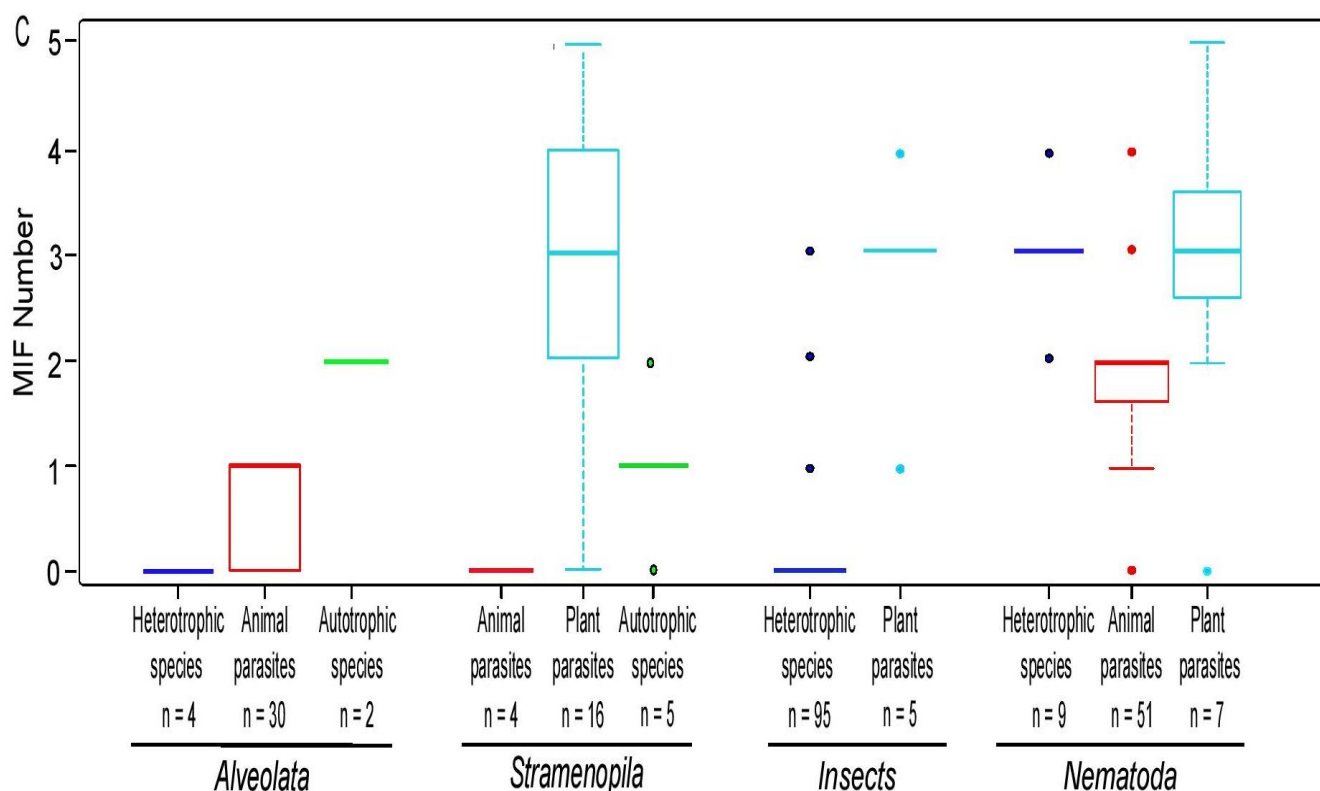


Figure 13 – Boxplot de la répartition du nombre de MIFs

Analyse de la répartition du nombre de protéines MIFs en fonction du mode de vie dans quatre taxons différents : Alvéolates, Straménopiles, Insectes et Nématodes. L'axe vertical correspond au nombre de protéines MIF par espèces. Chaque boîte est délimitée par les quartiles Q1 et Q3, qui correspondent à 50% de la distribution des protéines MIFs. La ligne horizontale dans le centre de la boîte correspond à la médiane de la répartition du nombre de protéines MIFs dans chaque groupe. Les deux barres verticales correspondent aux valeurs des premiers et derniers centiles (C1 et C99) et délimite 98% de la répartition. Les valeurs extrêmes (outliers) sont indiquées par un point et correspondent à moins de 2% de la répartition des protéines.

Contrairement à l'analyse globale (incluant les *Archeoplastida* et les champignons), cette étude montre que les espèces parasites de plantes possèdent un plus grand nombre de MIF (mediane=3) que les espèces parasites d'animaux.

Ainsi, l'ensemble de ces analyses indiquent que:

- (i) les plantes (*Archeoplastida*, espèces libres autotrophes) ont une médiane de trois MIFs, ce qui est plus élevé que le nombre médian de MIF des espèces ayant un autre mode de vie ou appartenant à un autre taxon.
- (ii) les espèces parasites de plantes, à l'exception des champignons, possèdent également un nombre médian de trois MIFs
- (iii) les espèces libres hétérotrophes et parasites d'animaux possèdent un nombre plus faible et/ou variable de MIFs.

Ces résultats soulèvent donc la question de contraintes évolutives particulières chez les plantes et les parasites de plantes, en terme de nombre de MIF.

V.1.2- Reconstruction phylogénétique des MIFs d'eucaryotes

V.1.2.1- Analyse globale

La phylogénie des MIFs a été reconstruite en utilisant la méthode d'Inférence Bayésienne (BI) et la méthode du maximum de vraisemblance (ML-Maximum Likelihood) sur un total de 720 séquences de haute qualité provenant de plantes, champignons, protistes, et métazoaires (détail en Tableau 8). Pour éviter de mauvaises interprétations dues à des branches peu soutenues, seule la topologie Bayésienne est représentée, et les deux valeurs de « approximate likelihood-ratio test (aLRT) » de l'analyse ML et de « posterior probability » (PP) de l'analyse BI sont indiquées aux branches correspondantes.

Une première observation générale est que la phylogénie des MIFs (Figure 14) ne suit pas celle des espèces, notamment en ce qui concerne les protistes, champignons, plantes et les métazoaires.

Une seconde observation générale est que la majorité des taxons ne forment pas de groupes monophylétiques. Seuls les *Viridiplantae* (plantes vertes), les Straménopiles et les Alvéolates forment des groupes monophylétiques fortement soutenus.

Enfin, cette phylogénie générale ne permet pas de confirmer une claire distinction des variants « MIF » et « D-DT » pour l'ensemble des métazoaires comme précédemment suggéré par des études phylogénétiques réalisées sur un nombre plus limité d'espèces [38]. Dans notre analyse, i) les MIFs de chordés forment un groupe à part entière et ii) de nombreuses séquences de métazoaires ne groupent ni avec le variant MIF ni avec le variant D-DT des chordés.

Ceci appuie donc l'idée que les MIFs de métazoaires proviennent probablement d'au moins une duplication précédant la diversification des métazoaires (entre 757 et 1.147 MYA selon TimeTree), mais la distinction entre « MIF » et « D-DT » n'est pas confirmée ici.

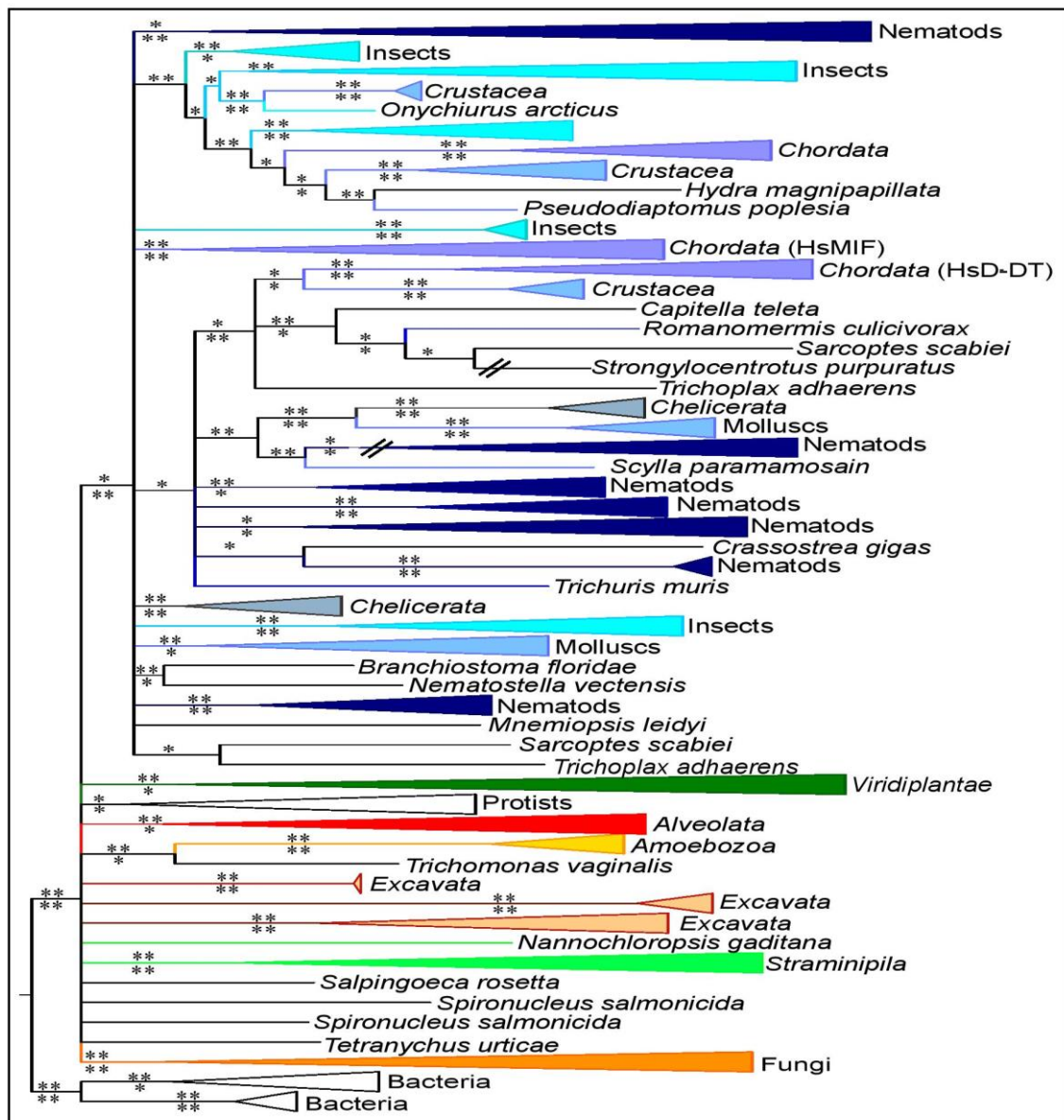


Figure 14 –Reconstruction phylogénétique des MIFs

Reconstruction phylogénétique par inférence bayésienne (BI). L'arbre a été enraciné en utilisant les séquences bactériennes des protéines MIFs. Les valeurs de confiance statistiques sont indiquées au-dessus des branches pour la méthode BI et en dessous des branches pour la méthode ML. Une étoile correspond à une valeur de PP supérieure à 0.5 ou une valeur aLRT supérieure à 0.7. Deux étoiles correspondent à une valeur de PP supérieure à 0.7 ou supérieure à 0.9 pour le test aLRT. Le même code de couleur sera utilisé pour l'ensemble des figures de phylogénie : Métazoaires en dégradé de bleu, Straménopiles en vert clair, *Archeoplastida* en vert foncé, Alvéolates en rouge, champignons en orange, Excavates en orange clair, et les *Amoebozoa* en jaune. Les Cryptophytes, *Haptophyceae* et *Picozoa* – regroupés sous le nom de protistes et les bactéries sont en blanc.

Dans la suite de l'analyse, les taxons qui contiennent à la fois des espèces libres (autotrophes ou hétérotrophes) et des espèces parasites sont examinées plus en détail : les Straménopiles, Alvéolates, champignons, nématodes et insectes. Les plantes ont également été examinées afin de comparer l'évolution des MIFs de parasites de plantes avec celles de leurs hôtes.

Pour comparer au mieux la phylogénie des MIFs avec l'arbre des espèces correspondantes, des analyses phylogénétiques ont été réalisées en parallèle sur les séquences de l'ARN 18S des espèces correspondantes.

V.1.2.2- Les MIF de plantes

L'ensemble des *Viridiplantae* possèdent des MIFs formant un groupe monophylétique fortement soutenu. Au sein des *Viridiplantae*, les protéines MIFs ne suivent pas la taxonomie et forment trois groupes très soutenus correspondant à trois copies des protéines présentes chez l'ensemble des Angiospermes (Figure 15). Ces résultats confirment des résultats similaires obtenus dans une étude précédente réalisée sur un échantillonnage plus réduit de *Viridiplantae* [8].

Résultats et Discussions

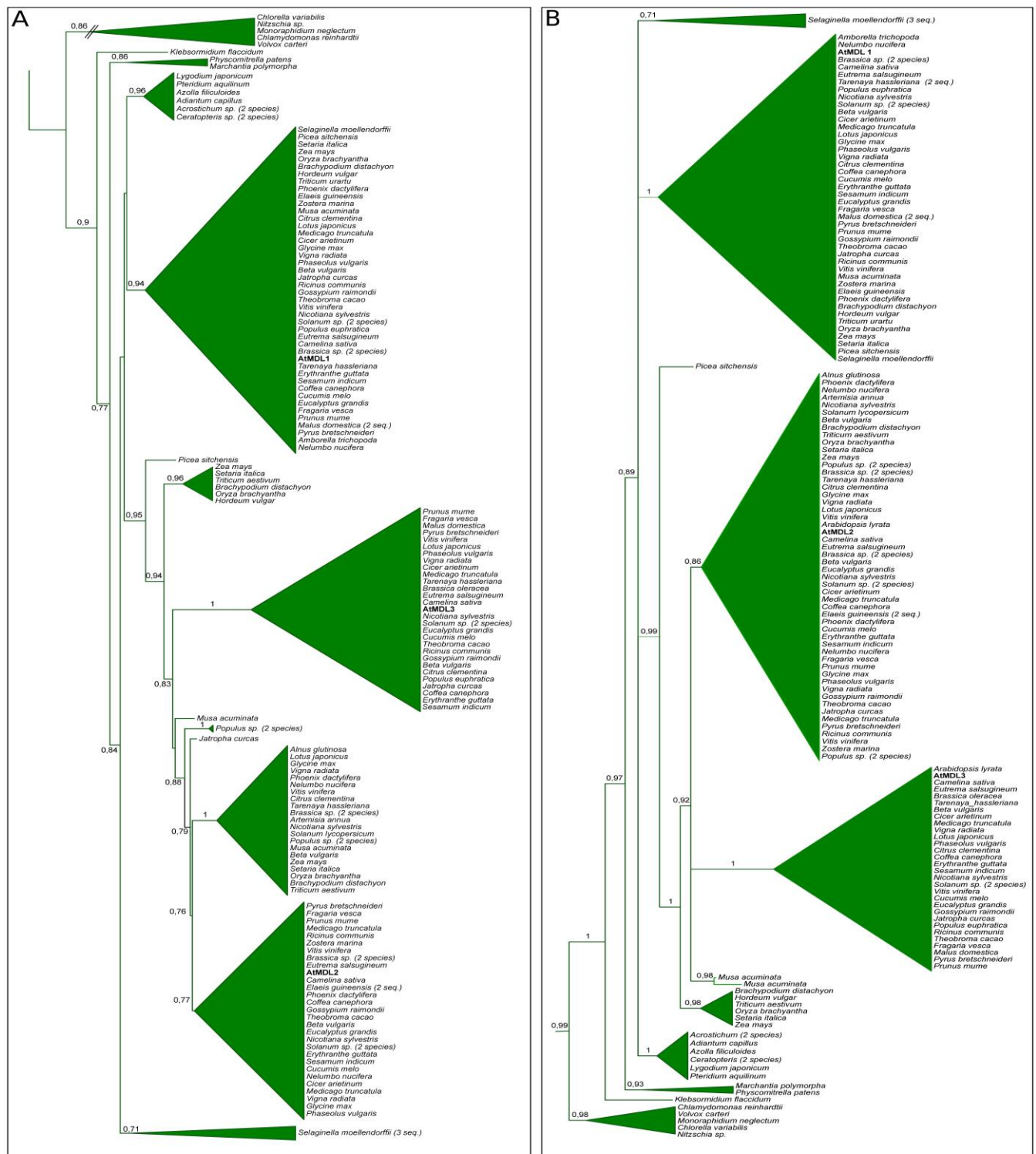


Figure 15 – Reconstruction phylogénétique des MIFs de Viridiplantae

Phylogénie réalisée via la méthode du maximum de vraisemblance (A) et la méthode Bayésienne (B). Les séquences correspondant aux protéines MIFs des *AtMDL* sont notées en gras. Les valeurs statistiques sont indiquées au-dessus des branches correspondantes si celles – ci sont supérieures à 0.5 ou 0.7, pour la méthode BI ou ML, respectivement.

V.1.2.3- Les MIFs de champignons

Cette étude a mis en évidence l'absence de protéines MIFs chez la majorité des champignons (Figure 16A). En effet, sur les 267 espèces étudiées, seules 30 % possèdent un MIF.

Il ne semble pas exister de lien entre la présence des MIF chez les champignons et un mode de vie spécifique (parasites obligatoires, vie libre, ...). Au sein d'un même taxon, comme les Bolétales (sous-famille du genre des *Agaricomycotina*), des espèces au mode de vie identique peuvent posséder ou non des MIFs. Ainsi, *Boletus edulus* et *Paxillus involutus* sont deux espèces de Bolétales, parasites de plantes mais seul *P. involutus* possède des protéines MIF.

De plus, l'arbre phlogénétique des MIFs chez les champignons qui possèdent ces protéines suit la phylogénie des espèces, basée sur les séquences d'ARN 18S.

La grande proportion de groupes de champignons ayant perdus les gènes MIFs, ainsi que la similarité entre la phylogénie des MIFs et celle des espèces correspondantes suggèrent une très faible importance fonctionnelle des MIFs chez les champignons, quel que soit leur mode de vie (libre ou parasitaire).

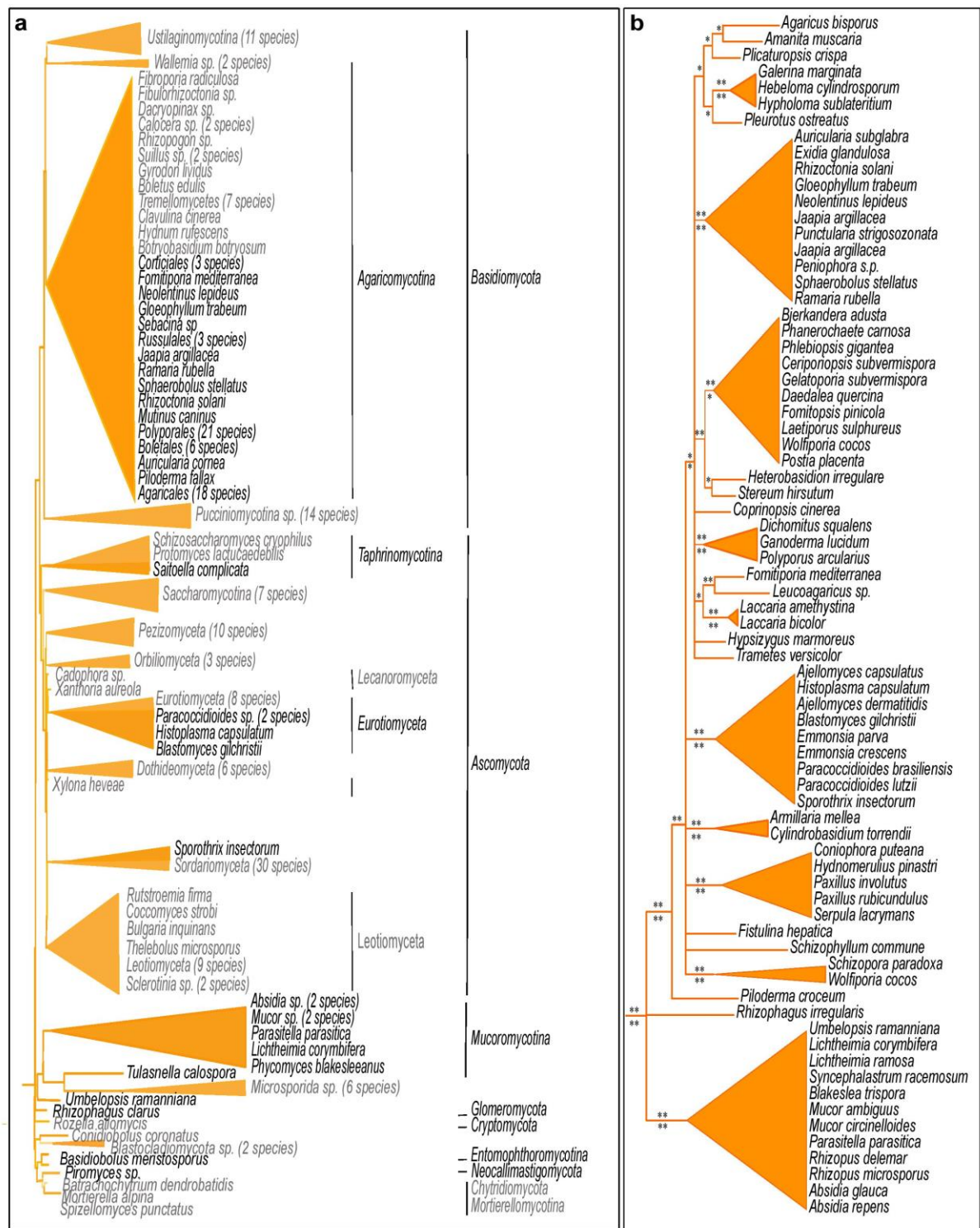


Figure 16 – Reconstruction phylogénétique des MIFs de champignons

Reconstruction phylogénétique (BI) des espèces basée sur les séquences des ARN18S (A) et des séquences MIFs (B). Les valeurs de confiance statistiques sont indiquées au-dessus des branches pour la méthode BI et en dessous des branches pour la méthode ML. Une étoile correspond à une valeur de confiance statistique supérieure à 0.5 ou 0,7 pour les valeurs de PP ou aLRT, respectivement. Deux étoiles correspondent à une valeur de probabilité supérieure à 0.7 ou 0,9 pour les valeurs de PP ou aLRT, respectivement.

V.1.2.4- Les MIFs de nématodes

L'arbre phylogénétique des nématodes établi sur la base des séquences d'ARN 18S correspond aux arbres précédemment établis, à l'exception des *Telenchima* (qui devraient constituer un groupe monophylétique) [168–171].

L'analyse a mis en évidence la présence de protéines MIFs chez l'ensemble des nématodes disposant d'un génome déjà séquencé et publié, à l'exception notable des espèces *Globodera pallida* et *G. rostochiensis*. L'analyse détaillée des génomes et des transcriptomes de ces espèces et des espèces proches permettent de conclure que les MIFs sont effectivement absentes des *Heteroderidae/Hoplolaimidae*. Cette observation est d'autant plus étonnante que les MIFs sont présentes en plusieurs exemplaires chez les espèces du genre *Meloidogyne* et les membres de la famille la plus proche, les *Pratylenchidae*. Ces informations indiquent que les gènes *MIF* ont été perdus chez l'ancêtre commun des *Heterodeidae*, *Rotylenchulidae* et *Hoplolaimidae*. L'ensemble d'espèces de ces familles de nématodes sont des parasites de plantes.

La phylogénie des MIFs de nématodes ne suit pas celle des espèces (Figure 17). Comme la majorité des métazoaires, les nématodes possèdent au moins deux copies de MIFs. L'une des ces deux copies regroupe avec le variant D-DT décrit chez les chordés. Toutefois, la seconde copie ne regroupe pas toujours avec le variant MIF des chordés.

Chez les nématodes, les gènes *MIF* semblent avoir subi plusieurs duplications successives, dont une ancienne duplication, datant vraisemblablement de la diversification des métazoaires (entre 757 et 1147 MYA). Certains phyla semblent avoir subi des duplications supplémentaires récentes. En effet, l'ensemble des *Meloidogyne* et des *Steinernema* disposent d'au moins deux ou trois copies de MIFs issues d'une duplication récente ayant vraisemblablement eu lieu peu de temps après la séparation de ce groupe des autres nématodes. Ainsi, cette duplication des MIFs chez les *Meloidogyne* date probablement de 78 MYA [171]. De plus, chez certaines espèces de *Meloidogyne*, comme *Meloidogyne incognita*, la sur-représentation de la protéine est également probablement liée à la duplication complète du génome [171]. Toutefois, la présence de deux protéines MIFs chez *Meloidogyne hapla*, espèce n'ayant pas subi de duplication de son génome, confirme que l'ensemble des *Meloidogynes* a au moins subi une duplication spécifique des protéines MIFs qui n'est pas liée à la duplication de leur génome.

Les analyses précédentes ne permettent pas de dater la séparation des *Steinernema* des autres espèces proches. Les dates ont été estimées à partir de la littérature déjà réalisée [171] sur ces espèces et via le logiciel Timetree [172].

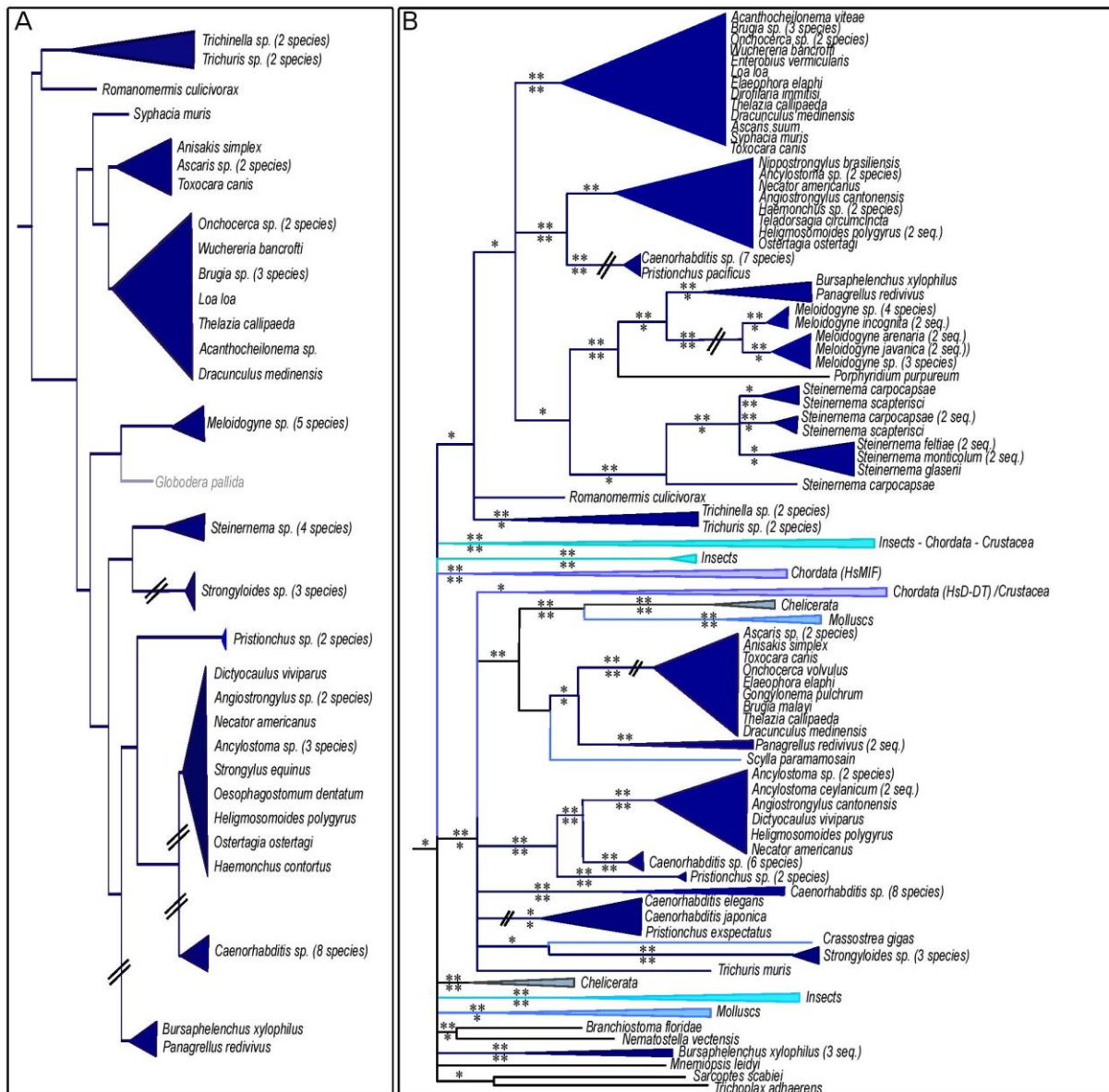


Figure 17 – Reconstruction phylogénétique des espèces de nématodes et des MIFs

Analyse par Inférence bayésienne des ARN18S des nématodes (A) et des protéines MIFs correspondantes (B). L'arbre des espèces des nématodes (A) est cohérent avec les phylogénies précédemment publiées et acceptées [168–170, 173]. *Globodera pallida* (en gris, en A) est la seule espèce de nématodes de cette analyse qui ne comporte aucune protéine MIF. Les valeurs de confiance statistiques sont indiquées au-dessus des branches pour la méthode BI et en dessous des branches pour le test de ratio de probabilité (test aLRT) réalisé pour la méthode de maximum de vraisemblance. Une étoile correspond à une valeur de confiance statistique supérieure à 0.5 ou 0,7 pour les valeurs de PP ou aLRT, respectivement.. Deux étoiles correspondent à une valeur de probabilité supérieure à 0.7 ou 0,9 pour les valeurs de PP ou aLRT, respectivement. L'arbre correspondant réalisé avec la méthode du maximum de vraisemblance est présent en Annexes.

V.1.2.5- Les MIFs d'insectes

Les insectes possèdent un nombre de protéines MIF variant de 0 à 4. De nombreux clades ne possèdent pas de MIFs, notamment les Hyménoptères, les Diptères, et les Phthiraptères. Cette absence indique que les gènes *MIF* ont probablement été perdus chez les ancêtres communs respectifs de ces différents clades. Il faut toutefois noter que chez les Phthiraptères, une seule espèce possède, à ce jour, un génome entièrement séquencé et annoté. Il est donc possible que seule cette espèce ne possède pas de MIF et que ce ne soit pas un phénomène caractéristique de l'ordre de *Phthiraptera* dans son ensemble.

La phylogénie des MIFs d'insectes est très différente de celle des espèces. Les Coléoptères, les Blattoptères et les Orthoptères possèdent tous une seule copie de la protéine MIF regroupant avec trois des groupes monophylétiques des Aphididées.

Les MIFs de Lépidoptères forment deux groupes distincts phylogénétiquement (notés L1 et L2 dans la Figure 18). Ces deux différents groupes ne regroupent pas avec les variants MIFs et D-DT présents chez les chordées. Les Aphididées possèdent de nombreuses copies de MIFs. Trois des copies de ces protéines regroupent préférentiellement entre elles (A1 à A3) et l'une est fortement séparée des autres (A4). Ces résultats suggèrent que les différentes protéines MIFs des pucerons sont issues de duplications successives ayant eu lieu avant la diversification des pucerons.

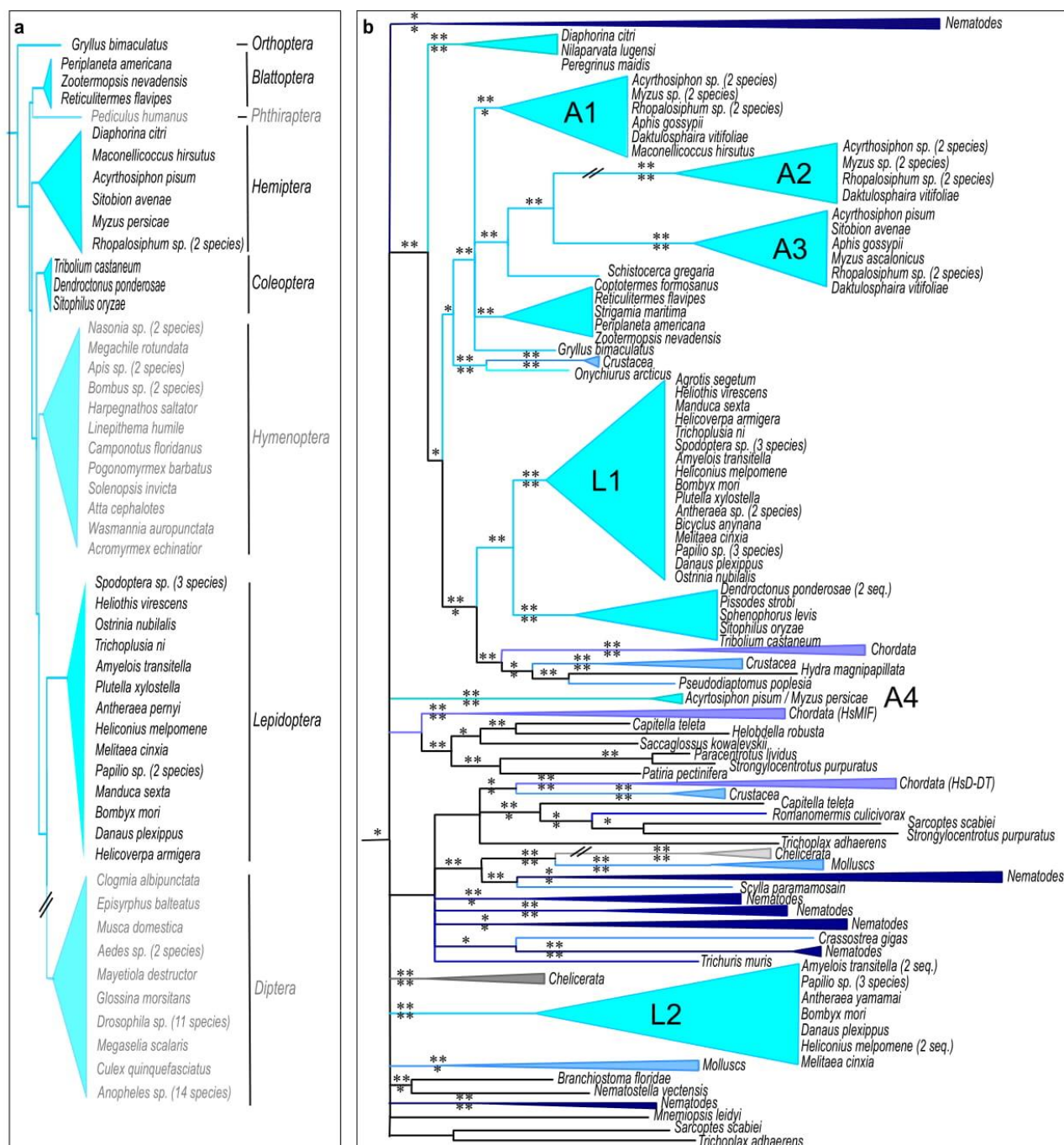


Figure 18 – Reconstruction phylogénétique des espèces d'insectes et des MIFs

Analyse bayésienne des ARN18S des insectes (A) et des protéines MIFs correspondantes (B). L'arbre des espèces des insectes (A) est cohérent avec les phylogénies précédemment publiées et acceptées [174]. Les espèces et les clades où aucune protéine MIF n'a pu être trouvée sont indiqués en gris. Les valeurs de confiance statistiques sont indiquées au-dessus des branches pour la méthode BI et en dessous des branches pour la méthode ML. Une étoile correspond à une valeur de confiance statistique supérieure à 0.5 ou 0,7 pour les valeurs de PP ou aLRT, respectivement. Deux étoiles correspondent à une valeur de probabilité supérieure à 0.7 ou 0,9 pour les valeurs de PP ou aLRT, respectivement. L'arbre correspondant réalisé avec la méthode du maximum de vraisemblance est présent en Annexes. Les groupes regroupant des pucerons et des lépidoptères ont été noté respectivement A1 à A4 et L1 et L2 pour aider à la visualisation.

V.1.2.6- Les MIFs des Straménopiles et des Alvéolates

Comme pour les insectes, le nombre de MIF par espèce chez les Straménopiles et les Alvéolates varie de 0 à 4. Ainsi, de nombreux phyla tels que les *Ciliophora* ne comportent aucune espèce possédant des MIFs. Les espèces ne possédant pas de MIF ont des modes de vie et appartiennent à des clades très différents. Ainsi, on ne retrouve pas de MIF chez des espèces libres telles que *Ectocarpus siliculosus* ou *Stylonychia lemnae*, ou encore des espèces parasites d'animaux telles que *Saprolegnia parasitica* ou *Babesia sp.*

Certaines espèces de Straménopiles et d'Alvéolates comprennent des MIFs dont la séquence est atypique ou incomplète (indiqué par une étoile dans la Figure 19). Ces séquences n'ont pas été incluses dans l'analyse phylogénétique.

La phylogénie des MIFs chez les Straménopiles et les Alvéolates correspond à celle des espèces (Figure 19), à l'exception des séquences de *Vitrella brassicaformis*, *Aureococcus anophagefferens*, *Aureoumbra lagunensis*, *Phaedactylum tricornutum* et *Thalassiosira oceanica*. Les MIFs de ces espèces sont dans un même groupe supporté par les deux méthodes phylogénétiques (Maximum de vraisemblance et Inférence Bayésienne) à l'extérieur du groupe des MIFs de Straménopiles et d'Alvéolates.

Au sein des Straménopiles, les MIFs issues du genre *Pythium* et *Phytophthora* ne forment pas un seul groupe monophylétique, contrairement à l'arbre des espèces. Cette topologie suggère une duplication des gènes *MIF* avant la diversification des *Pythiales* et des *Peronosporales* – estimée à environ 80 MYA (d'après TimeTree). De plus, au moins une autre duplication aurait eu lieu plus récemment formant ainsi les nombreuses copies des protéines MIFs que l'on observe au sein de ces deux genres.

Il est à noter que les différentes protéines analysées ont été séparées par motif MIF lors de l'étude des *Pythium* et des *Phytophthora*, pour la bonne réalisation de la phylogénie. En effet, sur les 10 espèces de *Pythium* et de *Phytophthora* analysées en détail, cinq ne présentent pas deux protéines distinctes comme les autres mais une seule, composée de deux motifs MIFs, séparés par seulement quelques acides aminés. Chez les autres espèces de ces deux genres, les deux gènes codant les MIFs ne sont séparés que par quelques paires de bases. Il est donc probable que cette observation ne soit due qu'à une erreur d'annotation. Toutefois, la structure spécifique des gènes *MIF* chez les *Pythium* et les *Phytophthora* reste intéressante à analyser dans le cadre d'une future analyse fonctionnelle de la protéine chez ces Straménopiles.

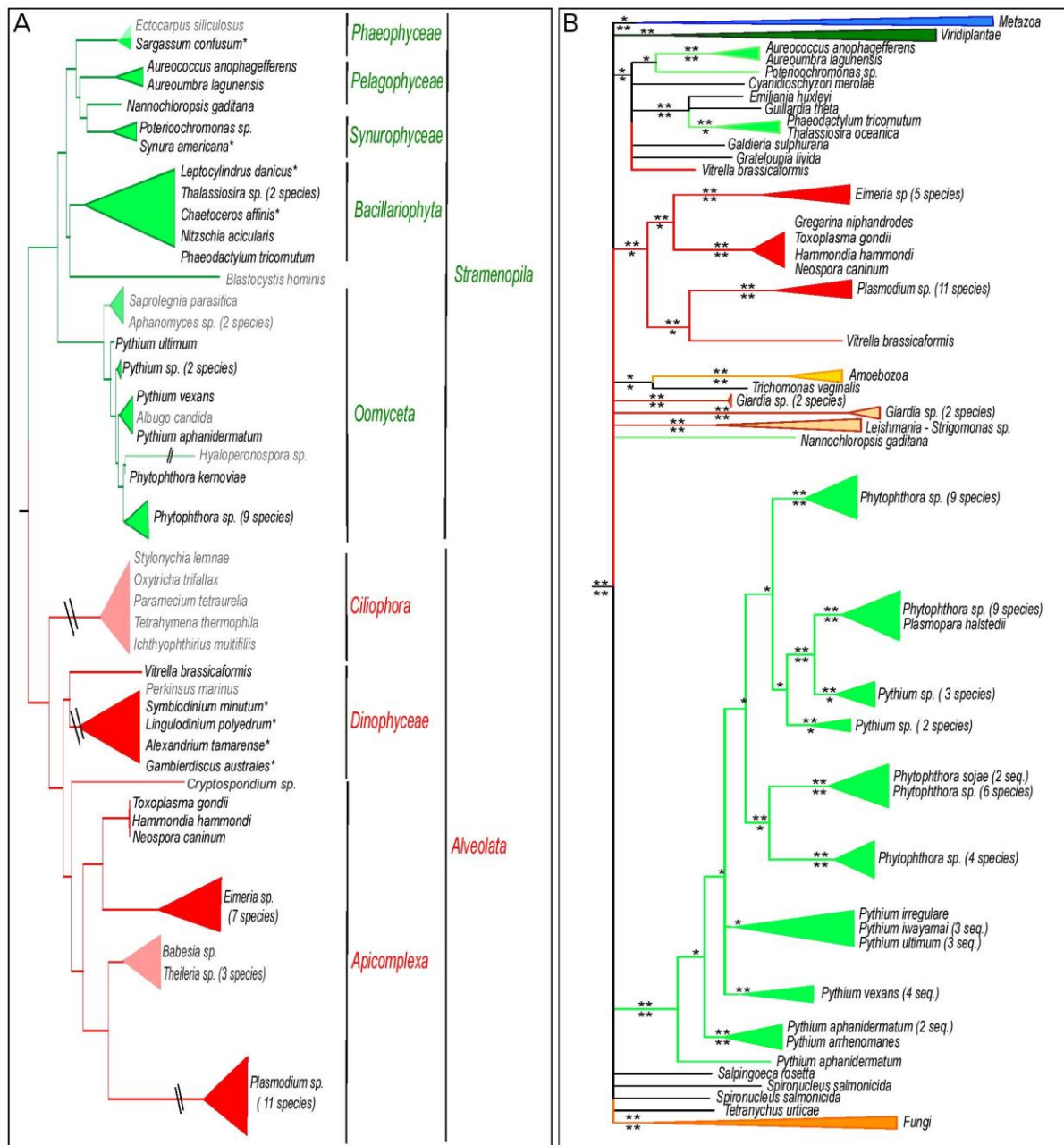


Figure 19 – Reconstruction phylogénétique des espèces et des MIFs d’Alvéolates et Straménopiles

Analyse bayésienne des ARN 18S des Straménopiles et des Alvéolates (A) et des protéines MIFs des espèces correspondantes (B). Les espèces où aucune protéine MIF n’a pu être trouvée sont indiquées en gris. Les branches de la phylogénie qui ne respectent pas l’échelle évolutive ont été marquées du symbole \\. Les séquences partielles ou atypiques sont notées d’une étoile (en A). Les valeurs de confiance statistiques sont indiquées au-dessus des branches pour la méthode BI et en dessous des branches pour la méthode ML. Une étoile correspond à une valeur de confiance statistique supérieure à 0.5 ou 0.7 pour les valeurs de PP ou aLRT, respectivement. Deux étoiles correspondent à une valeur de probabilité supérieure à 0.7 ou 0.9 pour les valeurs de PP ou aLRT, respectivement. L’arbre correspondant réalisé avec la méthode du maximum de vraisemblance est présent en Annexes.

V.1.3- Analyse des sites sous sélection

Pour déterminer si un motif, ou des acides aminés particuliers des MIFs seraient spécifiques des parasites d'animaux ou de plantes, j'ai analysé en détail les sites sous sélection positive (aussi appelée diversifiante) ou négative (aussi appelée purificatrice) chez cinq taxons différents : (i) les Alvéolates, parasites d'animaux, (ii) les nématodes, parasites de plantes ou d'animaux, (iii) les Straménopiles, parasites de plantes, et (iv) les Aphididées en tant que seuls insectes parasites de plantes présents dans l'étude. A titre de comparaison, les séquences de plantes ont aussi été étudiées.

Les résultats combinés des cinq logiciels utilisés indiquent que 94 sites sur les 122 présents sur les séquences MIFs des plantes sont prédits comme étant soumis à une pression de sélection purificatrice. Les séquences des MIFs de plantes sont donc fortement conservées dans l'ensemble du règne des *Viridiplantae*. A l'inverse, les séquences MIFs des autres groupes possèdent beaucoup moins de sites soumis à sélection purificatrice (de 31 à 75 sites différents).

Seuls six sites au total sont soumis à une sélection diversifiante (ou positive). De manière intéressante, en position 23, on retrouve un acide aminé qui est sous sélection positive à la fois chez les pucerons et chez les plantes alors qu'il est sous sélection purifiante chez les Alvéolates et les Straménopiles.

L'analyse des différents sites soumis à sélection, en fonction du mode de vie des différents taxons, ne permet pas de déterminer des acides aminés qui seraient spécifiques du mode de vie « parasite de plantes » ou « parasite d'animaux ». Toutefois, certains sites sous sélection sont différents au sein d'un même taxa (les nématodes notamment) ou correspondent à des acides aminés différents entre les espèces parasites de plantes et les espèces parasites d'animaux. L'acide aminé 101, par exemple, correspond à une asparagine (N) pour les nématodes parasites de plante et à un acide aspartique (D) pour les nématodes parasites d'animaux. Toutefois, ces différences ne se retrouvent pas entre tous les taxons ayant le même mode de vie.

Les acides aminés sous sélection ont été analysés au regard des connaissances actuelles sur les sites actifs des protéines MIFs humaines. L'ensemble des sites ayant été décrits dans la littérature comme étant liés à une fonction spécifique ont donc été représentés (Figure 20).

Il est intéressant de constater que les sites connus chez l'homme pour être nécessaires à la fonction tautomérase, la proline – 2, la lysine – 33, l'isoleucine – 65, la tyrosine – 37, la tyrosine- 96 et le tryptophane – 109 sont conservés chez la majorité des pucerons et des plantes, ce qui pourrait indiquer une conservation de la fonction tautomérase chez ces deux groupes.

La proline 2 nécessaire à cette fonction est le seul acide aminé sous sélection conservatrice pour la majorité des séquences. La majorité des autres sites nécessaires à cette fonction sont conservés pour la majorité des taxons. La présence de cette activité enzymatique chez les bactéries possédant des protéines MIFs, et la conservation de ces sites chez la majorité des eucaryotes montre l'importance fonctionnelle de cette fonction et suggère qu'il puisse s'agir de la fonction originelle de ces protéines.

La proline – 2 a aussi été décrite comme étant impliquée dans la fixation des MIFs au récepteur CD74 et CXCR4 chez l'homme. Même si les récepteurs CD74 et CXCR4 ne sont pas conservés dans l'ensemble des eucaryotes, la forte conservation de cette proline pourrait être également liée à la fixation des MIFs à de potentiels récepteurs chez d'autres espèces.

A l'exception des sites liés à la fonction tautomérase, l'ensemble des autres sites décrits chez l'humain comme étant liés à une fonction spécifique semble faiblement conservé chez les autres espèces. Le motif CXXC (position 57 à 60) est, par exemple, uniquement présent dans le groupe des nématodes parasites d'animaux. Ce site a été décrit précédemment comme étant lié à la régulation de la réponse au stress cellulaire (maintien de l'homéostasie, régulation de la concentration en peroxyde d'oxygène, ...) [72] et comme étant nécessaire à l'activité oxidoréductrice des MIFs. L'absence du motif CXXC s'observe chez la majorité des espèces eucaryotes étudiées. Cette observation semble indiquer que les fonctions dépendantes de ce site (oxydo-réduction, homéostasie, ...) ne sont pas les fonctions primaires – évolutivement - des protéines MIFs. Toutefois, des études récentes ont indiqué que la présence partielle du motif CXXC pouvait être suffisant pour la conservation de la fonction (conservation de la Cystéine 60 uniquement) [175, 176]. Des études dédiées seraient donc nécessaires pour vérifier l'activité d'oxido-reduction de ces MIFs.

Le motif « pseudo (E)LR » ressemble au motif ELR caractéristique des cytokines de type ELR+ CXC. Ce motif a été décrit précédemment comme étant nécessaire à la fixation de la protéine MIF sur son récepteur CXCR2. Il est absent chez l'ensemble des espèces testées pour cette analyse. A l'inverse, le motif N-like-loop (position 47 à 56) est partiellement conservé. Celui-ci a été décrit précédemment comme étant nécessaire à la liaison des MIFs au récepteur CXCR4. Toutefois, chez l'homme, l'activation du récepteur CXCR4 nécessite aussi la participation du motif RLR (en position 87 à 89). Ce motif n'est que partiellement conservé chez l'ensemble des espèces étudiées ici.

De la même manière, les sites nécessaires à la fonction nucléase [73] chez l'homme ne semblent que très peu conservés chez les espèces analysées ici. Ainsi, le motif PD/ExxxE nécessaire à cette fonction ainsi que le domaine zinc finger CxxCxxHx(n), domaine commun de réponse aux dommages à l'ADN, nécessaire lui aussi à la fonction nucléase, sont absents chez l'ensemble des espèces étudiées.

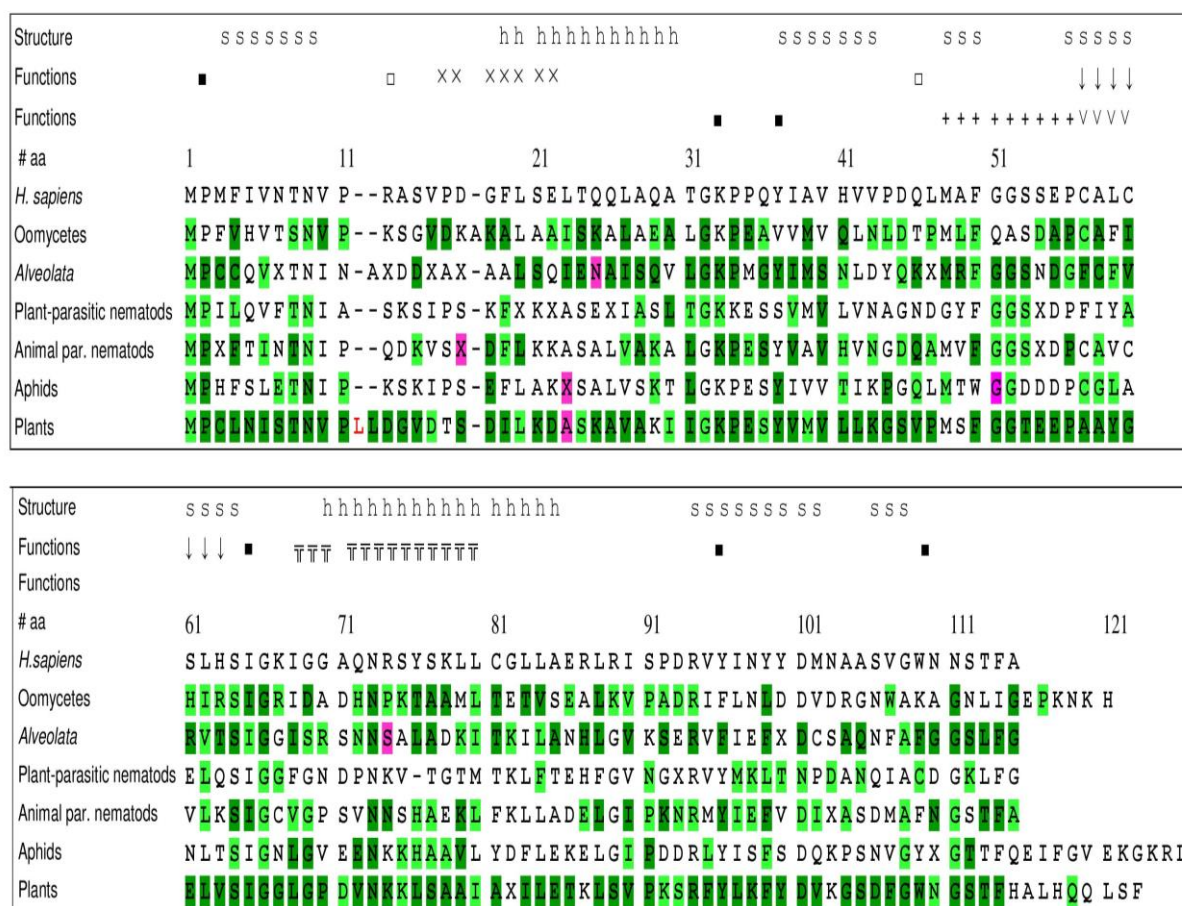


Figure 20 – Analyse des sites sous sélection positive ou négative

Alignement de la séquence du MIF humain (code à une lettre des acides aminés) et les séquences consensus des oomycètes, des Alvéolates, des nématodes parasites de plantes, des nématodes parasites d'animaux, des pucerons et des plantes. Les séquences consensus ont été déterminées via le logiciel Jalview version 2.9 et montrent l'acide aminé le plus fréquent à chaque position ou un X lorsque la fréquence de l'acide aminé le plus fréquent est inférieure à 40%. La numérotation est basée sur la séquence de la protéine MIF humaine. Les traits d'union (-) indiquent un gap dans les séquences. Les sites sous sélection purificatrice ou diversifiante sont notés en vert et rose respectivement. Le vert clair et le vert foncé indiquent un site sous sélection purifiante, détecté par, respectivement, deux à trois ou quatres à cinq logiciels différents. Les hélices alpha et les feuillets bêta de la protéine humaine sont indiqués respectivement par les lettres « h » et « s ». Les symboles représentent les différents sites de la protéine MIF considérés comme importants pour certaines fonctions ou activités chez l'homme. Ces sites permettent à la protéine d'avoir une activité tautomérase (symbole ■), une activité oxidoréductase (motif CXXC, symbole ∨), une activité nucléase (motif PE/DxxxxE, symbole ×) mais aussi de se fixer aux récepteurs CXCR4 (symbole ☐), aux récepteurs CXCR2 (motif pseudo ELR, symbole ☐). Ils permettent les interactions entre la protéine MIF et le récepteur CXCR2 (pattern N like loop, symbole +) et sont une zone de réponse aux dommages à l'ADN (domaine zinc finger CxxCxxHx(n), symbole ↓).

V.1.4- Conclusion

En conclusion de ce chapitre, nous avons pu voir que le nombre de protéines MIFs variait beaucoup d'une espèce ou d'un phylum à l'autre au sein des eucaryotes. Ainsi, les espèces autotrophes et les parasites de plantes possèdent en moyenne plus de protéines MIFs que les autres espèces (en moyenne, 3). Ce biais de répartition du nombre de protéines MIFs nous suggère un rôle important du nombre de protéines MIFs pour les espèces parasites de plantes et les plantes. De plus, nous avons pu voir que les protéines MIFs avaient une histoire évolutive complexe comportant de nombreux événements indépendants de duplications et de perte de gènes.

L'analyse des sites sous sélection purificatrice ou diversifiante a permis de caractériser la forte conservation des MIFs au sein de l'ensemble des *Viridiplantae* (les organismes verts). Cette conservation, au sein d'espèces taxonomiquement et écologiquement aussi éloignées que des algues vertes (*Chlamydomonas reinhardtii*) ou le palmier (*Phoenix dactylifera*) souligne l'existence probable de contraintes fonctionnelles très fortes pour les MIFs des organismes verts. Les futures études fonctionnelles devraient permettre de caractériser les fonctions des MIFs de *Viridiplantae* et de mieux comprendre la nature de ces contraintes fonctionnelles et évolutives.

Cette analyse n'a pas mis en évidence de sites ou motifs spécifiques des espèces parasites de plantes ou d'animaux. Ce résultat pourrait être lié au plus faible nombre de séquences utilisées pour les différents groupes. L'augmentation constante des bases de données de séquence permettra de répéter cette analyse avec un plus grand nombre de séquences dans un futur proche.

Parallèlement, cette analyse a aussi montré la faible conservation, dans l'ensemble du règne animal, des sites nécessaires aux fonctions des MIFs décrites chez l'humain. Pourtant, de nombreux parasites d'animaux (*Plasmodium*,...) ont été démontrés comme partageant des fonctions communes avec l'humain. Ces résultats semblent indiquer que

- (i) Les protéines MIF des parasites pourraient agir sur l'immunité de leurs hôtes par des mécanismes encore inconnus impliquant des sites (AA) qui n'ont pas encore été identifiés
- (ii) La conservation de la structure des MIFs, seule, pourrait expliquer la conservation de certaines fonctions majeures comme celles liées à la régulation de l'immunité.

Chapitre 2 : La modulation de la mort cellulaire par les MIFs

L'ensemble des analyses *in planta* ont été réalisées en collaboration avec Maelle Jouannet, post doctorante dans l'équipe et Arnaud de Ruffray, stagiaire de l'équipe.

V.2.1- Introduction

Les analyses de séquences réalisées dans le chapitre 1 n'ont pas révélé de motifs spécifiques et discriminants des différents groupes, en particuliers des parasites d'animaux ou de végétaux. De plus, à l'exception des motifs liés à l'activité tautomérase, les motifs liés aux autres fonctions majeures du MIF humain ne sont pas conservés chez les MIFs des différents groupes (végétaux, métazoaires parasites d'animaux ou de végétaux). La question des activités biologiques des différents MIFs reste donc entière et devra être abordée par des études fonctionnelles.

Pour ces études, nous nous sommes focalisés sur l'action des MIFs dans la régulation de la mort cellulaire. En effet, rappelons que chez l'homme, les protéines MIFs interfèrent avec la voie des p53, voie d'induction de l'apoptose des cellules. Elles contribuent donc à limiter l'apoptose, notamment celle induite par le NO lors d'une réponse immunitaire, ce qui favorise cette réponse en préservant les cellules immunitaires [41]. Par ailleurs, les protéines MIFs humaines ont récemment été décrites comme capables d'exercer une fonction nucléase, pouvant, elle, conduire à la mort de la cellule par parthanatose [73]. Ces protéines sont donc capables soit de favoriser, soit d'empêcher la mort cellulaire, par le biais d'activités biologiques différentes.

Notre équipe a récemment montré que les protéines MIFs de pucerons étaient capables d'inhiber la mort cellulaire induite par la cryptogéine chez la plante [20]. Chez les végétaux, cette mort cellulaire (stade final de la réponse hypersensible) est un mécanisme de défense qui permet d'éviter la propagation des agents pathogènes. Les MIFs de pucerons sont donc capables d'inhiber ce mécanisme de défense, ce qui joue un rôle majeur dans l'interaction hôte-parasite, et le succès de l'infection [20]. Ces résultats montrent ainsi que des MIFs peuvent inhiber la mort cellulaire indépendamment de la voie p53, qui n'existe pas chez les végétaux.

Ces résultats soulèvent donc la question de la conservation de l'action des MIFs sur la mort cellulaire, en système animal ou végétal.

Pour explorer cette question, notre objectif global est de tester l'effet des protéines MIFs de plusieurs espèces phylogénétiquement et écologiquement éloignées sur la modulation de la mort cellulaire, en système animal ou végétal. Dans le cadre de cette thèse, seules les analyses de mort cellulaire en cellule végétale sont présentées car les tests en cellule animale sont encore en cours.

V.2.2- Choix des modèles biologiques

Pour cette étude, neuf espèces, représentées par 20 séquences différentes de MIFs ont été choisies comme représentatives, à la fois, d'une grande diversité de taxons, mais aussi de modes de vie (libre, ou parasite de plantes ou d'animaux) différents.

Les protéines MIFs choisies sont les suivantes :

- MiMIF1 et MiMIF2, de *Meloidogyne incognita*. Cette espèce est un nématode parasite de plantes, qui provoque des galles racinaires. Son génome est composé d'une duplication complète et de plusieurs duplications partielles. Ainsi, il existe 5 gènes *MIF* dans le génome de *Meloidogyne incognita*, mais le gène MiMIF1.2 est identique à MiMIF1 à 99% (donc non testé) et le gène MiMIF2.2 est identique à MiMIF2 à 96% (donc non testé). Le cinquième gène, lui, n'est que partiel. Les deux gènes choisis sont situés sur le même scaffold, à une faible distance génique. Ces deux gènes étant probablement sous le contrôle d'un même promoteur, ils ont également été testés pour leur potentiel effet synergique sur la modulation de la mort cellulaire.
- BxMIF1 et 2, de *Bursaphelenchus xylophilus*. Cette espèce est aussi un nématode parasite de plantes, ayant pour vecteur de nombreux insectes tels que les longicornes. C'est une espèce invasive, occasionnant d'important dégâts sur les pins notamment en Amérique. En dehors de l'impact écologique de l'espèce, les deux protéines ont aussi été choisies pour leur position dans les reconstructions phylogénétiques. En effet, BxMIF2 groupe avec les MIFs des nématodes parasites de plantes de type *Meloidogyne*, alors que BxMIF1 groupe avec les MIF des insectes (Figure 17). Il est donc particulièrement intéressant de comparer la potentielle activité sur la mort cellulaire de ces deux protéines. *Bursaphelenchus xylophilus* possède trois autres protéines MIF, mais celles-ci (BxMIF3 à 5) diffèrent très peu de BxMIF1 et ont la même localisation dans la phylogénie. Elles n'ont donc pas été testées.
- WbMIF1 de *Wuchereria bancrofti*, un nématode parasite qui provoque la filariose de Bancroft chez l'homme. Ce parasite provoque un épaississement de la peau et une hypertrophie des tissus cutanés due à l'obstruction progressive de la circulation lymphatique [177]. Cette pathologie peut conduire au développement d'un éléphantiasis (pachydermie d'aspect monstrueuse due à l'hypertrophie marquée de tissus cutanés) et au développement parallèle d'infections bactériennes (streptocoques, ...) dues aux nombreuses lésions cutanées. Les MIFs de cette espèce sont sécrétées au cours de l'infection par le parasite et sont nécessaires à la manipulation de l'immunité de l'hôte vertébrés par le parasite [11, 116, 175].
- CeMIF1 de *Caenorhabditis elegans*. Cette espèce est le nématode modèle. C'est une espèce libre, bactériophage. Elle a été choisie à titre comparatif afin de déterminer si l'inhibition ou l'activation de la mort cellulaire est spécifique des espèces parasites.
- SiMIF1, SiMIF2, SiMIF3 et SiMIF4 de *Solanum lycopersicum*. La tomate est une plante d'intérêt agronomique majeur, en particulier dans le sud de la France. Pour cette solanacée, nous disposons actuellement d'un grand nombre de données génomiques et transcriptomiques qui confirme l'existence des protéines MIFs, mais il n'existe pas d'étude fonctionnelle sur leur rôle. Ces protéines ont été choisies pour tester celles d'une plante (libre, autotrophe).
- AtMDL1, AtMDL2 et AtMDL3 d'*Arabidopsis thaliana*. Cette espèce est la plante modèle en biologie végétale. Tout comme pour *Solanum lycopersicum*, nous disposons actuellement d'un très grand nombre de données génomiques et transcriptomiques. Les analyses au laboratoire ont pu montrer qu'il existait deux variants d'épissage issus des

gènes *atmdl1* et 2. Les protéines correspondantes ont donc été notés AtMDL1.1, 1.2, 2.1 et 2.2.

- PiMIF1 et PiMIF2 de *Phytophthora infestans*. Cet oomycète parasite de plantes est responsable du mildiou sur les solanacées, notamment de la pomme de terre et la tomate. L'ensemble des *Phytophthora* analysés dispose de deux copies de MIFs issues d'une duplication récente. Les deux gènes codant pour les protéines MIF chez ces espèces sont très proches et parfois organisés en opéron. Ces deux protéines ont donc également été étudiées pour leur potentiel rôle synergique dans l'inhibition de la mort cellulaire.
- PfMIF de *Plasmodium falciparum*. Il s'agit de l'agent infectieux responsable du paludisme. Le paludisme est la plus importante parasitose humaine à l'échelle mondiale, responsable de plus de 400 000 morts par an (Chiffre de l'OMS, 2015). Il est estimé que près de la moitié de la population mondiale est exposée au risque du paludisme. PfMIF est capable d'agir sur l'immunité humaine, notamment en induisant une diminution du nombre de lymphocytes T [15]. Tout comme l'ensemble des parasites d'animaux chez les Alvéolates, *P. falciparum* ne possède qu'un seul gène *MIF*.
- HsMIF d'*Homo sapiens*. Cette séquence étant la plus étudiée et fonctionnellement caractérisée, elle a été intégrée dans l'analyse à titre de contrôle (puisque son action sur la mort cellulaire en système animal est mieux connue).

Pour une meilleure compréhension des résultats, une courte phylogénie a été réalisée ne comprenant que les MIFs sélectionnées pour l'analyse (Figure 21).

Toutes ces protéines MIFs ont donc été testées pour leur éventuel rôle dans la modulation de la mort cellulaire, à la fois sur l'induction et l'inhibition de celle-ci en cellule végétale.

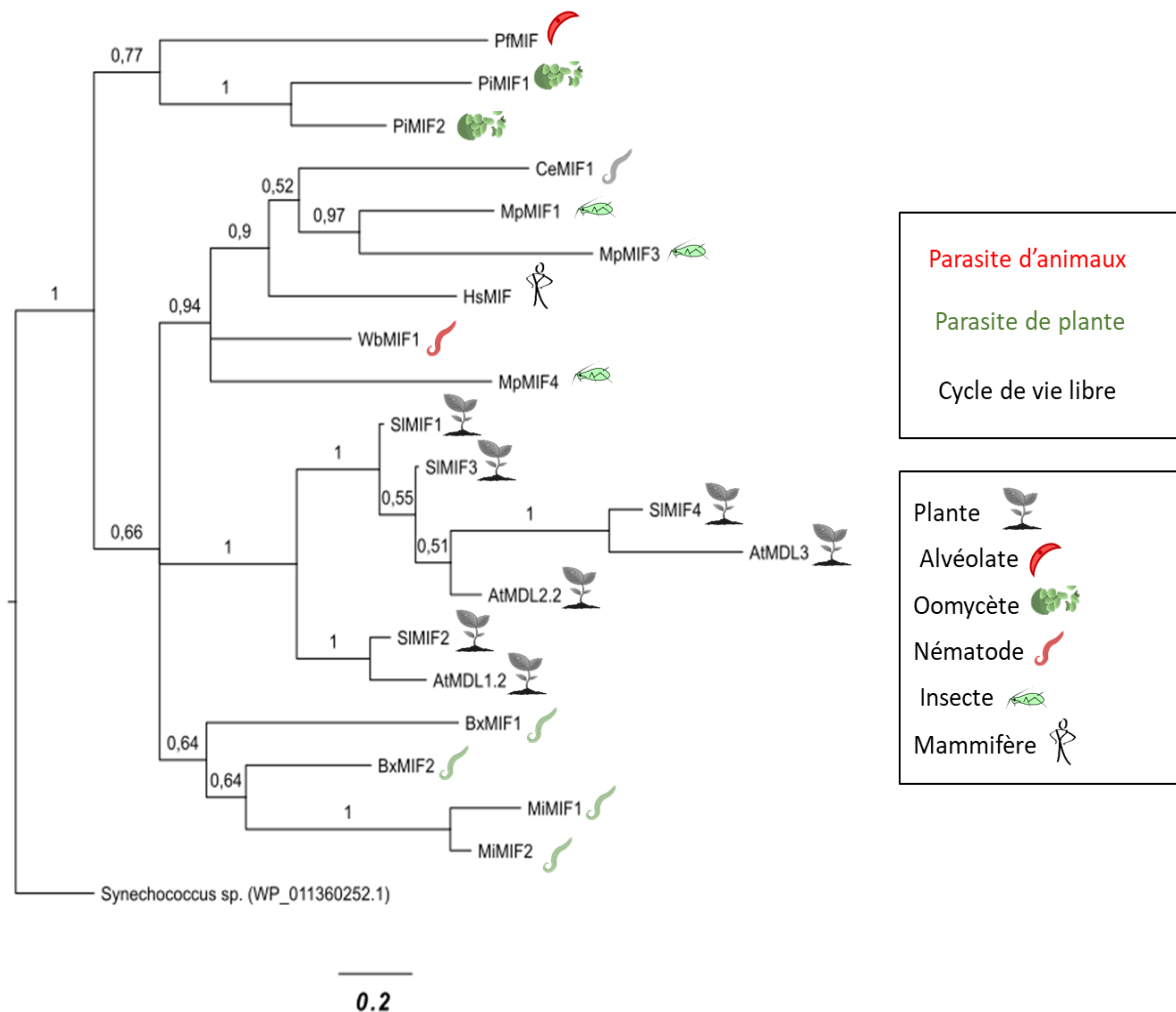


Figure 21 – Phylogénie des protéines MIFs testées pour leur rôle dans la modulation de la mort cellulaire

Reconstruction phylogénétique basée sur l'inférence Bayésienne. La séquence de la protéine MIF bactérienne de *Synechococcus* a été utilisée pour enraciner l'arbre. Les probabilités d'existence des branches sont indiquées au-dessus de celles-ci. Cette phylogénie restreinte suit la phylogénie générale présentée en Figure 14.

V.2.3- Effet des MIFs sur la modulation de la mort cellulaire en modèle végétal

V.2.3.1- Induction de la mort cellulaire

L'expression transitoire des différentes MIFs en feuilles de tabac montre une absence d'effet inducteur de la mort cellulaire pour toutes les protéines testées (Figure 22). Notons qu'une plus forte valeur du taux de chlorophylle indique un meilleur état physiologique. Ainsi, les seules différences significatives observées entre des MIFs et le contrôle GFP sont observées pour SIMIF1 et 3 et vont dans le sens d'une plus faible mort cellulaire. Un effet négatif sur la mort cellulaire serait possible du fait que l'agrotransfection elle-même (y compris avec la construction contrôle GFP) est susceptible d'induire un faible taux de mort cellulaire [178].

Une potentielle synergie entre les protéines MiMIFs, PiMIFs et BxMIFs n'est pas observée non plus.

Nous pouvons donc conclure que les protéines MIFs des différents organismes n'induisent pas de mort cellulaire en cellule végétale.

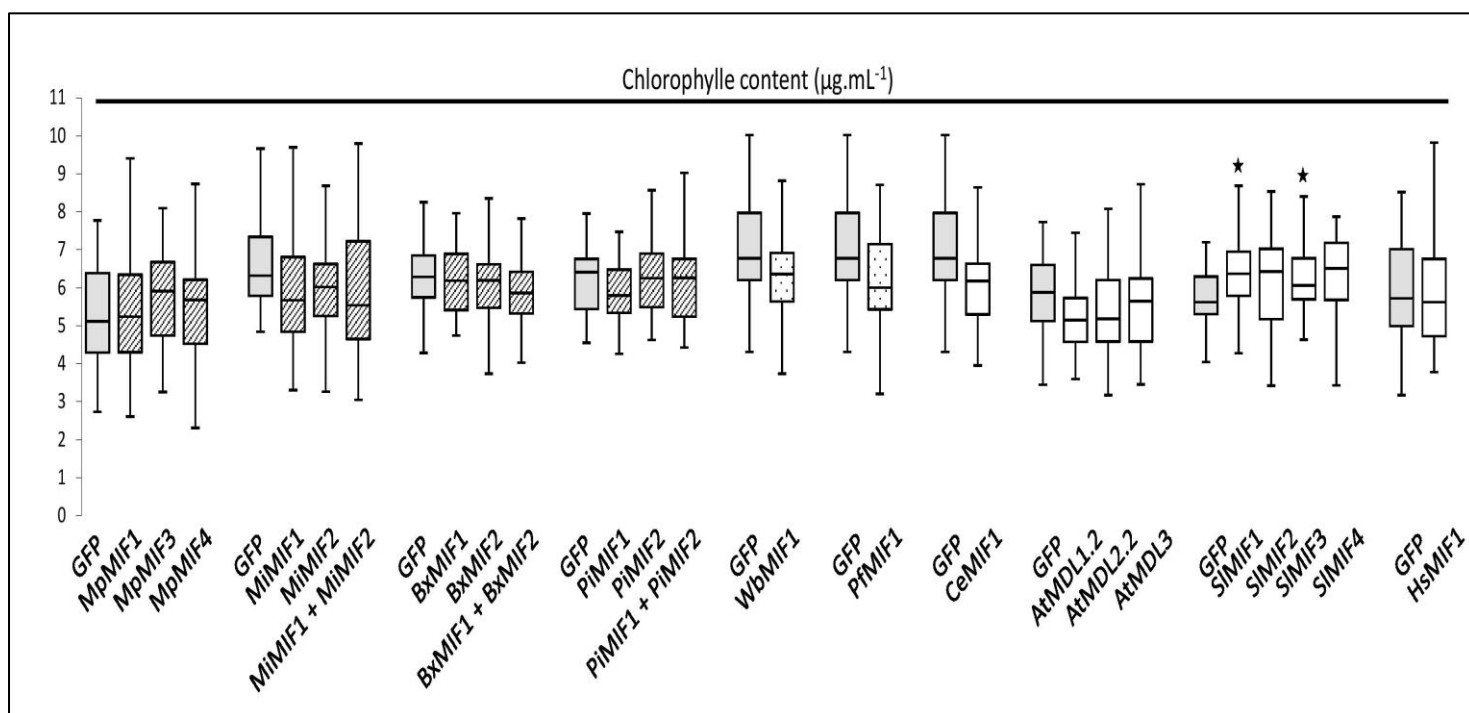


Figure 22 – Mesure d'induction de la mort cellulaire par les protéines MIFs en cellule végétale

Taux de chlorophylle des échantillons végétaux après surexpression de la GFP ou des protéines MIF. Des potentiels effets synergiques des protéines ont été testés pour les protéines MiMIFs, PiMIFs et BxMIFs. Les différences significatives entre les tissus après expression d'un gène contrôle (GFP) et des gènes codant les différentes protéines MIFs (test T de Student) sont indiquées par une étoile pour une p-value < 0,01.

V.2.3.2- Inhibition de la mort cellulaire

Pour analyser une potentielle inhibition de la mort cellulaire, nous avons induit la mort cellulaire par l'effecteur AvrPto [179] et co-infiltré les différentes protéines MIFs. La co-expression d'AvrPto avec la GFP (protéine non-relevante) a été utilisée en tant que contrôle pour l'induction de mort cellulaire en absence d'inhibiteur. Toutes les constructions MIFs testées ont été comparées avec ce contrôle.

Les analyses montrent que les protéines MIFs testées inhibent la mort cellulaire de façon significative (Figure 23), à l'exception de la protéine humaine HsMIF, de la protéine BxMIF2 de *Bursaphelenchus xylophilus*, de la protéine PiMIF1 de *Phytophthora infestans* ainsi que des protéines AtMDL1 et 2 d'*Arabidopsis thaliana*.

En ce qui concerne les MIF de *B. xylophilus* (BxMIF), de *Phytophthora infestans* (PiMIF), et d'*Arabidopsis thaliana* (AtMDL), notons que ces trois espèces possèdent, en plus des MIFs n'ayant pas d'effet, une protéine MIF ayant un effet significatif sur la répression de la mort cellulaire (Figure 23). Pour l'être humain, il serait nécessaire de compléter cette analyse en testant l'effet potentiel du variant HsD-DT, afin de vérifier si toutes les espèces possèdent au moins un MIF capable de réprimer la mort cellulaire.

A l'exception de l'être humain, toutes les espèces testées possèdent donc au moins un MIF capable d'inhiber la mort cellulaire en cellule végétale, y compris les protéines MIFs issues d'organismes qui ne sont pas en interaction avec les plantes comme le protiste *Plasmodium falciparum* ou le nématode libre *C. elegans*.

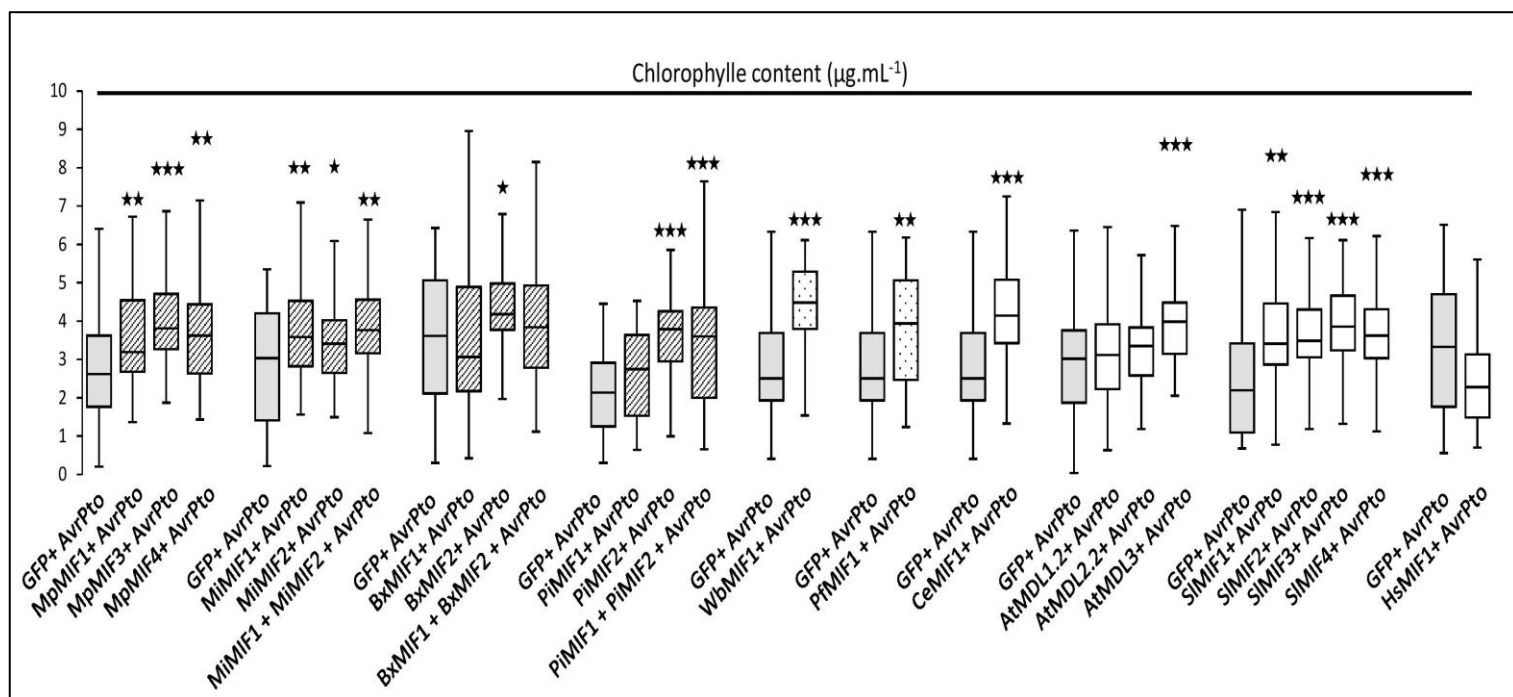


Figure 23 - Inhibition de la mort cellulaire des cellules végétales par les protéines MIFs

Taux de chlorophylle des échantillons végétaux après surexpression de la GFP ou des protéines MIFs en complément de l'inducteur de la mort cellulaire AvrPto. Des effets synergiques entre protéines proches d'une même espèce ont été testés pour PiMIFs, MiMIFs et BxMIFs. Les différences significatives (test T de Student) sont indiquées par une étoile pour une p-value significative à 5%, deux étoiles pour une p-value significative à 1% et trois étoiles pour une p-value significative à 1 pour 1000.

Sur le plan fonctionnel, ces résultats restent intrigants et méritent des études fonctionnelles poussées pour déterminer le mode d'action précis de ces protéines chez les cellules végétales. Notons que les travaux précédents de l'équipe avaient montré l'effet inhibiteur des MIFs de pucerons sur une mort cellulaire induite par la cryptogéine [20]. Contrairement à AvrPto qui est un effecteur, la cryptogéine est un PAMP reconnu par les plantes. Même si les voies et les potentiels types de mort cellulaire sont encore mal décrits chez les plantes [84], il est intéressant de constater que cet effet inhibiteur est efficace quel que soit l'inducteur de mort cellulaire. Enfin, il semble impossible de transposer nos connaissances sur le rôle inhibiteur de mort cellulaire dépendant de p53, bien connu chez les MIFs de mammifères, à un modèle plante où ces mêmes voies n'existent pas (e.g. absence de la protéine TP53 nécessaire à l'inhibition de l'apoptose par les MIFs *via* la voie p53 dépendante).

Sur le plan évolutif, les résultats suggèrent que l'inhibition de la mort cellulaire n'est pas une adaptation spécifique des parasites de plantes, mais résulte probablement de paramètres conservés (motifs, structure, paramètres physicochimiques) des MIFs.

V.2.4- Localisation subcellulaire des MIFs

La régulation de la mort cellulaire par les MIFs pourrait être associée à la localisation subcellulaire des protéines. En effet, chez l'être humain, les protéines MIFs induisent la mort cellulaire en interagissant directement avec l'ADN dans le noyau des cellules [73] ou en interagissant avec le complexe de récepteurs CD74-CD44 dans le milieu extracellulaire des cellules. Cette interaction peut conduire à l'induction de l'autophagie. L'inhibition de la mort cellulaire chez les êtres humains se fait elle exclusivement par une interaction protéine – protéine ayant dans le cytoplasme des cellules ou dans le milieu extracellulaire [41]. Elle peut notamment grâce à l'interaction des protéines MIFs avec la protéine TP53, qui conduit à l'inhibition de l'apoptose par la voie p53 – dépendante. De plus, chez les plantes, la localisation des protéines est souvent liée à leur éventuel impact sur la mort cellulaire [180].

La localisation des protéines en cellule animale comme végétale pourraient donc jouer un rôle déterminant sur les capacités des protéines à moduler la mort cellulaire. Pour aborder ce point, nous avons analysé la localisation subcellulaire grâce à l'expression transitoire des MIF couplées à la GFP en cellules de *N. benthamiana*.

La localisation subcellulaire a été étudiée par microscopie confocale à balayage laser dans les cellules d'épiderme de *N. benthamiana* exprimant les différentes protéines MIFs. Les cellules ont été agro-transfectées avec des constructions qui permettent l'expression des protéines MIFs fusionnées à la GFP. Les cellules d'épiderme possèdent une grande vacuole centrale qui remplit la quasi-totalité de la cellule. Le cytoplasme, avec ses organelles, est plaqué contre le plasmalemme et la paroi de la cellule. L'analyse effectuée ici permet la discrimination entre une localisation dans (i) le cytoplasme, (ii) le noyau, et (iii) les organelles si le marquage apparaît en forme de points.

A l'exception de SIMIF4 et AtMDL3, l'ensemble des protéines MIFs étudiées ont une localisation nucléo-cytoplasmique (voir Figure 24).

Les protéines SlMIF4 et AtMDL3 ont, elles, une localisation peroxysomale. Cette localisation a été confirmée par une étude de co-localisation entre les constructions GFP-MIF et un marqueur du peroxysome fusionnée à la RFP (voir Figure 25).

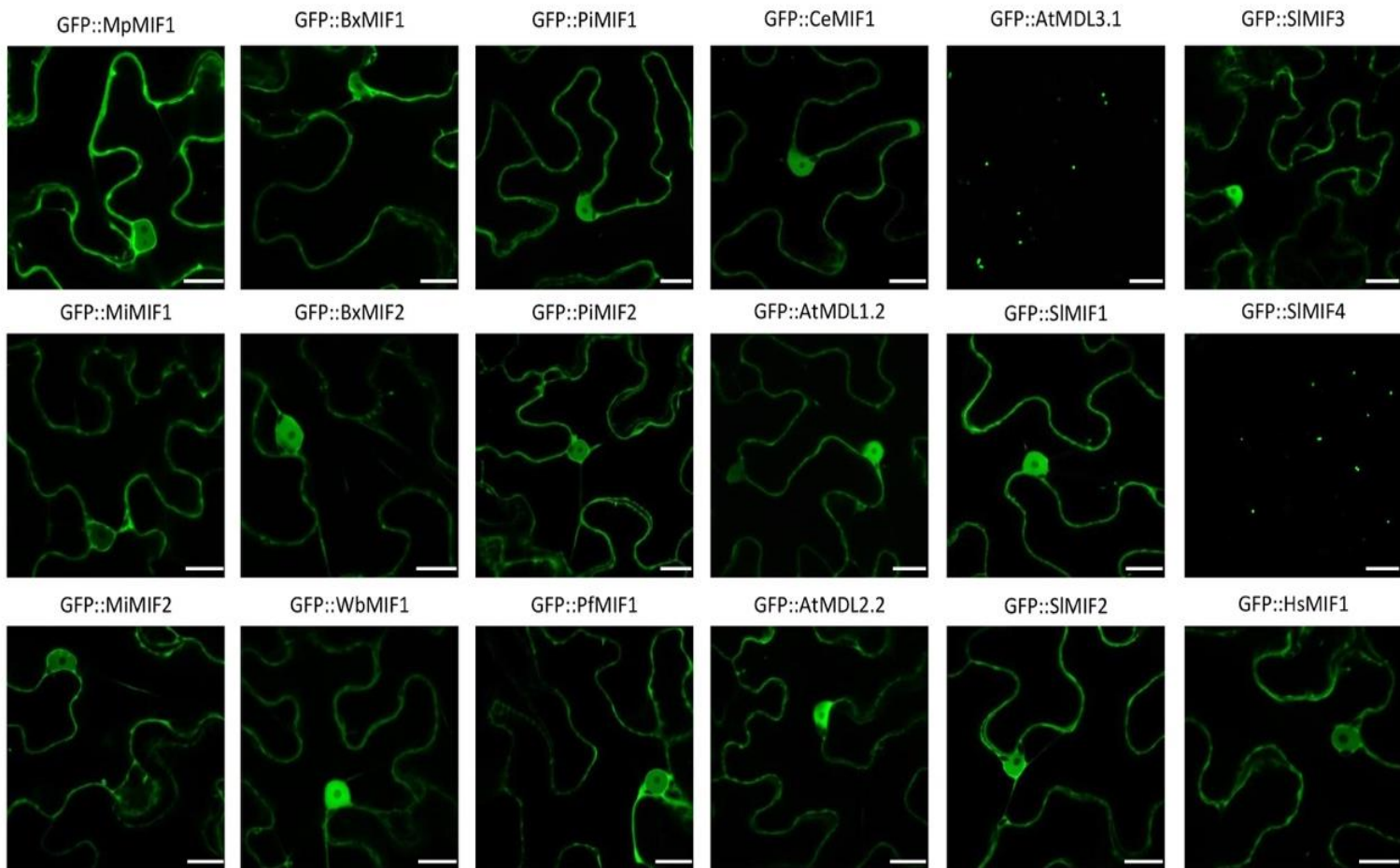


Figure 24 – Localisation subcellulaire des protéines MIFs en cellule végétale

Visualisation au microscope confocal des constructions GFP-MIF en cellules épidermiques de tabac. A l'exception des protéines SlMIF4 et AtMDL3, l'ensemble des protéines MIFs étudiées ont une localisation nucléo-cytoplasmique. La localisation des protéines AtMDL1.1 et 2.1 est identique à celle des variants du splicing alternatif présentés ici. Les barres correspondent à 20 μm .

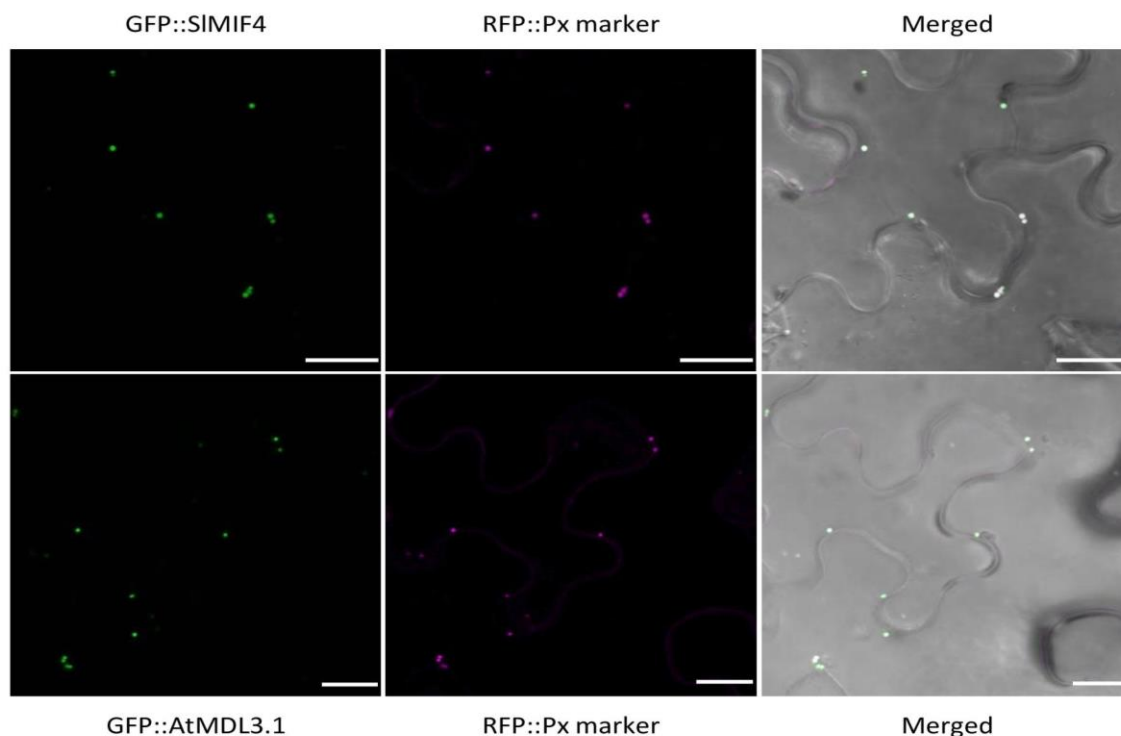


Figure 25 – Co-localisation des protéines MIFs avec un marqueur du peroxysome

Co-localisation des constructions GFP-SIMIF4 ou GFP-AtMDL3 (vert) avec la RFP (rouge) fusionné au « Peroxisome Target Signal 1 » (DRSCKDL) en cellules épidermiques de tabac. La superposition des deux fluorophores (en blanc dans l'image « merged ») indique une localisation peroxysomale des deux MIF. Les observations ont été faites au microscope confocal. Les barres correspondent à 20 μ m.

En conclusion, la localisation subcellulaire des protéines MIFs en cellule végétale n'est pas directement liée à leur activité d'inhibition de la mort cellulaire.

Le fait que la majorité des protéines MIFs ait la même localisation ne veut pas dire que celle-ci n'ait pas de rôle dans la mort cellulaire chez la plante. En effet, les protéines MIFs peuvent se lier à de nombreuses protéines pouvant avoir des rôles modulateurs de la mort cellulaire. Ainsi, les protéines MIFs sont capables d'agir en tant qu'inducteur de l'apoptose en se fixant à la protéine BCL2-L11 (aussi appelée BIM) ou en tant qu'inhibiteur de l'apoptose soit en se fixant aux complexes enzymatiques BAX/BID soit en agissant sur la voie des p53 [41]. Le fait que les protéines AtMDL3 et SIMIF4 soient localisées aux peroxysomes en cellule végétale, contrairement aux autres protéines, tout en ayant le même impact sur la mort cellulaire, suggère que les MIFs agissent (i) sur différents interlocuteurs dans différents compartiments de la même cascade de signalisation, ou (ii) sur deux cascades de signalisation séparées mais aboutissant à la même finalité. Une analyse détaillée des caractéristiques (motifs, structure, paramètres physico-chimiques) conservées chez l'ensemble des protéines MIFs est donc nécessaire pour mieux comprendre les mécanismes à l'origine de la capacité des MIFs à inhiber la mort cellulaire.

V.2.5- Recherche de relation structure-fonction

Les caractéristiques des séquences protéiques, en particulier la présence de motifs particuliers, la structure 3D ou les modifications post-traductionnelles, peuvent déterminer de nombreuses activités directes (e.g. activités enzymatiques) ou indirectes des MIFs. Par exemple, la pleiotropie de la protéine MIF humaine résulte en partie de sa capacité d'interaction avec de nombreuses protéines [41, 47, 72]. Ainsi, la capacité des MIFs à se fixer ou non à certaines protéines pourraient également expliquer les différences d'inhibition de la mort cellulaire. Afin d'explorer cette possibilité, une analyse *in silico* des caractéristiques des protéines MIFs a été réalisée.

V.2.5.1- Recherche de motifs

Afin de déterminer si un motif serait caractéristique des MIFs ayant la capacité d'inhiber la mort cellulaire en cellule végétale, les séquences protéiques des 20 MIFs utilisés ont été alignées et comparées. Tous les acides aminés différents entre les séquences consensus des protéines MIFs capables ou non d'inhiber la mort cellulaire ont été examinés.

Cette analyse ne révèle aucun motif ou acide aminé particulier qui serait :

- (i) spécifique des protéines ayant la capacité d'inhiber ou non la mort cellulaire chez *N. benthamiana* (Figure 26).
- (ii) associé à une fonction précise chez l'homme et à l'absence d'effet d'inhibition de la mort cellulaire.

Résultats et Discussions

	1	10	20	30	40	50	60	70
Fonctions	■	□	XXXXXX	■	■	□	↓↓↓↓↓	■
Fonctions	+++++++vvvv							
HsMIF	MPMFIVNTNVPASV-PDGFLSELTQQLAQATGKPPQYIAHVVPDQLMAFGGS-SEPCALCSLHSGIG- GA							
BxMIF1	MPVLEIRTNLKKEQV-PNGFLLKAAETIAQVTDKPIERVACHVQTDQLMSFGGS-ESPCGSVTIRSIGNVK- DR							
PiMIF1	MPNIQVTSNVSSVGVDKVKVMAISKSLATAMDKSEQVVMVHLNLDTPMLFQAS-DDPCAMIHLRSIGKVD- AQ							
AtMDL1.1	MPTLNLFTNIPVDAVTCSDILKDATKAVAKIIGKPESYVMILLNSGVPIAFAGT-E EPAAYGELISIGGLG- PG							
AtMDL1.2	MPTLNLFTNIPVDAVTCSDILKDATKAVAKIIGKPESYVMILLNSGVPIAFAGT-E EPAAYGELISIGGLG- PG							
AtMDL2.1	MPCLNLSTNVNLDGVDTSILSEASTVAKIIGKPENYVMIVLKGSVPMSFGGT-EDPAAYGELVSIGGLN- AD							
AtMDL2.2	MPCLNLSTNVNLDGVDTSILSEASTVAKIIGKPENYVMIVLKGSVPMSFGGT-EDPAAYGELVSIGGLN- AD							
Consensus	MPXLNLXTNVPXDGVDSXILSEAXTVAKIIGKPEXYVMIHNLXDVPXFGGT-EXPAAYGELXSIGGLG-AX							
BxMIF2	MPILKIATDLQSSQI-PSDFLAKASSLVATLTGKPESYVQVVLNSNNAISFGGK-VSSSAFVELGSIGGFG- GK							
MiMIF2	MPVLQIFTNISTESV-PPKFAKIASANIATLTGKPESYVMVLVNAGNVGCFGGS-TDPFIYAELQSIGGFT- DP							
CeMIF1	MPVFSINVNVKVP AEKQNEILKELSTVLGKLLNKPEQYMCIHFHEDQGILYAGT-TEPAGFAVLKSIGGVGSAK							
MiMIF1	MPILQVFTNIASKSI-PSEF-SKAAKIIANLTGKKEASVMVLVNAGNVGCFGGS-TDPFIYAELQSIGGFT- DP							
PiMIF	MPCEVITNVNLPDDNVQSTLSQIENAI SDVMGKPLGYIMSNYDYQKNLRFGG-NEAYCFVRITSIGGIN- RS							
WbMIF1	MPYFTIDTNKPQDSI-SSAFLKAPNVVPKALGKPESYVSIHVNGGQPMVFGGS-EDPCPVCVLKSIGCVG- PK							
PiMIF2	MPFVHTSNVPKTDVDPKALRALSALGKPEAYVMVQLDLDMPMIFQAS-DAPCAFIHRSIGRIG- PD							
SIMIF1	MPCLNISTNVNLEGVDTSVSEATSTVAKLIGKPEAYVMIVLKGSVPMAFGGT-EQPAAYGELVSIGGLN- AD							
SIMIF2	MPTLNLFTNLPVDAVIASDILKDATKAVAKIIGKPESYVMILLNNGVPIAFAGT-EAPAAYGELISIGGLG- PS							
SIMIF3	MPCLNISTNVNLEGVDTSVSEATSTVAKLIGKPEAYVMIVLKGSVPMAFGGT-EQPAAYGELVSIGGLN- AD							
SIMIF4	MPCFNLSTNVNLDGVDTSDFSEATKAVSSIIGKPENFVMVVLKGSVDISFGGN-KEPAFAEIVSMGGIN- SD							
AtMDL3	MPCLYITTNVNFQVNTDPFYSEVTKAVASIVGRPNQLVMVVLKGSVEIVFGGN-KEAAAAYEIVSMGGIT- KQ							
MpMIF1	MPHFRLTNISKSKV-TPELLKKLSAAVAKTLGKPESYVVVTIVPDQLMTWGGD-DKPCGTATLMSIGCLG- VE							
MpMIF3	MPTLSITTNLPKYRI-PSTFLADVSTLVSQLQTPEQYIAVRVKAGQQMFWNN-ESPCALGNITGTGNFG- IE							
MpMIF4	MSILRLDTNVSHSDI-DDEFLVQSTEALAKTLKKPKSDILVFVNGNQPIIAGSAEPAVIVSLLSVDGIN- EI							
Consensus	MPCLNI+TNVNLGVDTSSEFLSEA+KAVAKLLGKPESYVMVVLNNGG+P+AFGGSAEPAAYAEIVSISIGGIGS+D							

	80	90	100	110	120
Fonctions	TTTTTTTT	■	■		
HsMIF	QNRYSKLLCGLLAERLRISPDRVYINYYDMNAANVGWNNSTFA				
BxMIF1	LDKL-AKELTDLIVKELVIPSDFYILFQELGVDDVAHKGTLVRELKQNPQ				
PiMIF1	HNSTTASMLTETVSQELNIPKDRIFMNIIDVLRSNWAKGGVLIPKQ				
AtMDL1.1	VNGKLSETISEILQIKLSIDSSRFYIKFYDSRPFFGYNGSTF				
AtMDL1.2	VNGKLSETISEILQIKLSIDSSRFYIKFYDSPVSPLLLIQSFA				
AtMDL2.1	VNKKLSAAVSAILETKLSVPKSRFFLKFYDTKGSFFGWNGATL				
AtMDL2.2	VNKKLSAAVSAILETKLSVPKSRFFLKFYDTKDHQSQEYAQCLHALHQHY				
Consensus	VNKKLS++LSEIL++KLSIP+SRFYIKFYD+++S++GWN+TFAEL+Q++PQ				
BxMIF2	TNEI-AEALTGLISSNLKIPKDRFYIKFTEIEASELSHNGSTF				
MiMIF2	--NKVTGEMTKLFTEHFGVPGSRVYMKLTGPDANKFAHNGKTFA				
CeMIF1	QNNAISAVVFPIIEKHLGIPGNRLYIEFVNLGAADIAYNQTFFA				
MiMIF1	--NKVTGEMTKLFTEHFGVPGSRVYMKLTGPDANQIACDGKLG				
PiMIF	NNSALADQITKLLVSNLVKSRRIYVEFRDCAQNFAFGSLFG				
WbMIF1	VNNSHAELYKLLADELKIPKNRCYIESVDIEASSMAFNGSTFG				
PiMIF2	LNPKTAASLTATAAKALKLPDRVFLNLDLDATNWGMAGNIG				
SIMIF1	VNKKLSAAIADILETKLSIPKSRFFLKFYDTKGSFFGWNGSTF				
SIMIF2	VNGKLSSTIAEILQTKLSIDSDRFYIKFYDSRPFFGFNGSTF				
SIMIF3	VNKKLSAAIADILETKLSIPKSRFFLKFYDTKAHQSQEYAQCLHAIHQY				
SIMIF4	VKKLIATLGGICQNRFSIPRTRFFLKVYD---TTMATKFSKL				
AtMDL3	VKRELIAVGSILHTHFSIHPTRFIFKVFIDINSLPLPSKL				
MpMIF1	QNKKHAAVLYPLLKELGIPDDRLYITFSDQSSNVGYTGTTFTILG				
MpMIF3	ENKHYASIIYDFVEKQLGIPQDRFYLSFVEQKPSNVGVKGTTLQEIIQ				
MpMIF4	DNKLHSAALFSLITKYLKINENRITVTFSPIESHAMGRNGKTLIQ				
Consensus	VNKKLAAALTKLLEKHL+IPKDRFY+KFYDI+ASN+FNSTFG+I+QY				

Figure 26 – Alignement des séquences protéiques MIFs

Alignement des séquences protéiques MIFs étudiées pour leur rôle dans la modulation de la mort cellulaire. Les protéines MIF incapables ou capables d'inhiber la mort cellulaire en cellule végétale sont notées respectivement en noir ou en rouge. Les symboles représentent les différents sites de la protéine HsMIF importants pour certaines fonctions ou activités. Ces sites permettent à la protéine d'avoir une activité tautomérase (symbole ■), une activité oxidoréductase (motif CXXC, symbole ∨), une activité nucléase (motif PE/DxxxxE, symbole ×) mais aussi de se fixer aux récepteurs CXCR4 (symbole ⌘), aux récepteurs CXCR2 (motif pseudo ELR, symbole □). Ils permettent les interactions entre la protéine MIF et le récepteur CXCR2 (pattern N like loop, symbole +) et sont une zone de réponse aux dommages à l'ADN (domaine zinc finger CxxCxxHx(n), symbole ↓). Les séquences consensus pour chaque catégorie (MIF capable ou incapable d'inhiber la mort cellulaire) sont notées en dessous des séquences protéiques correspondantes.

V.2.6- Analyse de la structure des MIFs

Puisque ni la localisation subcellulaire ni la présence de motifs particuliers ne semble être lié aux différences d'activité inhibitrice de la mort cellulaire en cellule végétale, nous avons exploré la structure 3D des différentes protéines.

V.2.6.1- Analyse générale de la structure prédite des MIFs

A l'heure actuelle, seule une dizaine de protéines MIFs ont pu être cristallisées pour déterminer leur structure 3D [26, 31, 33, 34, 36]. Basé sur ces structures confirmées, les logiciels Phyre2 et Swiss - Model ont été employés pour prédire la structure tertiaire et quaternaire des 20 protéines MIFs.

Le logiciel Phyre2 estime la qualité de sa prédiction en fonction de la qualité de l'alignement entre la protéine à prédire et son modèle cristallisé. Pour l'ensemble des protéines MIFs, les protéines prédites sont alignées avec leur modèle sur au minimum 90% de leurs séquences. De plus, sur les zones alignées, le logiciel estime sa qualité de prédiction à 100%. Les zones où la prédiction des structures est moins bonne correspondent à des zones où le logiciel ne parvient pas à aligner les protéines entre elles. Ces zones se situent toujours en C-terminal de la protéine.

Le logiciel Swiss – Model n'estime pas la qualité de sa prédiction. Il prédit, cependant, en plus de la structure tertiaire de la protéine, sa structure quaternaire.

La qualité des structures prédites a été vérifiée par une analyse de conformation du squelette polypeptidique des protéines (diagramme de Ramachandran), et par HCA (Hydrophobic Cluster Analysis) qui permet de détecter des conflits de charge dans une structure 3D. Les deux méthodes d'analyse n'ont pas indiqué d'anomalies dans les structures prédites des protéines.

Les structures prédites par les deux logiciels sont très similaires. On retrouve la structure caractéristique de la protéine MIF humaine : 6 feuillets β et 2 hélices α formant le motif $\beta\alpha\beta\beta\alpha\beta$. Le logiciel Swiss – Model prédit que l'ensemble des protéines MIFs étudiées forment des trimères (voir Figure 27).

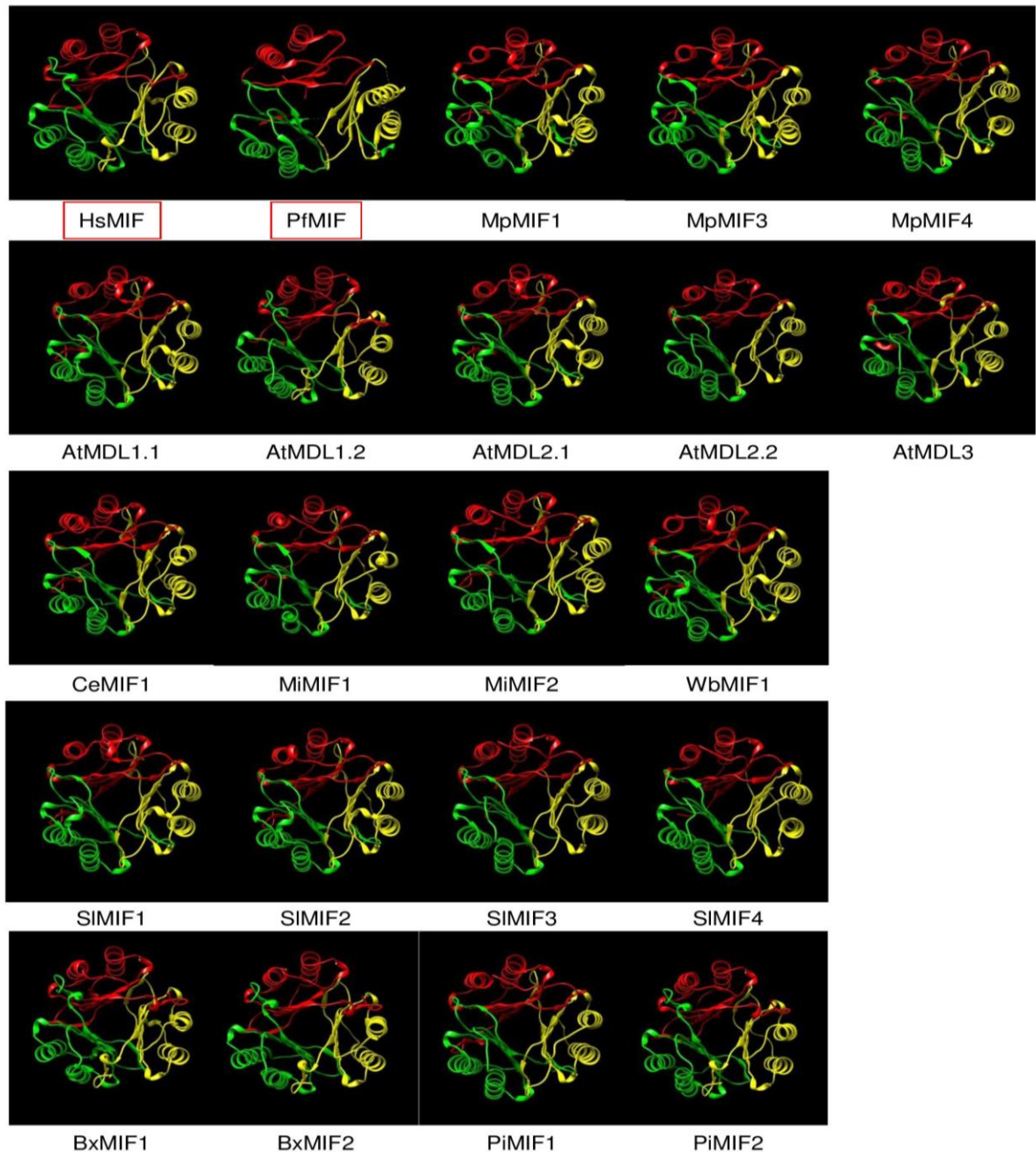


Figure 27 – Structure prédites des protéines MIFs

Les structures des protéines MIFs ont été prédites par le logiciel SWISS-MODEL. Les protéines dont la structure a été déterminée expérimentalement sont encadrées en rouge (*Homo sapiens* et *Plasmodium falciparum*). Toutes les protéines sont prédites comme étant des trimères.

Pour poursuivre cette étude et déterminer des variations fines de structures, les protéines ont toutes été comparées à la structure de la protéine MIF humaine à l'aide du logiciel MatchMaker (Figure 28). La valeur du RMSD (Root-mean-square deviation of atomic positions), valeur qui correspond à la moyenne des distances qu'il y a entre les atomes de la première structure et celui des atomes correspondant sur la seconde structure, a été calculée entre la protéine MIF humaine et l'ensemble des protéines prédites. Plus cette valeur est importante, plus les structures comparées sont différentes.

Les protéines MIFs analysées s'alignent parfaitement sur la structure de la protéine MIF humaine, hormis pour leur régions C-terminales. Cela se visualise notamment par la valeur de RMSD qui est très faible (en moyenne 0.64 Angströms) pour l'ensemble des comparaisons. Les protéines se discriminent par la présence ou non d'une courte hélice alpha ou d'une zone de « coils » de taille plus ou moins longue (voir Figure 28). La majorité des protéines MIFs prédites présente, en C-terminal, une zone de coils comprenant une courte hélice alpha en son centre (voir Figure 28B). Cette structure se retrouve à la fois chez les protéines MIF appartenant à des espèces parasites de vertébrés (*Plasmodium falciparum*) et de plantes (*Myzus persicae*). Seules les protéines MIFs de plantes (*Solanum lycopersicae* et *Arabidopsis thaliana*) présentent des structures protéiques particulières dans la zone C-terminale. Ainsi, nous pouvons constater que chez les deux espèces de plantes étudiées, on retrouve une protéine MIF avec une zone de coils et une courte hélice alpha en C-terminal, mais aussi une protéine ne disposant qu'une zone C-terminale très courte (SIMIF3 et AtMDL2.2) et enfin d'une protéine comportant simplement une zone de coils (SIMIF4 et AtMDL1.2). La protéine AtMDL2.2 a la particularité de ne pas comporter cette zone C-terminale et de se terminer directement après le dernier feuillet bêta. Pour ces protéines, le logiciel Phyre2 estime que la qualité de sa prédiction dans la partie C-terminale est faible (inférieure à 20%). Cette zone ne s'aligne pas avec les modèles de protéines cristallisées décrits à ce jour. Ces analyses ne permettent pas de mettre en évidence de lien entre un mode de vie et la présence ou non d'une structure spécifique en C-terminal. Il ne semble pas non plus y avoir de lien entre une structure particulière et la capacité (ou non) des protéines à inhiber la mort cellulaire. Toutefois, la présence chez les plantes de plusieurs protéines présentant des motifs différents en C-terminal de la protéine soulève la question de l'importance de cette zone dans des fonctions potentielles des protéines MIFs chez les plantes.

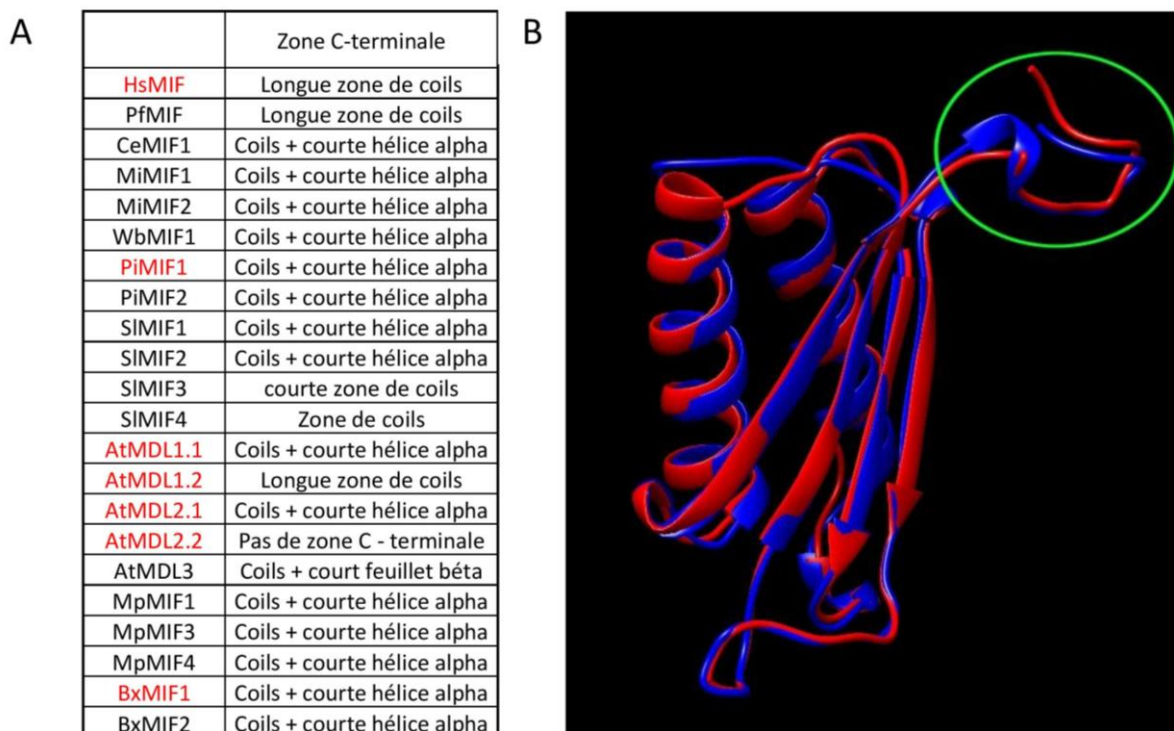


Figure 28 – Analyse de la zone C-terminale des protéines MIFs

En A, Tableau des différentes structures décrites dans la zone C-terminale des protéines MIFs. En B, Alignement de la structure d'un monomère de la protéine MIF humaine (en rouge) et d'un monomère de la protéine SIMIF1 (en bleu). L'encadré en vert représente la zone C-terminale qui a fait l'objet du tableau présenté en A.

En conclusion, bien que les modèles sur lesquels se basent la prédiction de chacun des deux logiciels soit différents, les protéines obtenues par les deux logiciels sont similaires et très semblables entre elles. Il ne semble pas y avoir de lien entre une structure spécifique des protéines et un paramètre particulier (mode de vie, inhibition de la mort cellulaire, ...). La structure ne semble donc pas être un facteur majeur pour déterminer l'origine de l'inhibition de la mort cellulaire par les protéines MIFs. Toutefois, cette analyse a permis de révéler une structure particulière de la zone C-terminale des protéines de plantes. Ce résultat est néanmoins à nuancer. En effet, toutes les structures ont été prédites sur la base d'une dizaine de protéines cristallisées provenant soit de mammifères, soit de parasites d'animaux. Il est donc possible que les prédictions de structures de MIF de plantes ou de parasites de plantes soient moins fiables. Toutefois, cette baisse de la qualité de la prédiction théorique n'est pas constatée par le logiciel Phyre2. Ainsi, les protéines AtMDLs sont toutes considérées comme ayant une qualité de prédiction supérieure à 95%.

Pour rechercher d'éventuelles différences plus fines, une analyse des sites actifs pourrait être réalisée. Toutefois, ce type d'analyse est très fortement soumis à la qualité de la prédiction notamment sur les zones de coils, qui sont les zones les plus variables des protéines. La réalisation d'au moins une structure cristallisée de MIF de plante pourrait permettre de limiter l'impact de la qualité de la prédiction.

V.2.6.2- Analyse de la trimérisation potentielle des MIFs

Une constatation intéressante de l'analyse des structures 3D est que toutes les protéines MIFs étudiées sont prédites pour former des homo-trimères. Chez l'être humain, les protéines MIFs se fixent à leurs récepteurs ou à d'autres protéines uniquement en forme homo-trimériques [32]. Les formes monomériques n'agissent qu'à travers leurs différentes activités enzymatiques (nucléase, tautomérase, et oxydo-réductase) [69–73]. Chez l'homme, la stabilité et le maintien de la forme homo-trimérique se fait exclusivement à travers des liaisons hydrogènes [181, 182]. Afin de déterminer la stabilité des formes homo-trimériques, je me suis intéressée aux acides aminés impliqués dans la formation des trimères. L'objectif a été de comparer le nombre de sites capables de former des liaisons hydrogènes nécessaires à la formation des trimères, et de déterminer si une stabilité prédite du trimère corrèle avec la capacité des protéines à inhiber ou non la mort cellulaire en cellule végétale.

Le logiciel Chimera (voir Matériels et méthodes) a été utilisé pour déterminer le nombre et les positions des liaisons hydrogènes entre les monomères des différentes protéines. Pour les différents homo-trimères, le nombre prédit de liaisons hydrogène varie entre 5 et 13 (voir Figure 29). Chez l'être humain, le nombre de liaisons hydrogènes est de dix. Cette protéine a été décrite comme étant préférentiellement trimérique. Nous pouvons donc estimer qu'un nombre de liaisons hydrogènes équivalent (entre 8 et 12) est probablement le signe d'une protéine MIF stable en forme trimérique. A l'inverse, une protéine MIF ne comportant que peu de liaisons hydrogènes (6 ou moins) aura vraisemblablement tendance à pouvoir passer plus facilement sous forme monomérique. Il s'agit toutefois d'une hypothèse car aucune étude *in vivo* n'a encore eu lieu sur le nombre de liaisons hydrogènes nécessaires au maintien des protéines MIF en forme trimérique. Pour l'ensemble des protéines MIFs comportant entre 8 et 10 liaisons hydrogènes, celles-ci se retrouvent aux mêmes sites que sur la protéine MIF humaine. Chez l'être humain, trois de ces liaisons se trouvent sur la partie C-terminale. Pour l'ensemble des protéines MIFs comprenant moins de 7 sites permettant la mise en place d'une liaison hydrogène, les zones de perte se situent toujours dans la partie C-terminale de la protéine.

Il n'existe pas de corrélation entre le nombre de liaisons hydrogène entre les monomères des différents MIF et leur capacité à inhiber ou non la mort cellulaire (Figure 29A). Par exemple, SIMIF3, AtMDL2.2 et AtMDL3 ont montré des activités diverses en cellule végétale, mais possèdent le moins de liaisons hydrogène entre leur monomères (respectivement 6, 5 et 6 liaisons hydrogènes seulement). L'inhibition de la mort cellulaire en cellule végétale ne semble donc pas être déterminée par la structure tertiaire ou quaternaire de la protéine MIF.

Il est à noter que les protéines MIFs comportant le moins de liaisons hydrogènes (5 ou 6) se retrouvent uniquement chez les plantes. Pour les protéines SIMIF3 et AtMDL2.2, ce faible nombre de liaisons hydrogène peut s'expliquer par la taille de la partie C-terminale de la protéine. Ce nombre de liaisons inférieur au nombre moyen de liaisons indique que les protéines MIFs des plantes changent plus facilement d'une forme trimérique à une forme monomérique. La présence chez *Arabidopsis thaliana*, de deux variants d'épissage AtMDL2.1 et 2.2, dont le nombre de liaisons hydrogènes est opposé (11 et 5) pourrait donc indiquer une nécessité pour la plante de disposer de protéines monomérique ou trimériques. On constate que ce phénomène ne se retrouve pas chez les espèces parasites de plantes pour lesquelles le nombre de liaisons hydrogène varie peu au sein de la même espèce. Par exemple, les protéines MIFs de *Myzus persicae*, comprennent toutes entre 11 et 13 liaisons hydrogènes.

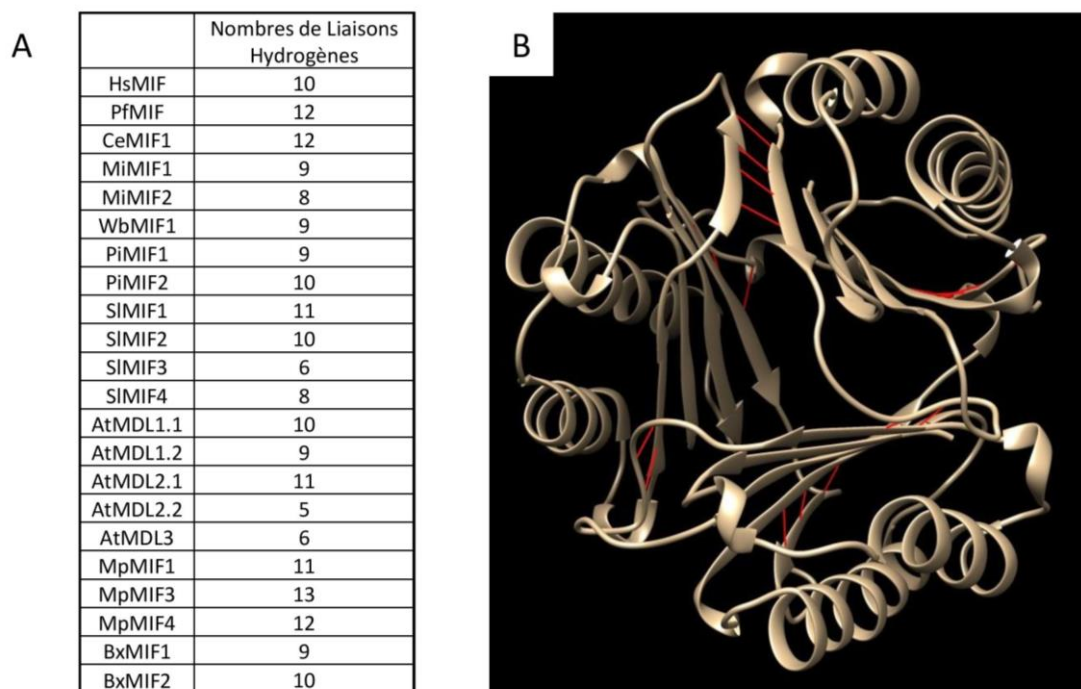


Figure 29 – Liaisons hydrogènes prédites des protéines MIFs

Le tableau (A) indique le nombre prédit de liaisons hydrogène entre les monomères. En B, la figure représente la structure cristallisée de la protéine MIF humaine (en blanc) et les liaisons hydrogènes entre ces différents monomères (en rouge).

En conclusion, les différences de capacité à inhiber la mort cellulaire en cellule végétale ne peuvent donc pas être expliquées par une structure particulière. Toutefois, les structures prédites ne sont basées que sur une dizaine de structure cristallisées de MIF de mammifères ou de parasites de mammifères. La qualité de la prédiction de la structure des protéines MIFs pour les plantes ou leurs parasites est donc relative. Une grande partie des activités de HsMIF étant liée à sa capacité à se fixer à de nombreuses protéines, la moindre modification de structure peut avoir un très fort impact physiologique. Ces interactions pourraient aussi être régulées par la présence de modifications post-traductionnelles. Chez les êtres humains, très peu de choses ont pour l'instant été décrites concernant l'impact des modifications post-traductionnelles sur les fonctions des protéines MIFs. Elles agissent toutefois en amont de voies de signalisation comportant un grand nombre de kinases telles que les MAP kinases. Pour déterminer si les différences de capacité des protéines à inhiber la mort cellulaire peuvent être liées à d'éventuelles modifications post-traductionnelles, j'ai étudié les probabilités de phosphorylation, de glycosylations et d'acétylations des différentes protéines MIFs en utilisant différents logiciels de prédiction.

V.2.7- Analyse des modifications post-traductionnelles

Les protéines MIFs sont connues pour agir directement ou indirectement sur leurs récepteurs ou des protéines à travers une interaction protéine-protéine. Ce type d'interaction nécessite la

présence d'acides aminés précis et peut être modifié par la présence de modifications post-traductionnelles. Pour cette analyse, je me suis donc concentrée sur les trois types de modifications post-traductionnelles les plus courantes : la phosphorylation, la glycosylation et l'acétylation.

V.2.7.1- Analyse de la phosphorylation des MIFs

La phosphorylation est l'un des mécanismes majeurs de régulation des fonctions protéiques. C'est aussi l'une des modifications post-traductionnelles les plus courantes. De nombreuses protéines peuvent ainsi passer d'une forme active à inactive (ou inversement) en fonction de leur degré de phosphorylation. Les protéines MIFs sont connues comme agissant sur de nombreuses kinases ou des phosphatases. Par contre, aucune étude n'a pour l'instant été réalisée sur la possibilité que ces protéines MIFs soient ou non phosphorylées.

Pour analyser les possibles phosphorylations des protéines MIFs, j'ai utilisé deux logiciels complémentaires : le logiciel YinoYang (voir Tableau 4) et le logiciel NetPhos. Le logiciel NetPhos utilise des algorithmes différents et est moins stringent quand à la confirmation qu'un site pourrait être phosphorylé.

Cette analyse suggère que certaines protéines pourraient être phosphorylées et ceux avec des probabilités très fortes pour les deux logiciels utilisés. Ainsi, la position 13 de la protéine BxMIF2 est prédite comme ayant une forte probabilité d'être phosphorylée. L'ensemble des sites confirmés par les deux logiciels se situe en début de protéine (position 11 à 22). Si aucun de ces sites ne correspond à une fonction déjà décrite chez les protéines, ils se situent toutefois sur la zone d'interaction entre les différents monomères des protéines MIFs. Les phosphorylations prédites pourraient donc influencer la formation de dimères ou de trimères chez les protéines MIFs.

Il ne semble toutefois pas y avoir de lien direct entre la présence de phosphorylation au sein des différentes protéines et une inhibition de la mort cellulaire en cellule végétale.

	Logiciel NetPhos				Logiciel YinoYang
	Nombres sites totaux	Nombre sites >0,7	Nombre sites >0,9	Positions des sites avec une probabilité >0,9	Positions (Probabilité >90%)
BxMIF1	7	3	3	49-55-112	-
BxMIF2	15	4	3	13-31-36	13
CeMIF1	5	3	0		-
MiMIF1	6	2	0		-
MiMIF2	10	3	2	31 - 37	-
PfMIF	10	5	3	22-45-58	22
WbMIF1	9	4	1	14	14
PiMIF1	7	2	1	12	-
PiMIF2	6	2	1	38	-
SIMIF1	10	4	0		-
SIMIF2	14	5	2	38-103	-
SIMIF3	10	5	0		-
SIMIF4	13	6	1	90	-
AtMDL1.1	15	3	1	37-38-103	-
AtMDL1.2	16	3	1	37-38-103	-
AtMDL2.1	16	7	1	50	-
AtMDL2.2	16	8	2	50-108	-
AtMDL3	12	2	0		-
MpMIF1	15	5	2	13-37	-
MpMIF3	11	5	1	98	-
MpMIF4	10	5	4	11-13-53-101	11
HsMIF	12	4	2	14-91	14

Tableau 4 – Phosphorylations prédites chez les protéines MIFs

Résultats des analyses de potentielle phosphorylations via le logiciel NetPhos (colonne verte) et YinoYang (colonne bleue). La colonne « nombre de sites totaux » indiquent tous les sites ayant une probabilité >0,5. Les autres colonnes indiquent la probabilité d'existence (>0,7 ou >0,9). La colonne bleue comporte l'ensemble des sites prédits par YinoYang avec une probabilité d'existence supérieure à 0.9.

V.2.7.2- Analyse de la glycosylation des protéines MIFs

a) Les C et N- glycosylations

La glycosylation des protéines est l'une des modifications post traductionnelles les plus courantes. Elle consiste en l'ajout de glucides sur une protéine via une liaison covalente. Cette modification peut débuter lors du passage de la protéine dans le réticulum endoplasmique rugueux ou intervenir lors du passage de celle-ci dans l'appareil de Golgi. Les glycosylations ont un rôle majeur dans la fonction de la protéine. Elles rendent les protéines plus hydrophiles, modifiant leurs activités et la distribution dans la cellule. L'ajout d'une glycosylation sur les protéines MIFs pourraient faciliter leur sécrétion dans le milieu extracellulaire ou perturber les interactions protéines-protéines. Les glycosylations potentielles des MIFs ont donc été

analysées dans le but de déterminer si elles pourraient être liées à l'inhibition de la mort cellulaire observées en cellule végétale.

Les glycosylations de type N et O - glycosylation ont été étudiées en détail. Les glycosylations de type C (ou mannosylation) n'ont pas été étudiées car ce type de modification n'a lieu que sur les tryptophanes. Or, seules cinq des protéines MIFs étudiées comportent des tryptophanes. Ce type de modification n'a, de plus, pour l'instant été découvert que sur très peu de protéines (interleukine 2, ribonucléase, ...).

Chez l'homme, les N-glycosylations s'établissent sur les asparagines présentes dans les motifs Asn-Xaa-Ser/Thr et plus exceptionnellement dans les motifs Asn-Xaa-Cys. Cette modification débute dans le réticulum endoplasmique. Les O - glycosylations consistent en l'ajout de glucides sur les chaînes latérales des sérines ou des thréonines d'une protéine. Cette opération a lieu lors du passage de la protéine dans l'appareil de Golgi.

Pour les N – glycosylations, les analyses ont mis en évidence une variation entre 0 et 2 sites prédits (voir Tableau 5) chez l'ensemble des protéines étudiées. Parmi ces sites fortement soutenus, la majorité se situe en début de protéine (position 5 et 9). Ces positions ne correspondent à aucune fonction précédemment décrite chez l'homme.

Chez l'homme, seule la position 73 est prédite comme pouvant être N- glycosylée. Or, l'arginine 73 a été décrite précédemment comme étant nécessaire à la fixation de la protéine MIF sur le récepteur CXCR4. L'analyse de la N-glycosylation n'a pour l'instant jamais été décrite chez l'homme mais le fait que ce site soit prédit comme potentiellement glycosylé pose la question du rôle potentiel de cette modification dans l'interaction de HsMIF avec le récepteur CXCR4.

Sur l'ensemble des protéines étudiées ici, la présence ou non de N – glycosylation ne semble pas liée à la capacité ou non de la protéine à inhiber la mort cellulaire en cellule végétale.

Pour les O – glycosylations, les analyses ont mis en évidence que seules cinq protéines sont prédites comme pouvant être O – glycosylées (voir Tableau 5). De plus, la probabilité de glycosylation de ces protéines reste assez faible (maximum 74% pour SIMIF1). Il est intéressant de voir qu'une seule position a été prédite comme pouvant être O - glycosylée : la position 26. Cette position n'a pas été caractérisée comme ayant une fonction spécifique chez l'homme. Elle se situe toutefois à l'interface entre les différents monomères des protéines MIFs. La présence de cette glycosylation pourrait donc avoir un impact dans la formation des trimères des protéines MIFs. La présence ou non de O – glycosylation ne semble pas liée à la capacité de la protéine à inhiber la mort cellulaire en cellule végétale.

En résumé, les N et les O – glycosylations ne semblent pas liées à la capacité des protéines MIFs à inhiber ou non la mort cellulaire en cellule végétale. Il est toutefois notable que seules les protéines MIFs de plantes (SIMIF1, 3 et AtMDL2.1 et 2.2) ont une probabilité d'avoir une N – glycosylation et une O – glycosylation. Ce type de modification doit fortement perturber l'accessibilité des protéines MIFs à d'éventuelles autres protéines et rendre les protéines plus hydrophiles. La présence de ce type de modification sur deux des quatre protéines MIFs de plantes suggérerait qu'elles n'interagissent vraisemblablement pas avec le même récepteur.

	Inhibition mort cellulaire	N-Glycosylation via le logiciel NetNGlyc				O-Glycosylation via le logiciel NetOGlyc				
		oui/non	Nombres sites prédits	Probabilité (conclusion)		Position	Oui/non	Nombre	Probabilité	Position
BxMIF1	non	non	0				non	0	-	-
BxMIF2	oui	oui	1	faible		109	oui	1	50%	25
CeMIF1	oui	non	0				non	0	-	-
MiMIF1	oui	oui	1	faible		28	non	0	-	-
MiMIF2	oui	oui	1	faible		9	non	0	-	-
PfMIF	oui	oui	1	faible		70	non	0	-	-
WbMIF1	oui	non	0				non	0	-	-
PiMIF1	non	oui	1	forte		9	non	0	-	-
PiMIF2	oui	non	0				non	0	-	-
SIMIF1	oui	oui	2	forte	faible	5 - 111	oui	1	74%	26
SIMIF2	oui	non	0				non	0	-	-
SIMIF3	oui	oui	1	forte		5	oui	1	75%	26
SIMIF4	oui	oui	1	forte		5	non	0	-	-
AtMDL1.1	non	oui	1	faible		111	non	0	-	-
AtMDL1.2	non	non	0				non	0	-	-
AtMDL2.1	non	oui	1	forte		5	oui	1	56%	26
AtMDL2.2	non	oui	1	forte		5	oui	1	56%	26
AtMDL3	oui	non	0				non	0	-	-
MpMIF1	oui	oui	1	forte		9	non	0	-	-
MpMIF3	oui	oui	2	forte		61	non	0	-	-
MpMIF4	oui	oui	1	forte		9	non	0	-	-
HsMIF	non	oui	1	forte		73	non	0	-	-

Tableau 5 – Analyse de la glycosylation des protéines MIFs

Tableau présentant les N-glycosylations (en bleu) et les O-glycosylations (en vert) des protéines MIFs. Le logiciel NetNGlyc utilise neuf prédicteurs différents pour prédire la présence de sites présentant potentiellement des N- glycosylations. Les résultats présentés sont donc la conclusion des neuf prédicteurs. Pour chacune des analyses, le tableau comprend la probabilité d'existence de glycosylations et l'acide aminé sur lequel la modification pourrait avoir lieu. Les protéines capables d'inhiber la mort cellulaire sont rappelées dans la première colonne.

b) Prédiction de O-GlcNAc glycosylation

L'analyse des glycosylations et des phosphorylations des protéines MIFs suggère que certaines protéines pourraient être à la fois glycosylées et phosphorylées. La *O-N-acétylglucosaminylation* consiste en l'addition d'un résidu de N-acétylglucosamine (GlcNAc) sur le groupement hydroxyle d'une sérine ou d'une thréonine d'une protéine. Cette modification par les acétylglucosaminyltransferases est en compétition avec les phosphorylations effectuées par les sérine/thréonine kinases.

La présence de ce type de modification pourrait fortement jouer sur l'accessibilité des protéines MIFs à leurs récepteurs ou à une autre protéine avec laquelle elle seraient en interaction. De plus, ce type de modification est connu pour augmenter de manière significative en cas d'inflammation, de perturbations du métabolisme et lors de la mise en place de différents cancers [183]. Les *O-N-acétylglucosaminylations* sont aussi fortement impliquées dans la mise en place de l'apoptose chez les êtres humains [183]. Le rôle potentiel de ce type de modification sur les protéines MIFs n'a par contre jamais été étudié.

L'état de phosphorylation d'une protéine détermine fréquemment son activité. Une protéine *O-N-acétylglucosaminylée* ne peut pas être phosphorylée et reste inactive dans une cascade de signalisation impliquant des kinases et phosphatases. Elle peut être activée par l'activité hydrolytique de la *O-N-acétylglucosaminidase* qui clive la moitié GlcNAc et rend les serine et thréonine accessibles à la phosphorylation. Pour déterminer si de tels GlcNAc/Phospho « switch » pourraient déterminer l'activité MIF, j'ai analysé les protéines à la fois pour des *O-N-acétylglucosaminylations* et des phosphorylations potentielles. Des prédictions par le logiciel YinOYang ont permis d'étudier les sites susceptibles au GlcNAc/Phospho « switch ».

L'analyse a mis en évidence que les protéines MIFs disposent d'un à cinq sites pouvant être sujets de *O-N-acétylglucosaminylation*. Toutefois, le nombres de sites prédits et les niveaux de confiance de ces prédictions varient d'une protéine à l'autre. Ainsi, la protéine WbMIF1 est prédite comme n'ayant que deux sites d'*O-N-acétylglucosaminylation* potentiels dont la probabilité n'a été soutenue que moyennement. A l'inverse, la protéine PiMIF1 est prédite comme possédant cinq sites dont trois sont très fortement soutenus par le logiciel. Aucun des acides aminés prédits comme pouvant être acétylglucosaminylés ne correspond à des sites connus comme liés à des fonctions spécifiques chez les protéines MIFs humaines.

Le nombre de phosphorylations potentielles au sein des protéines MIFs est faible allant de zéro à un site fortement soutenus par ce logiciel (probabilité supérieure à 80%). Les sites prédits pouvant être à la fois phosphorylé ou *O-N-acétylglucosaminylé* ne sont jamais soutenus fortement. Par exemple, les acides aminés 13 et 36 de la protéine BxMIF2 sont indiqués avec une forte probabilité d'être phosphorylés (98 et 80% respectivement), mais une faible probabilité d'être acétylglucosaminylé (voir Tableau 6). Au total, seules trois protéines MIFs disposent de sites prédits comme pouvant faire l'objet d'un GlcNAc/Phospho « switch ». Pour ces trois protéines, les sites faisant l'objet d'un potentiel conflit se situent au début de la protéine.

	GlcNAc O-glycosylation			Phosphorylation			Possible conflit entre les deux	
	Nombres	Position	Probabilité	Nombres	Position	Probailité	Nombres	Position
BxMIF1	0			0			0	
BxMIF2	3	13 - 55 - 112	++ / ++ / ++	2	13 36	98% 80%	1	13
CeMIF1	1	115	+++	0			0	
MiMIF1	1	8	++	0			0	
MiMIF2	1	111	++	0			0	
PfMIF	0			1	22	98%	0	
WbMIF1	2	14 - 113	++/++	1	14	92%	1	14
PiMIF1	5	7 - 8 - 75 - 76 - 86	++++/+++ / +++/++/++	0			0	
PiMIF2	4	7 - 13 - 25 - 77	++/++/+++ / ++	0			0	
SIMIF1	2	18 - 114	+++ / ++	0			0	
SIMIF2	0			0			0	
SIMIF3	3	18 - 26 - 90	+++ / + / +	0			0	
SIMIF4	2	29 - 110	++ / +++	2	18 29	56% 80%	1	29
AtMDL1.1	1	114	++	1	90	62%	0	
AtMDL1.2	1	114	+++	1	90	62%	0	
AtMDL2.1	3	18 - 26 - 78	+++ / ++ / ++	0			0	
AtMDL2.2	3	18 - 26 - 78	+++ / ++ / ++	0			0	
AtMDL3	0			1	25	67%	0	
MpMIF1	1	104	++	2	41 103	60% 51%	0	
MpMIF3	2	24 - 28	+++ / ++	0			0	
MpMIF4	2	24 - 49	++ / ++	1	11	99%	0	
HsMIF	2	31 - 113	++ / +++	2	14 31	100% 73%	0	

Tableau 6 – Tableau d’analyse des O-N-acétylglucosaminylation

Analyse de la présence potentielle d’addition de GlcNAc (colonne bleu), de phosphorylation (colonne verte) et d’un éventuel passage réversible d’une acétylglucosaminylation à une phosphorylation (colonne orange) réalisée via le logiciel YinOYang version 1.2. Les sites n’ayant qu’une très faible probabilité d’être acétylglucosylés n’ont pas été indiqués ici. Les probabilités d’existence de sites comportant des acétylglucosaminylations sont indiquées par des symboles +. La présence de deux à quatre symboles signifie que la probabilité d’existence du site varie de moyenne à très forte.

Les prédictions d’existence de glycosylations et de phosphorylations ne semblent pas avoir de lien avec l’inhibition de la mort cellulaire en cellule végétale. De la même manière, il ne semble pas y avoir de lien entre les sites soumis à de potentiels acétylglucosaminylations et l’inhibition de la mort cellulaire. Les régulations post – traductionnelles (phosphorylation et glycosylations) sont très impliquées dans la régulation des voies de signalisation conduisant à l’apoptose ou à l’autophagie. Malgré le rôle connu des protéines MIFs en tant qu’activateur de nombreuses kinases (MAP kinases, notamment) [50], celles-ci ne semblent pas pouvoir être régulées par la présence de kinases. Nous pouvons toutefois penser, dans la mesure où certaines des protéines MIFs étudiées semblent pouvoir être glycosylées ou phosphorylées, que les mécanismes

permettant aux protéines MIFs d'inhiber ou non la mort cellulaire en cellule végétale pourraient être différents d'une protéine à l'autre.

V.2.7.3- Analyse de l'acétylation des MIFs

Lors de l'analyse des structures de protéines, nous avons pu voir que la majorité des protéines MIFs étudiées étaient prédites comme étant préférentiellement trimériques et disposant de suffisamment de liaisons hydrogènes pour maintenir cette structure dans le temps. L'acétylation est un mécanisme, principalement décrit chez les histones, qui induit la modification de la structure de la protéine par l'ajout d'un groupement acétyl. Ce type de modification post – traductionnelle pourrait tout – à – fait induire la séparation des protéines MIFs en monomères et ainsi influencer ses fonctions dans la cellule. Il s'agit d'un des mécanismes les plus courants de régulation des protéines. Ce mécanisme est réversible et se fait grâce à une acétyltransférase.

Le logiciel NetAcet a été utilisé pour prédire les acétylations possibles sur les protéines MIFs. L'analyse de l'ensemble des protéines MIFs a mis en évidence le fait que ces protéines ne sont probablement pas acétylées. En effet, seule la protéine MpMIF4 est prédite comme ayant un site pouvant être acétylé et la probabilité d'existence de ce site reste très faible (seulement 50%).

L'acétylation ne semble donc pas être un mécanisme de régulation des protéines MIFs.

	N - Acétylation		
	oui - non	Sites	Probabilité
BxMIF1	-	-	-
BxMIF2	non	-	-
CeMIF1	non	-	-
MiMIF1	non	-	-
MiMIF2	non	-	-
PfMIF	non	-	-
WbMIF1	non	-	-
PiMIF1	non	-	-
PiMIF2	non	-	-
SIMIF1	non	-	-
SIMIF2	peu	-	-
SIMIF3	non	-	-
SIMIF4	non	-	-
AtMDL1.1	peu	-	-
AtMDL1.2	peu	-	-
AtMDL2.1	non	-	-
AtMDL2.2	non	-	-
AtMDL3	non	-	-
MpMIF1	non	-	-
MpMIF3	non	-	-
MpMIF4	oui	2S	50%
HsMIF	non	-	-

Tableau 7 – Prédiction des acétylations des protéines MIFs

Ce tableau indique la présence de sites acétylés prédits (colonne « oui – non »), la position des sites potentiels et l'acide aminé qui peut comporter des acétylations noté via le code à une lettre des acides aminés (colonne « Sites ») et la probabilité d'existence d'acétylation (colonne « Probabilité »). Seule la protéine MpMIF4 est prédite comme ayant une probabilité, égale à 50%, de posséder des acétylations.

V.2.8- Conclusions

En conclusion de ce chapitre, nous avons pu voir que la majorité des protéines MIFs inhibait la mort cellulaire induite par AvrPto. Seules les protéines humaine (HsMIF), BxMIF1 de *Bursaphelenchus xylophilus*, PiMIF1 de *Phytophthora infestans* et AtMDL1, AtMDL2 d'*Arabidopsis thaliana* ne semblent pas capables d'inhiber la mort cellulaire en cellule végétale.

Cette observation étonnante nous a conduit à explorer les mécanismes déterminant cette inhibition de la mort cellulaire en cellule végétale. Nous avons donc étudié la localisation des protéines en cellule végétale, recherché la présence de motifs caractéristiques des protéines inhibitrices de la mort cellulaire, analysé la structure des protéines, et enfin examiné les possibles modifications post-traductionnelles des protéines (voir Figure 31).

Cette étude a mis en évidence la conservation de la capacité des protéines MIFs à inhiber la mort cellulaire en cellule végétale au cours de l'évolution. Elle a, de plus, suggéré l'existence de particularités des protéines de plantes telles que i) une plus forte tendance à rester sous forme monomérique, et ii) la présence de structures particulières en C-terminal de la protéine. Certaines protéines de plantes se caractérisent aussi par la présence potentielle à la fois de N et de O – glycosylation. L'analyse des glycosylations des protéines MIFs a, de plus, permis de montrer la présence potentielle de N – glycosylations à la position 73 de la protéine MIF humaine. Cette modification pourrait jouer un rôle important dans la fixation de la protéine HsMIF à ses récepteurs.

De façon intrigante, aucun des paramètres étudiés dans cette analyse n'est lié (de façon corrélative) à i) la capacité ou non d'inhiber la mort cellulaire ou ii) aux caractéristiques de l'espèce (libre, parasite de plante ou d'animaux).

Partie VI : Conclusion générale et perspectives

La première partie de ce travail constitue la première étude globale de la diversité des MIFs dans les différents règnes d'eucaryotes que sont les végétaux, les champignons, les protistes et les métazoaires. Le criblage de données génomiques et transcriptomiques de plus de 800 espèces a permis de montrer que des protéines MIFs existent dans tous les phylum (hormis, éventuellement les *Rhizaria*), et que le nombre de MIF par espèce varie de 0 à 5.

Alors que la plupart des espèces de champignons (environ 70% des espèces étudiées) ne possèdent pas de MIF, les *Archeoplastida* (correspondant aux organismes photosynthétiques) possèdent trois protéines MIFs. Un résultat significatif de ce travail est que, à l'exception des champignons, les parasites de plantes, phylogénétiquement éloignés comme les straménopiles, les nématodes et les insectes, possèdent également un nombre médian de trois MIFs, alors que les autres catégories (espèces libres hétérotrophes ou parasites d'animaux), ont un nombre de MIF plus faible et/ou plus variable.

Les reconstructions phylogénétiques ont montré que les membres de la famille multigénique MIF ont une histoire évolutive complexe, comportant de nombreuses pertes ou duplications. De façon intéressante, il semble que les pressions de sélection, et donc les contraintes fonctionnelles soient très différentes en fonction des phyla. Par exemple, les nombreuses pertes observées chez les champignons suggèrent une faible importance fonctionnelle des MIFs chez ce règne d'eucaryotes. Ce constat est supporté par l'observation que la phylogénie des MIFs suit la topologie générale des espèces de champignons, indiquant de faibles pressions de sélection sur ces protéines. A l'inverse, l'existence de trois MIFs très conservés chez les plantes (MDL) suggère un rôle primordial de ces trois variants de MDL dans la survie des plantes. Cette hypothèse est également appuyée par nos résultats de sélection de sites puisqu'il apparaît que 94 des 122 résidus de la séquence MDL sont soumis à une forte sélection purifiante. La fonction des MDL de plantes n'a pas encore été étudiée à ce jour. Le seul travail actuellement publié [21] est une étude *in silico*. Elle suggère que les deux gènes *MDL* d'*Arabidopsis thaliana* (*AtMDL1* et 2) sont exprimés constitutivement et sont co-exprimés avec, globalement, le même set de gènes. Le troisième gène (*AtMDL3*) semble surexprimé en situations de stress, et co-exprimé avec un ensemble de gènes impliqués dans l'immunité des plantes [21]. Notre travail sur le rôle des MIFs dans la régulation de la mort cellulaire chez la plante a apporté un élément nouveau sur la fonctionnalité des *AtMDL*, puisque nous avons montré qu'*AtMDL3* a un effet important sur l'inhibition de la mort cellulaire, contrairement à *AtMDL1* et 2.

Ces résultats suggèrent que les trois *AtMDL* pourraient avoir des fonctions différentes, voir complémentaires, et ainsi permettre de réguler finement certains processus comme la réponse immunitaire ou la mort cellulaire.

En ce qui concerne les espèces parasites de plantes, elles semblent posséder un nombre médian de trois MIFs également, issues de duplications plus récentes. L'origine plus récente de ces duplications est déduite de la topologie des reconstructions phylogénétiques, et elle est également appuyée par la position, très proche de certaines copies dans les génomes de différentes espèces de straménopiles, pucerons, etc. Ces duplications se sont donc produites indépendamment dans les génomes d'espèces parasites de plantes.

Ainsi, nous pouvons proposer un modèle évolutif simplifié qui tient compte à la fois des résultats de l'étude sur le nombre de MIF chez les taxons, et des travaux de phylogénie (Figure 30).

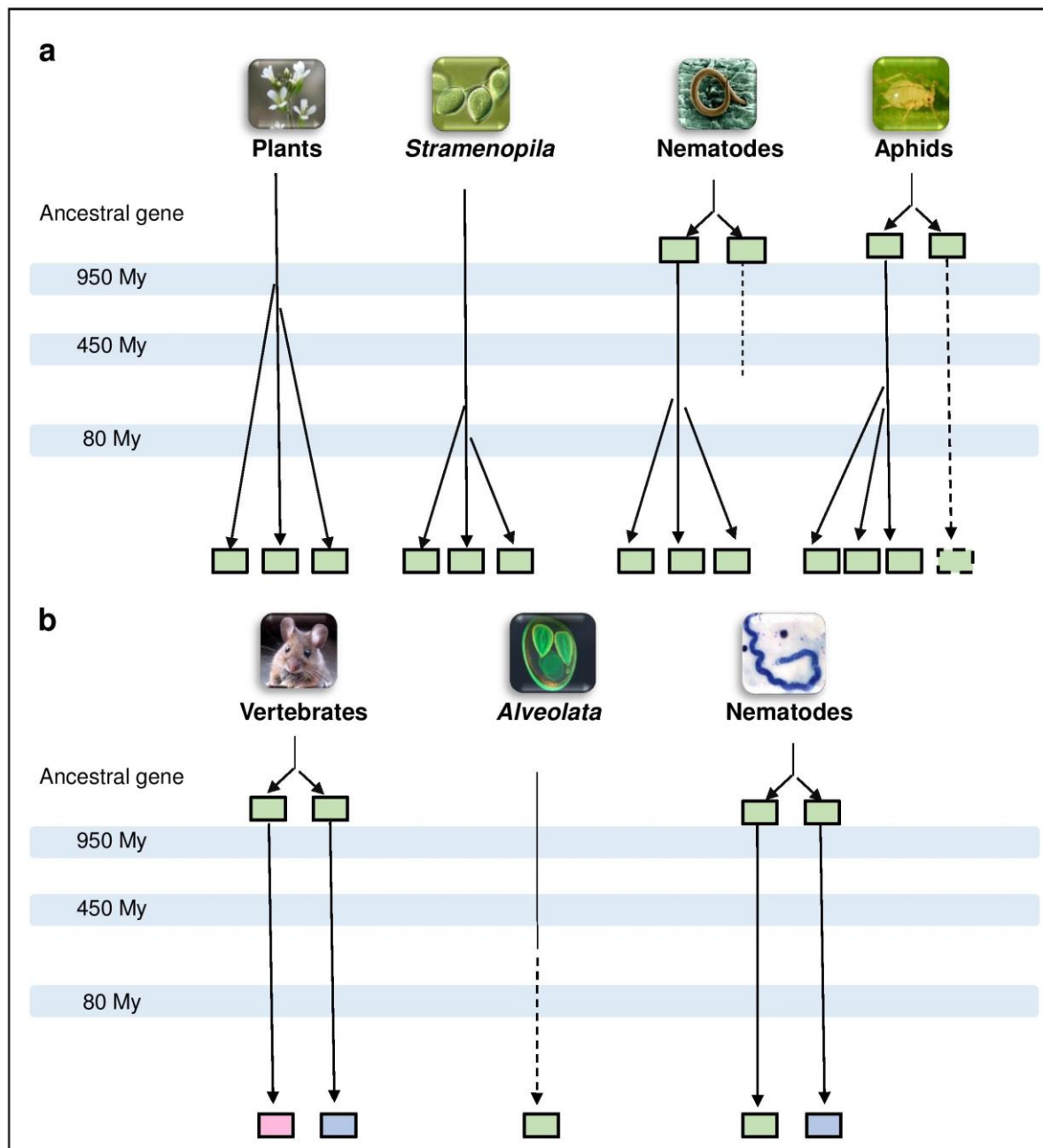


Figure 30 – Modèle évolutif simplifié montrant les évènements de perte et de duplications chez certains phyla d’eucaryotes

Ce schéma représente une version simplifiée des hypothèses d’évolution réconciliant les analyses statistiques sur le nombre de protéines MIF par espèce et les phylogénies à la fois chez les plantes et leurs parasites (A) et chez les parasites de vertébrés et leurs hôtes (B). La couleur vert clair représente les protéines MIFs ne pouvant être directement liées aux protéines MIFs ou D-DT des vertébrés. Les couleurs rose et bleu clair représentent les protéines pouvant être associées aux protéines des vertébrés MIF et D-DT respectivement.

Ces scénarios évolutifs soulèvent la question des contraintes fonctionnelles, et des fonctions précises des MIFs chez les plantes et les parasites de plantes. Tout comme pour les MDL de plantes, nos résultats sur la mort cellulaire suggèrent que toutes les espèces parasites de plante testées ont au moins une protéine MIF capable d’inhiber la mort cellulaire en cellule végétale

et une ou deux autres protéines capables ou non d'inhiber la mort cellulaire. Une première hypothèse serait que chez ces organismes, comme chez les plantes, les différents membres de la famille multigénique MIF/MDL sont nécessaires à la régulation de la réponse de la plante vis-à-vis des agresseurs. Cette hypothèse implique toutefois que tous les membres MIFs des parasites de plante sont en interaction avec la plante, ce qui n'est probablement pas le cas. Ainsi, notre équipe a montré que chez les pucerons, seul MpMIF1 (chez *M. persicae*) et ApMIF1 (chez *A. pisum*), sont exprimés dans les glandes salivaires [20]. Une autre hypothèse serait que chez les organismes parasites de plantes, certains MIFs soient effectivement impliqués dans le dialogue moléculaire avec la plante hôte (sécrétés ou injectés dans la plante et actifs sur certaines voies de l'interaction), alors que d'autres sont uniquement impliqués dans la régulation de l'immunité du parasite. Seules les futures études fonctionnelles sur les MIFs de ces différents modèles pourront éclaircir cette question.

Afin d'avancer dans la compréhension des caractéristiques fonctionnelles des MIFs, je me suis ensuite focalisée sur la modulation de la mort cellulaire.

L'observation marquante de cette analyse a été de constater que de nombreuses protéines MIFs provenant d'espèces diverses comme des protistes, des insectes ou des nématodes, étaient capables d'inhiber la mort cellulaire en cellule végétale, y compris les MIFs issus d'organismes qui ne sont pas en interaction avec les plantes, comme *P. falciparum*. Ce résultat suggère que cette activité d'inhibition de mort cellulaire n'est pas une adaptation, résultant d'une néofonctionnalisation des MIFs, des parasites de plantes, mais est, d'une certaine façon, « inhérente » à la structure même de ces protéines.

Etant donné l'importance du contrôle de la mort cellulaire chez les organismes, qu'ils soient animaux ou végétaux [85, 184], nous nous sommes focalisés sur cette activité et j'ai tenté d'identifier des caractéristiques de ces molécules, qui soient liées à cette activité.

La Figure 31 résume l'ensemble des résultats obtenus avec les différentes analyses et met en évidence l'absence de lien entre inhibition de mort cellulaire et propriétés ou caractéristiques particulières des protéines MIFs. Ainsi, en dépit de nos efforts pour explorer en détail, *in silico*, une possible relation structure-fonction, aucun élément explicatif n'émerge de ces résultats.

Il faut noter que ce travail ne représente qu'une partie de l'étude qui se poursuit actuellement par l'analyse de la localisation subcellulaire et de la capacité à inhiber (ou induire) la mort cellulaire en cellules animales, pour les mêmes protéines MIFs. Ces futurs résultats apporteront des informations complémentaires et permettront peut-être une interprétation plus claire de cette analyse structure-fonction.

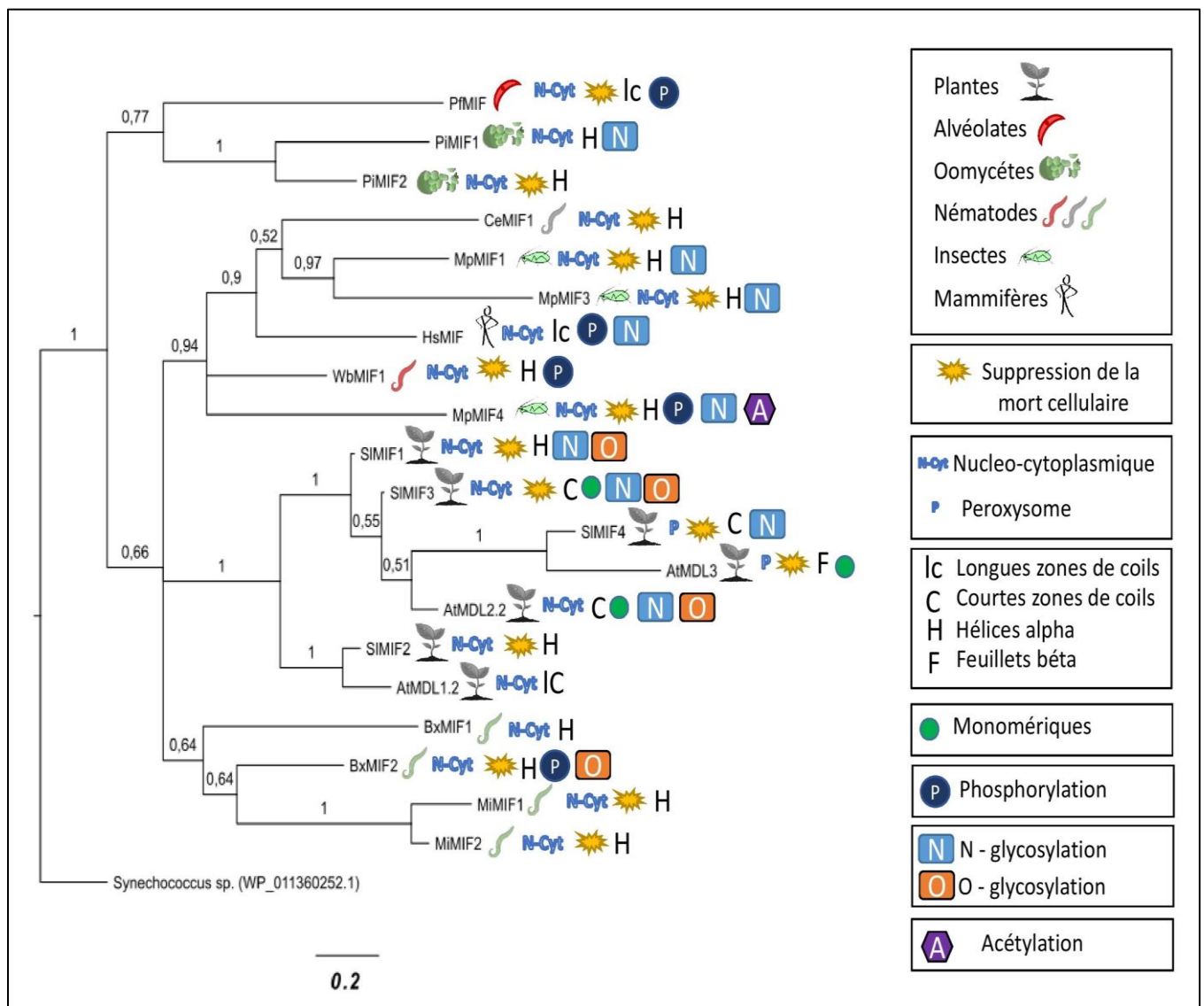


Figure 31 – Résumé des analyses de la mort cellulaire chez les MIFs

Phylogénie Bayésienne basée sur 115 positions, réalisée grâce au logiciel Mrbayes 3.1 et le modèle VT. La séquence de la protéine MIF bactérienne de *Synechococcus sp.* a été utilisée pour enraciner l'arbre. Les probabilités d'existence des branches sont indiqués au dessus de celle – ci. Cette phylogénie restreinte suit la phylogénie générale présent en Figure 14. L'ensemble des caractéristiques prédites des protéines MIFs sont indiquées.

Toutefois, si l'on se focalise sur les travaux en cellules végétales, dans l'état actuel de nos connaissances, nous pouvons proposer deux hypothèses, pour expliquer ces résultats :

- 1) La capacité des MIFs à inhiber la mort cellulaire en cellules végétales serait liée à une caractéristique structurale, non prédictible *in silico*. Ainsi, chez *Plasmodium yoelii*, un Alvéolate, responsable du paludisme chez la souris, de très faibles différences structurales au niveau du site actif de la protéine de *P. yoelii* (PyMIF) comparée à la protéine MIF de souris (*Mus musculus*) MmMIF ont été observées [40]. Ces faibles modifications (seulement deux acides aminés) conduisent les protéines à se fixer

différemment à leurs récepteurs. A de fortes concentrations, compatibles avec les valeurs observées lors de l'infection par ce parasite, les protéines MIFs de l'hôte MmMIF et du parasite PyMIF auront alors des fonctions opposées. Par analogie, on peut penser que de faibles différences d'acides aminés pourraient expliquer les différences de fonctions observées lors de la modulation de la mort cellulaire. Toutefois, ces différences ne sont observables qu'à la condition de disposer des structures cristallisées des protéines MIFs. La prédiction des structures est, en effet, biaisée par le faible nombre de modèles structuraux de protéines MIFs cristallisées dont aucune n'appartiennent à des protéines MIFs de plantes ou de parasites de plantes. Une analyse des sites actifs des protéines MIFs, ainsi qu'une analyse plus détaillée des charges pourraient nous permettre d'analyser plus en détail cette hypothèse.

- 2) Les différentes protéines MIFs testées pourraient agir, dans les cellules végétales de tabac, sur des voies de signalisation différentes, donc par le biais de récepteurs ou molécules partenaires différentes. Si cela est le cas, aucune relation structure-fonction ne pouvait être mise en évidence par notre approche. Cette hypothèse semble possible, si l'on considère la complexité fonctionnelle impressionnante de la protéine MIF la plus étudiée à ce jour, HsMIF, et le fait que chez les plantes, comme chez les mammifères, différentes voies de signalisation peuvent être impliquées dans le déclenchement de la mort cellulaire. Comme déjà mentionné, la mort cellulaire en plante est beaucoup moins caractérisée que celle en cellules animale [84, 185]. Cependant, nos éléments de connaissance actuels permettent, par exemple, de réaliser chez les êtres humains différentes voies impliquant des Mitogen-Activated Protein kinases (MAP kinases) peuvent être impliquées dans l'induction de la mort cellulaire (voir Figure 2) [41, 73]. Ainsi, les protéines HsMIF sont connues comme interagissant indirectement avec des MAP kinases à travers la fixation des MIFs sur les complexes de récepteurs CD74/CD44 ou CD74/ARB1. Ces interactions peuvent permettre l'induction de la mort cellulaire soit par autophagie soit par apoptose. De la même manière, chez les plantes, différentes cascades de signalisation incluent des MAP kinases et sont impliquées dans les défenses de la plante et l'induction de mort cellulaire programmée [186]. Il semble donc possible, que dans le test utilisé (expression en feuilles de tabac), différentes protéines MIFs puissent agir sur différentes MAP kinases induisant des formes de morts cellulaires différentes.

Il est à noter que nous avons analysé la répression de la mort cellulaire suite à une forte surexpression de MIF ectopique en cellule végétale. La forme trimérique de MIF est la forme active. Or, la formation d'oligomères entre protéines n'est jamais statique, mais suit une dynamique d'association et de dissociation. En présence d'un excès de MIF ectopique, il est envisageable qu'un de ces monomères remplace un monomère de MIF végétale dans sa forme trimérique, rendent ce trimère ainsi inactif (ou, au contraire, hyperactif). Si les MIFs trimériques de la plante sont impliqués dans la régulation de la HR, ce remplacement interfère ainsi directement avec la mort cellulaire. La similarité de structure prédite dans la présente étude, supporterait l'hypothèse d'une telle compétition non-sélective.

En conclusion, toutes les questions soulevées par ce travail qui a majoritairement été effectué par des approches *in silico*, doivent maintenant être adressées par des études fonctionnelles détaillées.

Conclusion générale et perspectives

* Comme nous l'avons déjà vu, l'analyse comparée de la structure de ces différents MIFs semble essentielle et des collaborations sont actuellement en cours pour réaliser ce travail. L'identification potentielle de portions de séquences ou de sites/résidus importants peut être suivie par des expériences de mutagenèse dirigée et un test des fonctions ciblées (e.g. mort cellulaire).

* L'une des grandes questions est celle de la fonction des MIFs de plantes. Sont-elles impliquées dans la réponse immunitaire des plantes ou sont-elles impliquées dans la régulation d'autres processus généraux comme la prolifération cellulaire, la mort cellulaire (dépendante ou non de la réponse immunitaire) ? Ces questions sont en cours d'étude par l'un des collaborateurs de l'équipe (Prof. Ralph Panstruga, Aachen), dans le cadre d'un projet ANR-DFG obtenu conjointement pendant le cours de ma thèse. Ces travaux vont notamment utiliser des lignées de mutants AtMDL d'*Arabidopsis thaliana* (simples, doubles ou triples) et explorer les voies de signalisations affectées par chacun des AtMDL.

* Une question majeure est celle du mode d'action des MIFs d'espèces parasites de plantes. L'hypothèse la plus simple est que les MIFs de parasites de plantes utilisent la même machinerie cellulaire et agissent sur les mêmes voies de signalisation que les MDL de plantes. Cela peut se produire soit, en interagissant avec les mêmes interactants que les MDL, soit, en formant des hétérotrimères dont l'activité serait différente de celle des homotrimères. En effet, l'ensemble des protéines MIFs étudiées étant des protéines trimériques et se localisant pour la majorité d'entre elles dans le cytoplasme de la cellule, les protéines MIFs pourraient interagir directement entre elles, et former des hétéro-oligomères. Afin de tester cette hypothèse, j'ai effectué des premières analyses *in silico*. Le nombre de liaisons hydrogène nécessaire à la formation des hétérotrimères a été mesuré dans le cadre d'une interaction AtMDL-MpMIF. Malgré les nombreuses combinaisons possibles entre ces six protéines, cette analyse a montré que la formation des hétérotrimères concernés était théoriquement possible.

La priorité de cette étude semble donc de rechercher les protéines interagissant avec les MIFs de pucerons et ceux d'*A. thaliana*. Ce travail peut débiter par une analyse d'immunoprécipitation suivie par de la spectrométrie de masse pour identifier (et comparer) les potentiels interactants.

Parallèlement, une recherche des voies de signalisation affectées par les MIFs de pucerons devrait être réalisée. Une approche ciblée serait de rechercher une éventuelle action sur les voies MAP kinases. Comme nous l'avons vu, ces voies sont impliquées dans la réponse immunitaire et dans la mort cellulaire, mais aussi soient partagées par les animaux et les plantes [41, 186]. Dans une approche non ciblée, une étude par RNAseq des gènes différentiellement exprimés en présence ou non de MpMIF pourrait apporter des éléments de réponse. L'utilisation du modèle *A. thaliana* et des importantes collections de mutants disponibles permettrait ensuite de valider l'implication de différentes voies de signalisation.

* A plus long terme, ce type d'analyses fonctionnelles devrait être réalisé sur les MIFs d'autres espèces comme celles utilisées dans mon travail, notamment celles des oomycètes et des nématodes. La compréhension du mode d'action précis des MIFs dans ces différents modèles d'interaction plante-parasites, ainsi que dans le cadre de l'immunité des invertébrés, devrait permettre de mieux comprendre l'histoire évolutive de ces protéines et les éventuels événements de néo-fonctionnalisation.

Enfin, si, en 2007, R. Buccala publiait un livre de plus de 300 pages intitulé « MIF : Most interesting Factor », les résultats très complexes obtenus dans ce travail de thèse me pousseraient à proposer « MIF : Most Intriguing Factor ».

Partie V : Annexes

Article issu de mon travail de thèse

Cross-kingdom analysis of diversity, evolutionary history and site selection within the eukaryotic MIF protein superfamily

Michelet C.¹, Danchin E.G.J.¹, Jaouannet M.¹, Bernhagen J.², Panstruga R.³, Kogel K.H.⁴, Keller H.¹ & Coustau C.^{1*}

¹ Institut Sophia Agrobiotech, UCA/INRA/CNRS, 400 Route des Chappes, 06903 Sophia Antipolis, France
Université Côte d'Azur, INRA, CNRS, ISA, 06903 Sophia Antipolis, France France.

² Department of Vascular Biology, Institute for Stroke and Dementia Research (ISD), Klinikum der Universität München (KUM), Ludwig-Maximilians-University (LMU), D-81377 Munich, Germany

³ RWTH Aachen University, Institute for Biology I, Unit of Plant Molecular Cell Biology, Worringerweg 1, D-52056 Aachen, Germany

⁴ Department of Phytopathology, Center of BioSystems, Land Use and Nutrition (iFZ), Justus Liebig University (JLU), D-35392 Giessen, Germany

* christine.coustau@inra.fr

ABSTRACT

Macrophage migration inhibitory factors (MIF) are multifunctional proteins regulating major processes in mammals, including control of the cell cycle and migration, activation of innate immune responses, and prevention of p53-mediated apoptosis. MIF proteins also play a role in innate immunity of invertebrate organisms or serve as virulence factors in parasitic organisms, raising the question of their evolutionary history and of a putative differential evolution of structure/function relationships. We performed a broad survey of MIF presence or absence and evolutionary relationships across 803 species of plants, fungi, protists, and animals, and explored a potential relation with the taxonomic status, the ecology, and the lifestyle of individual species. We show that MIF evolutionary history in eukaryotes is complex, involving ancestral duplications, multiple gene losses and recent clade-specific re-duplications. Of note, plants and plant parasites (other than fungi) harbour a median number of three MIFs, while heterotrophic and animal parasite species harbour a lower and variable MIF number. Intriguingly, MIFs seem to be essential and highly conserved with many sites under purifying selection in some kingdoms (e.g. plants), while in other kingdoms they appear more dispensable (e.g. in fungi) or present in several diverged variants (e.g. insects, nematodes), suggesting potential neofunctionalizations within the protein superfamily.

INTRODUCTION

Macrophage migration inhibitory factors (MIFs) are pivotal pro-inflammatory cytokines/chemokines involved in infectious and immune diseases in humans, such as septic shock, colitis, malaria, rheumatoid arthritis, and atherosclerosis, as well as tumorigenesis[24, 187–190]. They contribute to important physiological functions including control of the cell cycle, sensing of pathogen stimuli, activation of innate defence responses, recruitment and homeostasis of various immune cells, and prevention of p53-mediated apoptosis[54]. Despite a limited molecular mass and a lack of clear modular domains, human MIF contains several sequence motifs including three potential catalytic sites that provide this protein with a surprising functional complexity[24, 73]. Multi-functionality may also be explained by the fact that the protein is remarkably amenable to interactions with a number of target proteins, despite its overall architectural rigidity[191]. For example, human MIF appears to be involved, directly or indirectly, in 49 steps within signalling pathways, thus evidencing the extraordinary functional complexity of the protein[41].

The vertebrate MIF protein family is composed of the actual MIF and D-DT (D-dopachrome tautomerase), which are evolutionarily ancient proteins since *MIF/D-DT*-like genes exist in bacteria[130]. In some vertebrate parasites such as protozoa, nematodes, or ticks, MIF/D-DT proteins are secreted and participate in the modulation of the immune system of the vertebrate host[15, 130]. Intriguingly, free-living invertebrate species also seem to require MIF/D-DT to mount their immune responses, although most of the MIF membrane-expressed partner molecules ('the MIF receptors') identified in mammals are missing. In the gastropod snail, *Biomphalaria glabrata* for example, MIF contributes to the activation of immune responses, promotes immune cell proliferation, and inhibits p53-mediated apoptosis[16]. These functions seem to be general features of MIF in molluscs[80], but also occur in invertebrates such as crustaceans and echinoderms[79]. Other unexpected findings were that plants possess MIF genes (named MDL for MIF/DDT-like)[21], and that aphids secrete a MIF protein upon feeding to manipulate plant immunity[89]. Although the plant immune system shares some similarities with innate immunity in animals, such as the occurrence of pattern-recognition receptors (PRR), nitric oxide signalling, mitogen-activated protein kinase cascades and an oxidative burst, it involves three major plant-specific hormonal pathways, namely the salicylic acid, jasmonic acid and ethylene signalling pathways[192]. In addition, some essential compounds of the animal immune system are missing in plants, such as a circulatory system, specialized immune cells, and prototypical immune cell receptors of the G protein-coupled receptor (GPCR) class. Therefore, the roles that both plant and insect MIFs play during host-parasite interactions and how they may interact with plant signalling pathways remain unclear and deserve dedicated studies[20, 21].

Our current knowledge on non-mammalian MIFs raises questions on the evolutionary history of these proteins, and on a putative differential evolution of structure/function relationships. First, the pattern of diversification, expansion or loss of *MIF* genes across eukaryotic kingdoms is presently unknown. While MIF proteins are

diversified and functionally essential in some species such as aphids, they are missing in other insects such as *Drosophila*, therefore suggesting a complex evolutionary history with differential losses and duplications[128].

Second, because MIF proteins, when present, can be essential immune regulators involved in host survival or parasite virulence, or possibly have yet unknown functions, differential selection pressures may have driven the evolution of MIF sequences according to the ecology or the free-living versus parasitic lifestyles of the respective species. Accordingly, the conservation or diversification of specific amino acid (aa) sites or motifs of MIF proteins across kingdoms might be informative on the earliest MIF function(s) and might give insights into the evolution of MIF functions.

In order to investigate these questions, we performed a broad survey of the presence or absence of MIF and established evolutionary relationships across 803 species of plants, fungi, protists, and animals. We first investigated the variations in MIF number within phyla and explored a potential relation with the taxonomic status, the ecology, and the lifestyle of individual species. We then conducted a general phylogenetic analysis of MIF proteins across kingdoms, and finally examined the structure/function relationship by analysing site selection on sequences from selected phyla including free-living species and parasites of plants and animals.

RESULTS

MIF presence and number across eukaryotic kingdoms

To link the abundance of MIF sequences with the taxonomy and ecology of species across eukaryotic kingdoms, we first investigated the presence of MIFs in species from protists, fungi, animals and plants. In case of the former three taxonomic groups, we focused on phyla comprising both free-living species and parasites of plants and animals. In total, we analysed 803 species for the presence and number of MIF sequences in relation to information on their taxonomy, their environment (fresh water, marine or terrestrial) and their lifestyle (free-living heterotrophic or autotrophic species, plant parasite or animal parasite; see Methods). In addition, we considered the presence or absence of a sequenced genome and the number of available expressed sequence tags (ESTs) when no sequenced genome was available, thus gathering together data that provided a solid framework for our study (Supplementary Table S1). Depending on the availability of a sequenced genome and on the quality of assembly and annotation, comprehensive identification of gene copy number was not always possible. We thus do not report here the number of *MIF* gene copies, but the number of distinct MIF protein sequences retrieved after BLAST searches and we therefore use the term “MIF number” throughout the manuscript.

Species possessing MIF protein sequences were detected in all main eukaryotic phyla with the exception of *Rhizaria* (Figure 1a). However, the absence of MIF from the entire taxon is uncertain as only five species have a sequenced genome (Supplementary Table S1). To investigate a potential relation between the number of MIF sequences and taxonomic or ecological parameters, we removed from the analysis several basal eukaryotic taxonomic groups that were represented by only few species with available sequence data (i.e. *Haptophyceae*, *Apusozoa*, or *Parabasalida*), thus leaving a remaining total of 677 species out of the initial 803 species. The results from a Multiple Correspondence Analysis (MCA)[193] performed on all parameters indicate that the species environment is poorly related to MIF number (Figure 1b). Although the median number of MIF per

species is 1 for the marine and terrestrial environment, and 0 for the fresh water environment (Supplementary Figure S1a), a stepwise regression analysis confirms that “environment” is not a significant parameter for the prediction of MIF number ($P>0.05$, generalized linear regression model). By contrast, MCA analysis shows that the distribution variation in MIF number is partly explained by both the parameters “taxa” and “lifestyle”, that are strongly associated in fungi and *Archaeplastida*. Indeed, fungi comprise many plant-parasitic species and possess no, or one MIF at most, while *Archaeplastida*, exclusively represented by autotrophic species, mostly possess three MIFs, which significantly differs from the other groups ($P<0.01$, generalized linear regression model) (Supplementary Figure S1b).

In order to explore further the potential importance of “lifestyle” as a parameter, while simultaneously reducing the confounding effect due to the phyla fungi and *Archaeplastida*, we focussed on four other major phyla that show variation in MIF number and cover a substantial diversity of both free-living and parasitic species. These are the *Stramenopila*, *Alveolata*, and, among metazoans, the nematodes and insects (Figure 1c). Interestingly, while the global analysis, including the particular cases of fungi and *Archaeplastida*, suggested a reduced number of MIFs in plant-parasitic species (median value=0, Supplementary Figure S1c), this phyla-specific study shows a higher number of MIF in plant parasites (Figure 1c). In *Stramenopila*, insects, and nematodes, the number of MIFs in organisms having a plant-parasitic lifestyle has a median value of 3 and is significantly higher than in the animal parasite groups ($P<0.01$; Figure 1c). Taken together, these findings suggest i) that free-living autotrophic species have a median number of 3 MIFs, ii) that, with the exception of fungi, plant parasites also have a median number of 3 MIFs, while iii) both heterotrophic and animal-parasitic species have a lower and variable number of MIFs.

MIF phylogenetic reconstruction across kingdoms

MIF phylogenetic reconstructions were performed using the Bayesian Inference (BI) and the Maximum Likelihood (ML) methods on high quality sequences selected from the initial 803 species of plants, fungi, protists and animals (Supplementary Table S1 and Methods). To avoid misinterpretation of poorly supported branches in the phylogenies, we used the Bayesian topology as a reference and reported the approximate likelihood-ratio test (aLRT) support values from the ML analysis together with the posterior probability (PP) values of the Bayesian analysis at corresponding branches.

A first observation is that the global phylogeny of MIF proteins does not correspond to the expected species tree (Figure 2). A second observation is that MIF sequences from *Viridiplantae*, *Alveolata*, *Stramenopila* and fungi form highly supported monophyletic groups (Figure 2), whereas the monophyly of other phyla is either not or only poorly supported. Although the monophyly of metazoan MIFs is only moderately supported by the BI method, it is highly supported by the ML analysis (Figure 2), and it is consistent with observations from previous studies on metazoan MIFs[38]. A third observation is that this cross-kingdom analysis does not support the clear segregation of metazoan MIF and D-DT variants (Figure 2) previously reported in studies performed on a subset of taxa[38]. First, chordate MIFs form a group on their own. Second, MIF sequences from most metazoan phyla, such as nematodes, insects, or molluscs, do not form monophyletic groups. Third,

many metazoan phyla possess one (or more) MIF(s) clustering with chordate D-DT variants, as well as one or more MIF(s) that cannot be related to the chordate MIF or D-DT variants (Figure 2). This observation supports the idea that the metazoan MIFs originate from an ancient duplication event that occurred prior to metazoan diversification[194], approximately between 757 and 1.147 million years ago (MYA; according to analysis using TimeTree [172]), but the distinction between metazoan “MIF variant” and “D-DT variant” is not supported by our analysis.

In contrast to the situation in metazoans, the plant MIFs form a highly supported monophyletic group (Figure 2, Supplementary Figure S2). However, the plant MIF tree itself does not match the expected species taxonomy. With the exception of some *Poaceae*, plant MIF phylogeny confirms that sequences cluster into three distinct groups that are related to the *Arabidopsis thaliana* MIFs, AtMDL1, AtMDL2 and AtMDL3[21] (Supplementary Figure S2). This suggests two subsequent duplications in a common ancestor of the species that were considered.

In order to examine better the evolutionary history of *MIF* genes in relation to possible constraints linked to a parasitic lifestyle, we then focussed on phyla containing both free-living species, plant parasites and animal parasites, namely the fungi, *Stramenopila*, *Alveolata*, nematodes and insects. In the following phyla-specific studies, parts of the general tree (Figure 2) are expanded and detailed to show the precise phylogenetic relationships between subsets of the MIF sequences. As mentioned above, most fungal species do not possess *MIF* genes, or only a single copy. The phylogeny of these few MIF sequences is consistent with the expected fungal species tree (Supplementary Figure S3).

MIFs in *Stramenopila* and *Alveolata*

The phylogenetic reconstruction of *Stramenopila* and *Alveolata* species, based on 18S ribosomal RNA (18S rRNA) analysis (Figure 3a) is consistent with the largely accepted species tree of these phyla[167]. Some phyla such as *Ciliophora* and some particular genera and species are missing MIF sequences, despite several genomes publicly available (Figure 3a). This finding suggests that several independent gene loss events occurred in the *Stramenopila* and *Alveolata*. Species lacking MIF show a diversity of lifestyles. Some are free-living species such as *Ectocarpus siliculosus*, or *Paramecium tetraurelia*. Others are obligatory plant parasites, such as *Hyaloperonospora arabidopsidis*, or animal parasites such as *Saprolegnia parasitica*, and *Theileria* sp. Several *Stramenopila* and *Alveolata* species (marked by one star in the Figure 3A) appear to possess either a partial or an atypical MIF sequence with a MIF domain present within a longer unknown protein. Because of the uncertainty of these sequences, they were not included in the phylogenetic reconstruction analysis.

The MIF phylogenetic reconstruction overall matches the species tree, as the sequences from *Alveolata* and *Stramenopila* form two highly supported monophyletic clusters (Figure 3B). However, MIFs from one *Alveolata* species (*Vitrella brassicaformis*) and four *Stramenopila* species (*Aureococcus anophagefferens*, *Aureoumbra lagunensis*, *Phaeodactylum tricornutum* and *Thalassiosira oceanica*) form a separate polyphyletic group. The main clustering of typical *Alveolata* or *Stramenopila* sequences is highly supported by both BI and ML methods (Figure 3b, Supplementary Figure S4). The positioning of MIFs from the five species *V. brassicaformis*, *A.*

anophagefferens, *A. lagunensis*, *P. tricornutum* and *T. oceanica* outside of the main *Stramenopila* and *Alveolata* clusters is also confirmed by both methods (Figure 3b, Supplementary Figure S4).

Within the *Stramenopila*, the oomycete sequences from the genera *Pythium* and *Phytophthora* do not form two separate monophyletic groups. Instead, they are interspersed multiple times. This suggests that duplication events occurred prior to the diversification of *Pythiales* and *Peronosporales*, which is estimated to have occurred approximately 80 MYA, based on TimeTree analysis[195]. Subsequent species-specific duplication events must have occurred more recently as both *Pythium* and *Phytophthora* species show multiple copies clustering together (Figure 3b).

For *Alveolata*, the MIF phylogeny broadly corresponds to the species tree, with the exception of MIFs from *Vitrella brassicaformis*. This dinoflagellate species possesses two MIF proteins and interestingly, one sequence clusters with other *Alveolata* sequences, while the second sequence groups with the “rogue” MIF forming a separate group (Figure 3b, Supplementary Figure S4).

Nematode MIFs

The nematode species tree obtained with the phylogenetic analysis of 18S rRNA genes (Figure 4a) is globally consistent with the currently accepted phylogeny of nematodes[168]. The only exceptions are the *Tylenchina*, which form three distinct groups, one with *Meloidogyne* and *Globodera*, one with *Steinernema* and *Strongyloides* and a last one with *Bursaphelenchus* and *Panagrellus*. All nematode species with publicly available genome data possess *MIF* genes, with the noticeable exception of the cyst nematodes *Globodera pallida* and *Globodera rostochiensis* (Figure 4a). An extensive search of genomes and transcriptomes confirmed the absence of *MIF* sequences from these species and from their close relative in the *Heteroderidae* /*Hoplolaimidae* group. This finding and the presence of *MIF* in the root-knot nematodes (genus *Meloidogyne*) and in *Pratylenchidae*, belonging to the closest outgroups, suggests that *MIF* genes have been lost in a common ancestor of the *Heterodeidae*, *Rotylenchulidae* and *Hoplolaimidae*.

The phylogenetic reconstruction of *MIF* sequences from nematodes (Figure 4b) clearly differs from the 18S rRNA phylogenetic reconstruction (Figure 4a). As expected for many metazoans, most nematodes possess two or more *MIF* copies. One of them clusters with the D-DT variant from chordates. A link between the other nematode *MIF* sequence and the *MIF* variant from chordates is neither supported by BI nor ML analysis (Figure 4b, Supplementary Figure S5a). *MIF*s from nematodes seem to have undergone serial duplications at various times of their evolutionary history. In addition to an ancient duplication event occurring prior to nematode diversification, probably before the metazoan diversification (between 757 and 1147 MYA), some phyla harbour several sets of *MIF* sequences apparently originating from more recent gene duplication events. For example, species of the genera *Meloidogyne* and *Steinernema* have two and three *MIF* variants, respectively, apparently resulting from duplication events that occurred after the separation of these genera from the other nematodes, and before the radiation and expansion of associated species. For *Meloidogyne*, the duplication probably arose before the diversification of the genus, about 78 MYA [171]. Unexpectedly, one *MIF* variant from the *Bursaphelenchus* species clusters with *Meloidogyne* *MIF*s, while a cluster of three sequences from the second variant is distant from other nematode *MIF*s.

Insect MIFs

Insects show a diversity of MIF number, ranging from none to four, with several clades such as *Hymenoptera*, *Diptera* and *Phthiraptera* lacking MIF sequences (Figure 5a), suggesting multiple independent losses of *MIF* genes in the common ancestors of each of these orders. The phylogenetic reconstruction of MIFs from insects clearly differs from the species tree and shows that none of the insect MIF sequences can be related to the MIF or D-DT variant from chordates (Figure 5b, Supplementary Figure S5b). The *Coleoptera*, *Blattodea* and *Orthoptera* genera possess a single MIF sequence, while *Lepidoptera* species possess two variants belonging to two distantly related monophyletic groups (Figure 5b).

As previously shown, aphids possess multiple MIF sequences[128], and the present study reveals that none of the aphid MIFs relates to MIF or D-DT variants from other phyla. Three of the aphid MIFs are closely related among each other, while a fourth one is strikingly different[128], supporting the suggestion that after the first duplication event expected to occur prior to metazoan diversification (between 757 and 1.147 MYA), subsequent serial duplications occurred before aphid diversification estimated around 90 MYA[196].

Conservation of amino acids and motifs within MIF sequences

To explore the possible conservation or diversification of particular aa and sequence motifs within MIF sequences from plant or animal parasites, we analysed MIF sequences from five taxa, comprising *Alveolata* (parasites of animals), nematodes (plant and animal parasites considered separately), oomycetes (plant parasites), and aphids (plant parasites). For comparison, we also analysed site selection in plants. Analyses were performed using five complementary site selection models (see Methods for details).

Combined results of the calculations on the basis of the five different site selection models confirmed the conservation of plant MIFs, as 94 out of the 122 residues are predicted to be under purifying selection by four to five selection models (Figure 6). By contrast, MIF proteins from the other taxonomic groups considered (oomycetes, *Alveolata*, plant-parasitic nematodes, animal-parasitic nematodes and aphids) show a considerably lower number of strongly conserved aa residues (31 to 75 amino acids under purifying selection). A comparatively low number of aa positions (six in total in the different taxonomic groups) are under positive (diversifying) selection. These positions are different in the various taxonomic groups and concern aa that are in part under purifying selection in other taxa. Interestingly, the residue corresponding to the leucine in position 23 in human MIF is predicted to be under positive selection in both plant and aphid sequences (Figure 6).

Remarkably, the MIF tautomerase active sites, encompassing residues Pro-2, Lys-33, Ile-65, Tyr-37, and Tyr-96, (symbol ■ in Figure 6) are globally conserved for most of the species, suggesting potential tautomerase activity of these MIF proteins. The presence of this enzymatic activity in bacterial MIF[129] and its conservation in most eukaryotic species support an ancestral origin of this function. Although a physiologic function of the tautomerase activity has not yet been identified in mammals, it is of note, that the Pro-2 residue has also been suggested to be involved in binding of mammalian MIF to the cognate receptor CD74 and the chemokine receptor CXCR4[46].

Except for the tautomerase active sites, most residues or motifs known to be important for human MIF function are poorly conserved in the other groups. For example, the CXXC motif (Cys-57 and Cys-60) (✓ symbol in Figure 6) is only present in the consensus sequence of animal-parasitic nematodes. This site was previously described to be associated with the regulation of cellular redox homeostasis[72]; it is the active site of the MIF oxidoreductase (TPOR; thiol-protein oxidoreductase) activity. The limited presence of this motif might indicate that the regulation of redox homeostasis is probably not the primary ancestral MIF function.

The pseudo-(E)LR motif (Arg-12, Asp-45), which resembles the ELR motif of ELR+ CXC-type chemokines, was previously described to be important for an engagement of the chemokine receptor CXCR2 by MIF[45]. This motif is absent from the MIFs of the species we tested here, whereas the N-like loop motif, spanning residues 47-56 of human MIF, is partially conserved (e.g. compare LMAFGGSSEP in human MIF with PMSFGGTEE in plant MIF or LMTWGGDDDP in aphid MIF). The N-like loop motif is also involved in human MIF binding to the chemokine receptor CXCR4, but interaction of MIF with this receptor requires additional contributions from the residues RLR in position 87-89[197]. Again, this extended N-like loop is only partially conserved in MIFs from the species studied herein.

A Tyr-100 residue in human MIF was recently identified as a solvent channel gating residue that dynamically interacts with and influences the CD74 activation site of MIF and thus is important for an allosteric mechanism of receptor stimulation[198]. This Tyr-100 residue is present in plant MIFs, but otherwise not conserved in the MIFs of the species tested here.

Finally, a nuclease activity was uncovered for human MIF, which has been suggested to involve the PD/ExxxxE motif (Pro-17 to Glu-22)[73]. This motif is absent from MIFs in other species, just like the zinc finger domain CxxCxxHx(n) (Cys-57 to His-63) commonly found in DNA damage response proteins, which seems also be required for nuclease activity[73].

DISCUSSION

Our large-scale data mining effort involving over 800 species confirmed that MIF proteins are present in all eukaryotic kingdoms and revealed that their number varies from 0 to 5 in the taxa analysed. While most fungi have none or only one MIF, most of the *Archaeplastida*, represented by free-living autotrophic species, possess three MIFs. A major observation from this analysis is that, with the exception of fungi, plant parasites belonging to phylogenetically distant groups such as, *Stramenopila*, insects or nematodes, also have a median number of three MIFs, while both free-living heterotrophic species and animal parasite species have a lower and/or variable number of MIFs. This similarity of MIF number between plants and plant-parasitic species raises the question of particular functional constraints of MIF numbers in plant parasites.

Previous phylogenetic studies on MIF/D-DT (D-DT now also often referred to as MIF-2[38, 199]) in metazoans were either restricted to specific taxonomic groups such as insects[128], arthropods[81], or molluscs[200], or were focused on the relation between sequences from a single species and known sequences from vertebrates[79]. Several of these phylogenies supported a classification of MIF and D-DT variants from metazoa[19, 104]. The present extended cross-kingdom analysis shows that MIF sequences from most

metazoan phyla, such as nematodes, insects, or molluscs, do not form monophyletic groups. While a number of MIF sequences from metazoan such as crustacean, molluscs, echinoderms and nematodes cluster with the chordate D-DT sequences, many other sequences including sequences from metazoans, plants, and protists cannot be related to the chordate MIF or D-DT.

Similar to variations in MIF number, the MIF evolutionary history appears different according to phyla. In fungi, the phylogenetic MIF tree broadly follows the species tree. This observation, together with the numerous independent losses of MIF, provides evidence that MIF is not functionally essential and therefore subject to selection in this taxonomic group. By contrast, as previously reported, both the number and the phylogenetic tree of plant MIFs (MDLs) suggest that they are functionally important and originate from two serial duplication events[21]. This expected functional importance is further supported here by the observation of a strong purifying selection acting on 94 out of 122 residues of MDLs. While ancient duplications resulted in the existence of the three plant MDLs, the presence of three or more MIFs in plant parasites such as oomycetes (protists), nematodes or aphids (insects) seems to result from different evolutionary events involving recent duplications (Figure 7a). By contrast, the phylogeny of MIF in animal parasites such as *Alveolata* (protists) or nematodes does not include recent duplications (Figure 7b).

Altogether, our results support the idea that the occurrence of three MIF copies is functionally important for both plants and plant parasites, regardless of their taxonomic status. To date, studies on MIF functions in plants and plant parasites are extremely scarce. Plant MDLs have not been functionally studied yet. However, an in depth *in silico* analysis suggested that two *Arabidopsis thaliana* MDL genes (*AtMDL1* and 2) are constitutively expressed and share an overlapping set of co-expressed genes[21]. The third gene (*AtMDL3*) exhibits stress-inducible transcript accumulation and is co-expressed with a number of genes implicated in plant immunity[21]. To our knowledge, to date a single functional study on MIF from plant parasites showed that aphid MIF secreted during feeding can interfere with plant immune responses[89]. However, the present comparison of MIF sequences from plants, plant parasites or animal parasites, did not identify particular motifs that are specific for plant parasites since none of the 30 to 60 aa found to be under negative selection in the three groups of plant parasites (oomycetes, nematodes, aphid) appear to be shared exclusively by plant parasites. Also, with the exception of the tautomerase active sites, most sites known to be functionally important in human MIF are not conserved. This situation supports an ancient origin of MIF tautomerase function, but does not allow predicting other potential activities of these MIF proteins. Therefore, future functional studies should help deciphering MIF multifunctional activities in the various kingdom.

In conclusion, this extensive analysis of MIF sequences from plant, fungi, protist and metazoan kingdoms shows that the evolutionary history of eukaryotic MIF sequences is complex, involving ancestral duplications, multiple gene losses and recent clade-specific reduplications. Although the number of MIF variants per species is variable, ranging from none to five in the taxa analysed, factors underlying the variation in MIF number remain unclear for most kingdoms, clades or species. Of note, the single general observation is that both plants (free-living autotrophic species) and plant parasites harbour a median number of three MIFs, while heterotrophic and animal-parasitic species harbour a lower and variable MIF number. Maybe even more surprising is that

MIFs seem to be essential and highly conserved, with many sites under purifying selection in some kingdoms (e.g. plants), while in other kingdoms the protein appears more dispensable (e.g. in fungi) or present in several diverged variants (e.g. in insects and nematodes). This suggests potential neofunctionalizations apart from the ancestral and possibly conserved tautomerase activity. Future comparative functional studies across kingdoms should bring insights into functions and function-associated sites of these evolutionarily and functionally complex proteins and may ultimately help to pinpoint currently unrecognized molecular target sites that could be applicable in MIF-based therapeutic or plant protection strategies.

METHODS

In silico identification of MIF proteins

MIF sequences were searched across eukaryotes, with a special emphasis on phyla including both, free-living species, and parasites of plants and animals. The human MIF family members (MIF and D-DT/MIF-2) [Genbank accession numbers: NP_002406.1 and NP_001077861, respectively] were systematically used as queries in BLAST[201] searches. We also used as queries published MIF sequences of five phylogenetically distant species representing four main clades of the eukaryotic tree of life: The *Alveolata Plasmodium berghei* [Ensembl accession number : CDS52161], the *Excavata Leishmania major* [Ensembl accession numbers: LmjF.33.1740 and LmjF.33.1750], the metazoan mollusc *Biomphalaria glabrata* [Genbank accession numbers : XP_013094629.1 and ACR81565.1], and the metazoan nematode *Ancylostoma caninum*[29] [wormbase parasite accession numbers : ANCCAN_11528 and ANCCAN_25101], as well as the plant *Arabidopsis thaliana* [Genbank accession numbers : NP_200527.1, NP_195785.1 and NP_566955.1].

BLAST searches were performed on general sequence libraries such as Ensembl[135] (<http://www.ensembl.org/index.html>), the National Center for Biotechnology Information (NCBI)[136] (nr database), as well as in various taxon-specific repositories. These comprised dedicated archives for fungal sequences (Joint Genome Institute (JGI[137]) at <http://jgi.doe.gov/> and FungiDB[138] at <http://fungidb.org/fungidb/>), *Diptera* (Flybase[139] at <http://beta.flybase.org/>), *Hymenoptera* (HymenopteraMine[142] at <http://hymenopteragenome.org/hymenopteramine/begin.do>), nematodes (wormbase[143] at <http://www.wormbase.org/> and wormbase parasite[144] at <http://parasite.wormbase.org/>), aphids (Aphidbase[140] at <http://bipaa.genouest.org/is/aphidbase/>) and human parasites (GeneDB[141] at <http://www.genedb.org/Homepage>). MIF protein sequences were also searched in publicly available databases from genome sequencing consortia: *Cyanophora paradoxa* (*Cyanophora* genome project[145] at <http://cyanophora.rutgers.edu/cyanophora/home.php>), *Cryptosporidium hominis* (CryptoDB[146] at <http://cryptodb.org/cryptodb/>), *Globodera pallida* (in the Sanger server at http://www.sanger.ac.uk/cgi-bin/blast/submitblast/g_pallida), *Pristionchus pacificus* (at <http://www.pristionchus.org/blast/>) and *Meloidogyne* species (at <http://meloidogyne.inra.fr>). BLASTP and TBLASTN searches were performed with a cut-off E-value of $1E^{-10}$. For each query sequence, a BLASTP search was performed between the sequence and the NCBI's nr databases. The presence of MIF domains was verified using Interproscan[202]. MIF sequences were also identified using TBLASTN against EST libraries, whole genome shotgun contigs (WGS), or transcriptome shotgun assemblies (TSA) to optimize appropriate diagnostics

on the presence/absence of MIF proteins and/or to improve sequence quality or length. Sequences of 18S rRNA from the corresponding species were retrieved from the SILVA[134] databases (at [https:// www.arb-silva.de/](https://www.arb-silva.de/)) to generate reference species trees with a comparable methodology.

Classification of species ecology and lifestyle

With respect to the ecology of species, we took into account whether the general environment is terrestrial, freshwater or marine. We examined the lifestyle of species regarding the mode of feeding, and classified organisms in free-living autotrophs, heterotrophs, and animal- or plant-parasites. To simplify, we considered as “parasite” all species engaged in a sustained interaction with a host, including, for example, mutualistic symbionts or species classically referred to as “pathogens”. The rationale was that these species share, to some extent, evolutionarily important constraints such as establishing a sustained molecular dialog with the host and evading or manipulating host immune responses, at least at an early stage of infection.

Statistical analyses

Statistical analyses were performed with the R package[203] version 3.1.3. The potential links between the number of MIF sequences present in each species, the taxonomy, the environment and the life style (free-living or parasite) were analysed *via* a generalized linear regression model (GLM[204] function) with a Poisson distribution. The best model based on Akaike information criterion (AIC[205]) was chosen using a stepwise algorithm[206] (“step” function). The Multiple Correspondence Analysis (MCA[193]) was performed using the function “dudi.acm” present in the package “ade4” (Analysis of Ecological Data[207] : Exploratory and Euclidean Methods in Environmental Sciences). The function “scatter” was used to obtain a graphical representation of the multivariate analysis.

Multiple sequence alignments and phylogenetic reconstructions

MIF protein alignments were performed using MUSCLE[148] with standard settings in Jalview version 2.9[208], and examined in Jalview to detect and manually curate incomplete protein sequence predictions, sequencing errors and poorly aligned sequences. The alignments were automatically cleaned using trimAl[209] on gappymode with standard settings. Sequences containing less than 85% of amino acid overlaps with the rest of the alignment were removed. The 18S rRNA sequences were aligned using the specific SINA[151] (SILVA Incremental Aligner) tool optimized for 18S sequences and available in the SILVA[134] databases (<https://www.arb-silva.de/>).

For phylogenetic reconstructions, we used the two complementary methods of maximum likelihood (ML) analysis[210] and Bayesian inference (BI) of phylogeny[211]. For both methods, the best aa substitution models in the MIF phylogeny were estimated from the alignments using ProtTest[212] version 3. The Wag model (without inclusion of a proportion of invariable sites or gamma distribution of the rates of evolution) was systematically estimated to be the fittest for the MIF analyses. The ML analysis was performed using the PhyML[213] method implemented in the SeaView[214] (version 4) software with the following settings: aa frequencies (optimized), proportion of invariable sites and number of substitution rate categories estimated

from the datasets. BIONJ was chosen to create the starting tree and the best method between the NNI (Nearest Neighbor Interchanges) and SPR (Subtree Pruning and Regrafting) tree improvement methods were used to estimate the best topology, and the tree topology optimization option was chosen. Branch support values were calculated using an approximate likelihood-ratio test (aLRT). Bayesian phylogenetic analysis was performed using MrBayes[215] version 3.2.6 and four independent analyses starting from different random trees. The Markov chain Monte Carlo (MCMC) simulations were automatically stopped after average deviation of split frequencies reached a value <0.01 (this was reached after 10,500,000 simulations). For MIF phylogenetic reconstruction, MIF sequences from five different bacteria were used as an outgroup: *Synechococcus sp.*, *Thioalkalivibrio nitratreducens*, *Thioalkalivibrio paradoxus*, *Oscillatoria sp.* and *Microcoleus sp.* [Genbank accession number WP_011360252.1, WP_015260126.1, WP_006746032.1, WP_017717428.1 and WP_015181735.1, respectively].

As a reference taxonomic information, the species phylogenetic trees were performed on the 18S rRNA of species from selected phyla (e.g. *Alveolata* or insects), using the *Arabidopsis thaliana* sequence as an outgroup. For the 18S rRNA, the best nucleotide substitution models were estimated for alignments using JModelTest[216] with default settings and the starting tree. The best method between the NNI (Nearest Neighbor Interchanges) and SPR (Subtree Pruning and Regrafting) tree improvement methods was used to create the starting tree. For illustrative purposes, MIF and species trees were colored in a similar way, based on broad phylogenetic taxa (plants, fungi, *Alveolata*, metazoan, etc.).

Analysis of purifying/diversifying selection

Conservation or diversification of particular aa and sequence motifs within MIF sequences was investigated on sequences from *Alveolata* (parasites of animals), nematodes (plant and animal parasites considered separately), oomycetes (plant parasites), aphids (plant parasites) and plants. *Alveolata* sequences encompassed four *Eimeria* species, ten *Plasmodium* species as well as *Toxoplasma gondii*, *Hammondi hammondi* and *Neospora caninum* (17 sequences). Oomycetes were represented by all MIF sequences from *Phytophthora sojae*, *P. parasitica* and *P. infestans* (nine sequences). For nematode parasites of animals, we selected the MIF sequences from *Ostertagia ostertagi*, *Ascaris suum*, *Onchocerca volvulus*, *Brugia malayi*, *Loa loa*, *Trichinella spiralis*, and *Ancylostoma ceylanicum* (seven sequences). For plant-parasitic nematodes, we used the sequences of all MIF family members from *Bursaphelenchus xylophilus*, *Meloidogyne hapla*, and *Meloidogyne incognita* (six sequences). Aphids were represented by all MIF sequences from *Acyrtosiphon pisum*, *Myzus persicae*, *Rhopalosiphum maidis* and *Rhopalosiphum padi* (13 sequences). Finally, the analysis of plant MIFs was performed on sequences from seven monoliophyte (ferns) species, four chlorophyte (green algae) species, two bryophyte (moss) species, one marchantiophyte (liverwort), one spermatophyte, five monocotyledonous species and five eudicotyledonous species (total of 47 sequences). Multiple sequence alignments of MIF proteins were converted into codon alignments *via* PAL2NAL[217] by providing the corresponding coding sequences (CDS). The existence of purifying (negative) or diversifying (positive) site selection was investigated by a combination of complementary models implemented in the Selecton[153] and

Datamonkey[152] online servers. Selecton uses a BI approach to calculate the ratio between non-synonymous and synonymous substitutions. Analyses with the Selecton tool were performed with the positive selection enabled model (M8, $\beta + w \geq 1$), new tree and the default settings. The null hypothesis was tested using the M8a, $\beta + w = 1$ and the M7, β models.

For the calculations we adopted the following five models in the Datamonkey server: (1) The single likelihood ancestor counting (SLAC), (2) The fixed effect likelihood[154] (FEL), (3) The Random Effect Likelihood (REL), (4) the mixed-effects model of evolution[155] (MEME), which combines the fixed and random effects to identify instances of both episodic diversifying selection and pervasive positive selection at the individual branch site level, and (5) the fast unconstrained Bayesian approximation (FUBAR) using Markov Chain Monte Carlo routine. Analyses on the basis of FEL, REL, SLAC, FUBAR and MEME were conducted with the following settings: global DN/DS estimated with CI (likelihood profile based on 95% confidence interval), the significance level was set at 0.1 and ambiguities were resolved. The substitution model selection was determined by the automatic substitution model selection tool included in the Datamonkey server. The trees required by FEL, REL, MEME, FUBAR, and Selecton softwares were reconstructed using the ML and BI methods. After verifying that the resulting trees were similar, we chose the ML tree for site selection analyses.

Acknowledgements/Funding

This work was supported by the ANR-DFG collaborative project “X-Kingdom-MIF - Cross-kingdom analysis of macrophage migration inhibitory factor (MIF) functions” to C.C. and H.K. (ANR-16-CE92-0014-01), J.B. (DFG BE 1977/10-1), K.H.K. (DFG Ko 1208/26-1) and R.P. (DFG PA 861-15-1), which is co-funded by the French Agence Nationale de la Recherche (ANR) and the German Deutsche Forschungsgemeinschaft (DFG). The work was also supported by the French National Research Agency through the “Investments for the Future” LABEX SIGNALIFE: program reference # ANR-11-LABX-0028-0. J.B. also was partly supported by DFG projects SFB1123/A3 and SyNergy EXC1010. Authors are grateful to Dr. Marc Bailly, PhD, for valuable help with the MCA analysis and Dr. Dzmitry Sinitski for helpful discussions.

Author contributions statement

C.M., E.D., H.K. and C.C. designed and implemented the study. C.M. performed all the database searches, conducted the analyses, and drew the figures. All authors discussed the results. C.M., and C.C. wrote the manuscript with input from all authors. All authors reviewed the manuscript.

Additional information

All data used herein was procured from public databases. The data set supporting the phylogenetic analysis is available in the TreeBase repository (<http://purl.org/phylo/treebase/phyloids/study/TB2:S21992?x-access-code=4a0cbb6434be939b700c0fee1c363e8d&format=html>). Authors declare no competing interests.

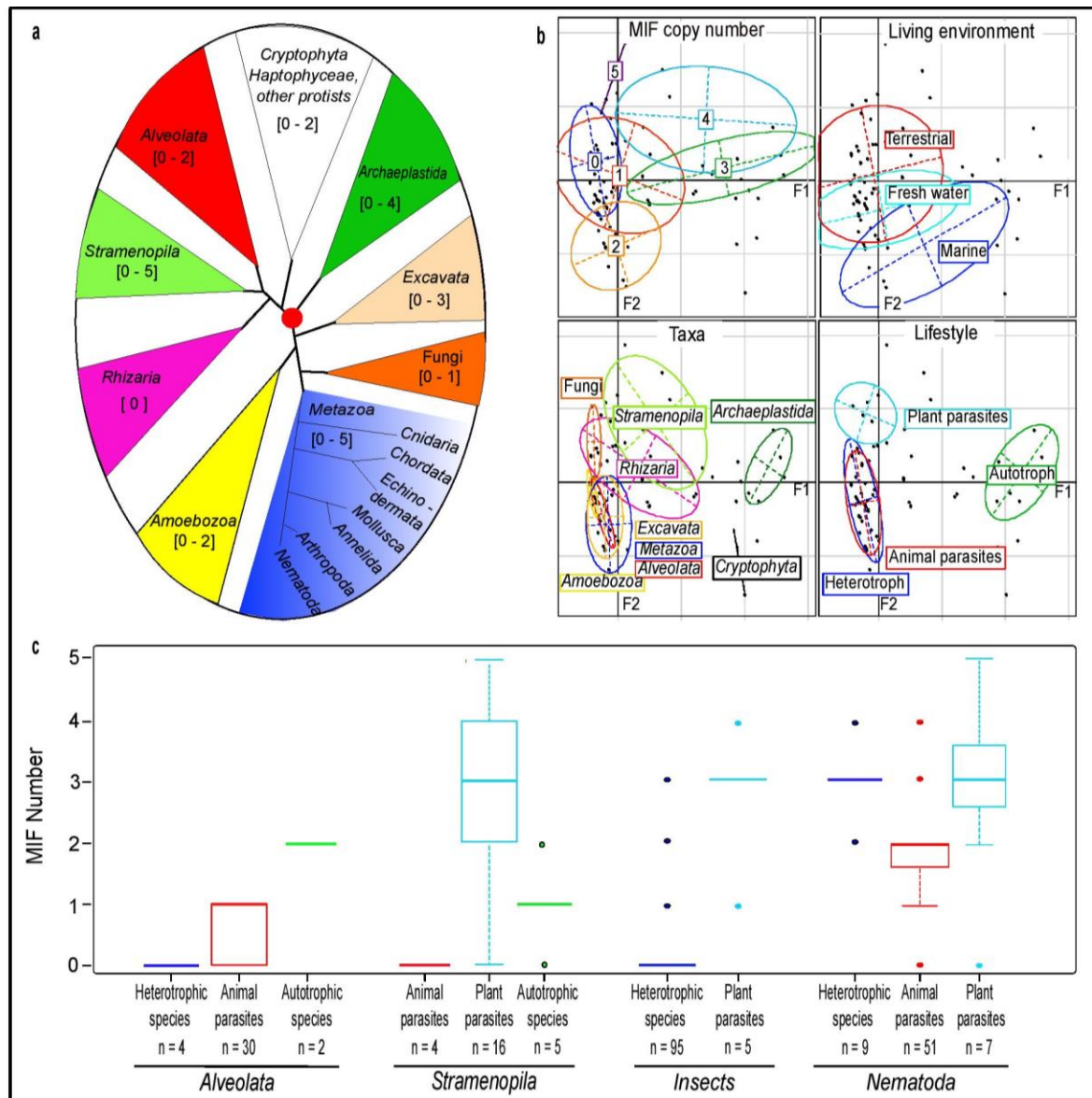


Figure 1: Analysis of MIF distribution across eukaryotic kingdoms.

(a) Rooted phylogenetic tree of eukaryotic kingdoms according to Burki *et al.*, 2014[167]. The size of the triangle areas representing collapsed clades is not proportional to the number of taxa within each clade. The range of MIF sequence number identified in the respective species is indicated within square brackets. **(b)** Graphical representation of Multiple Correspondence Analysis (MCA) data showing, respectively, the distribution of MIF number, the environment, major taxa, and lifestyle. The confidence ellipses delimitate the centroid around each variable. The adjusted inertia of the F1 and F2 axes are 52.54 % and 15.64 % respectively. **(c)** Box plot analysis of MIF number according to the species lifestyle in *Alveolata* and *Stramenopila*, insects and nematodes (n= number of species). The vertical axis corresponds to the number of unique MIF protein sequences per species. Each box is delimited by the quartiles Q1 and Q3, whose variation Q3-Q1 corresponds to 50% of the distribution. The horizontal line in the centre of the box represents the median value. The two ends of the vertical bars correspond to the values of the first and last percentiles (C1 and C99) and delimitate 98 % of the sample distribution. Extreme values (outliers) are indicated by a dot (below 2% of the sample distribution).

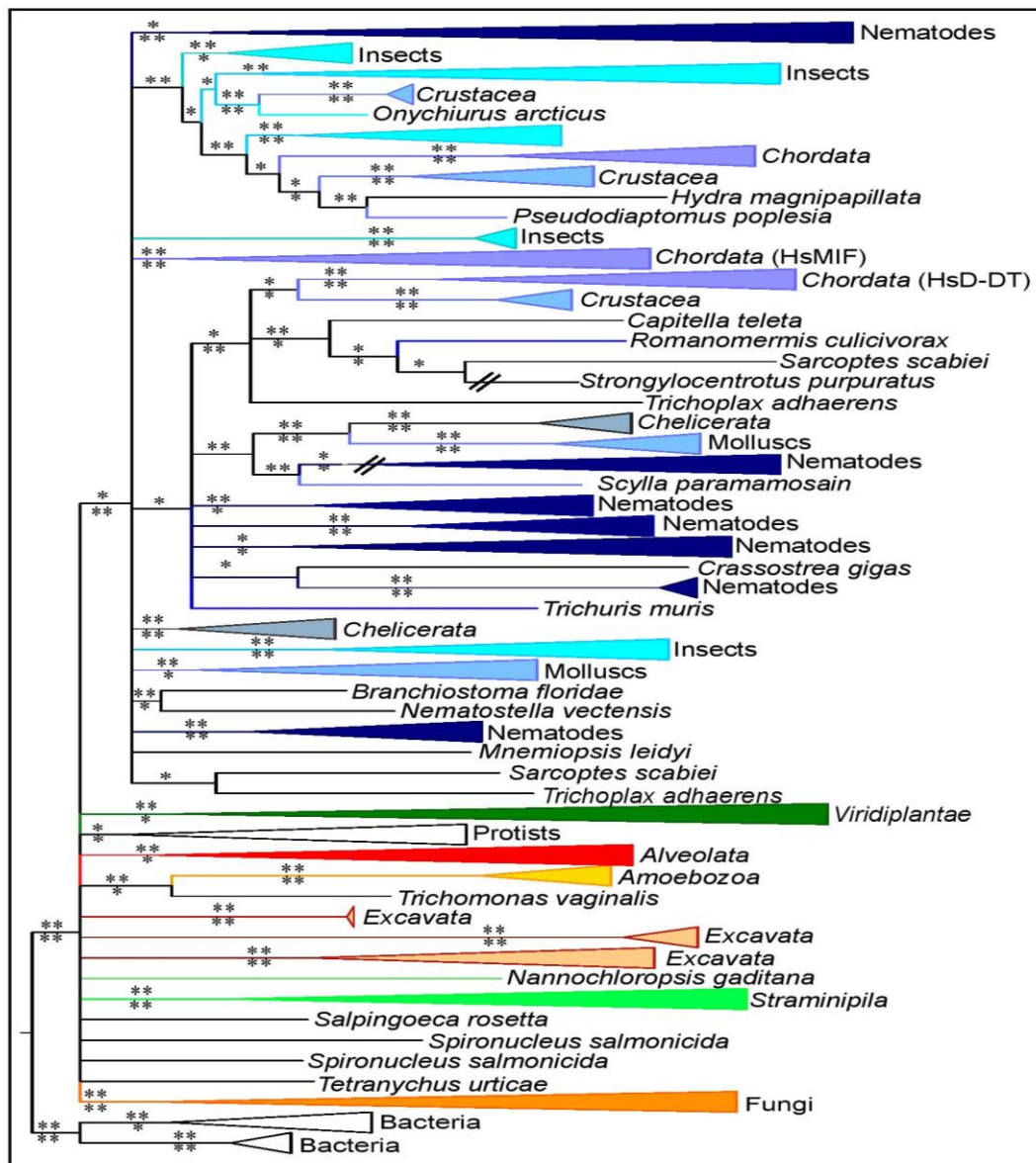


Figure 2: Overview of MIF phylogeny across eukaryotic kingdoms.

Phylogenetic analysis of 720 MIF sequences based on alignments of 113 amino acid characters using the Bayesian Inference (BI) method and rooted using bacterial sequences (see Methods for details). Posterior probability (PP) support values are indicated above the branch for the BI and corresponding approximate likelihood-ratio test (aLRT) support values below the branch for the ML analysis. *: PP > 0.5 or aLRT > 0.7. **: PP > 0.7 or aLRT > 0.9. MIF clusters are coloured according to the taxonomic phyla: metazoa (shades of blue), Stramenopila (light green), Archaeplastida (dark green), Alveolata (red), fungi (dark orange), Excavata (light orange), Amoebozoa (yellow), Rhizaria (pink), and Cryptophyta - Haptophyceae - Picozoa (white). Colour codes of phyla are identical for all figures of the study.

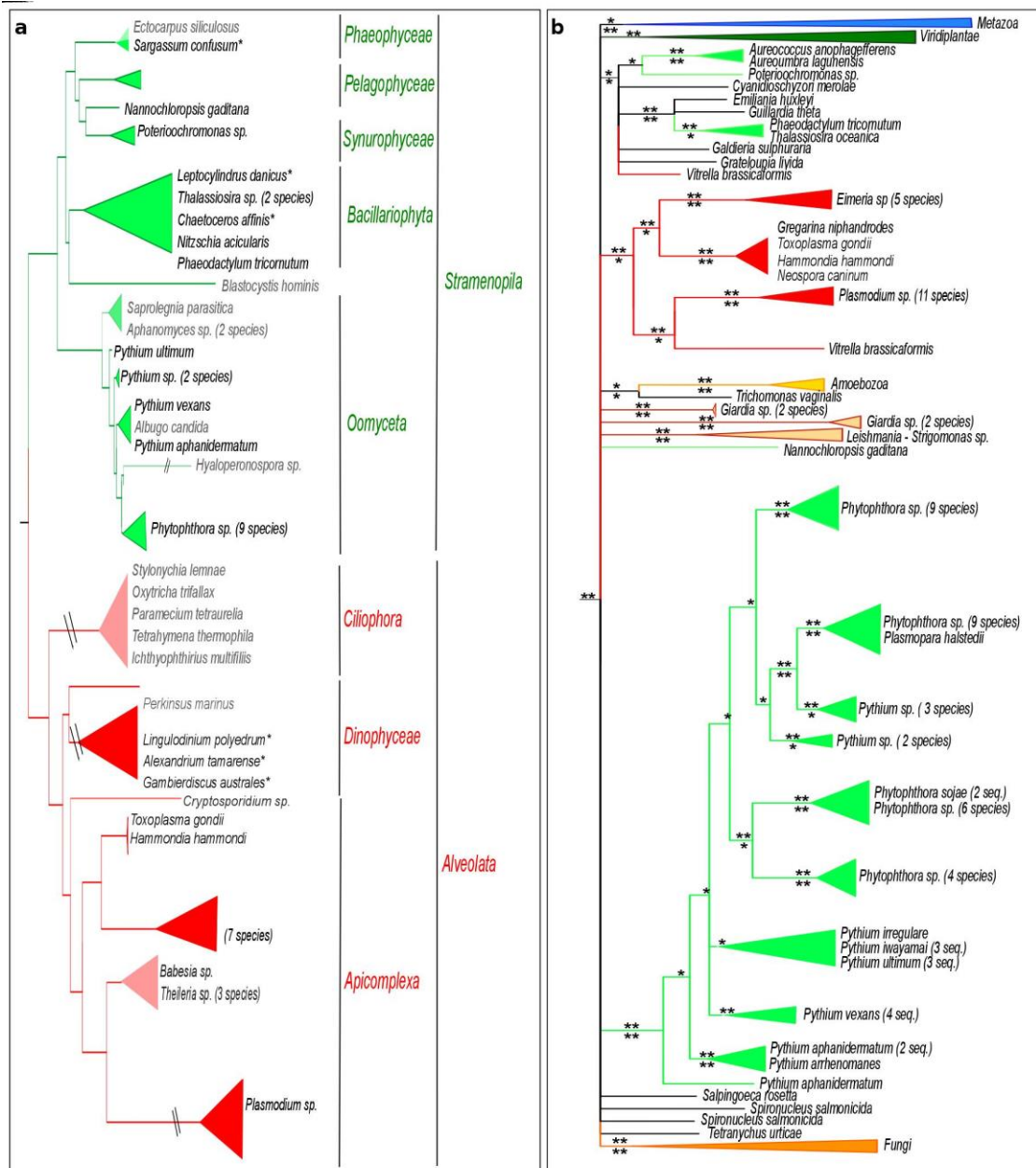


Figure 3: Stramenopila and Alveolata species tree and corresponding MIF phylogeny.

(a) BI tree topology based on *Stramenopila* and *Alveolata* 18S rRNA sequences. (b) BI tree topology for MIF sequences of the same species. Species lacking an identified MIF sequence appear in grey. The symbol // indicates that the respective branch does not follow the scale of the figure. Partial or atypical sequences are marked by a star in the species tree and are absent from the MIF tree. Colours of MIF clusters correspond to the species from which the MIF sequence originated. Posterior probability (PP) support values are indicated above the branch for the BI and corresponding aLRT support values below the branch for the ML analysis. *: PP > 0.5 or aLRT > 0.7. **: PP > 0.7 or aLRT > 0.9. The corresponding tree using the ML method is presented in the Supplementary Figure S4a.

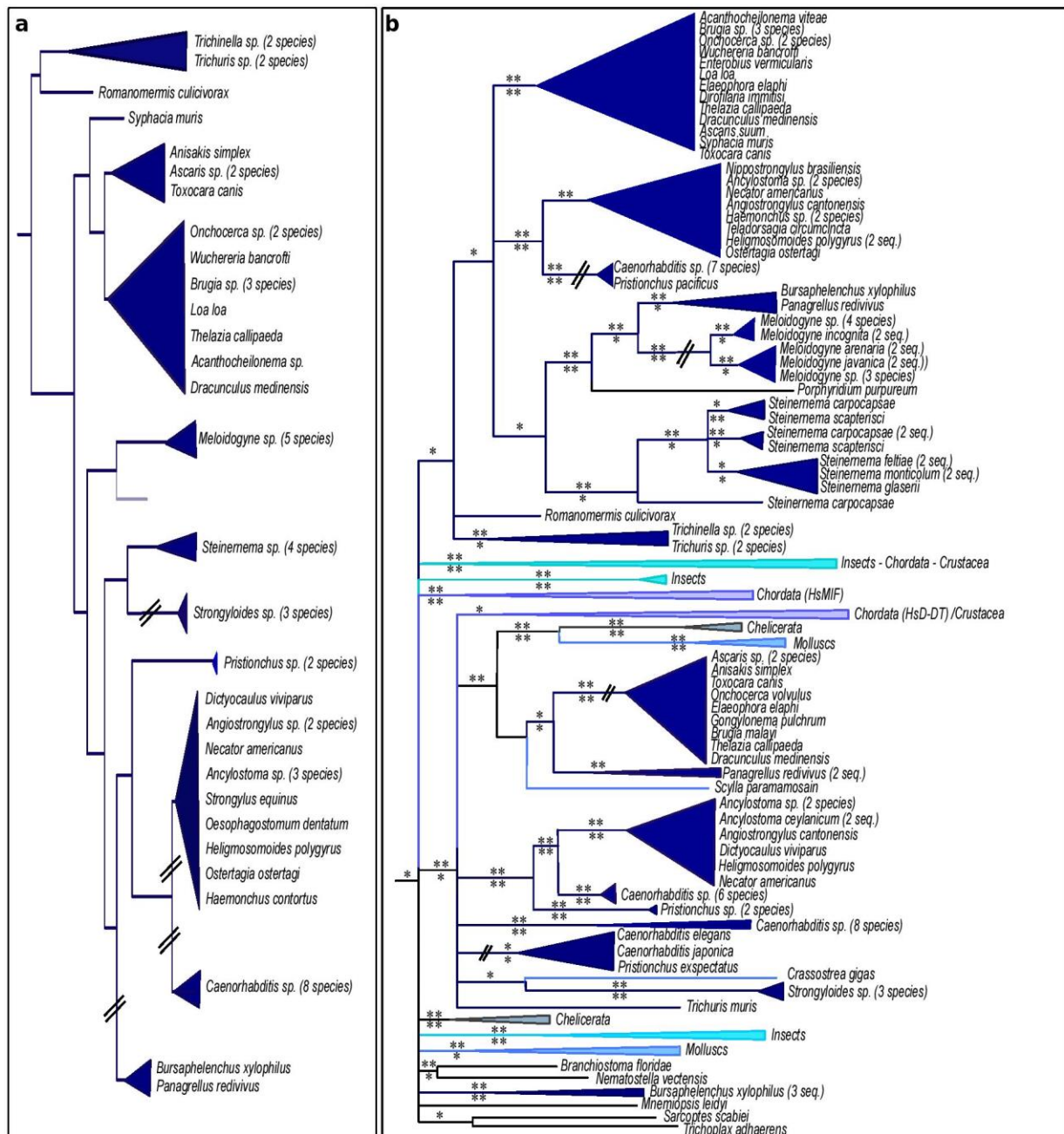


Figure 4: Nematode species tree and corresponding MIF phylogeny.

(a) BI tree topology based on nematode 18S rRNA sequences. (b) BI tree topology for MIF sequences of the same species. *Globodera pallida*, lacking MIF protein sequence is shown in grey. Colours of MIF clusters correspond to the species from which the MIF sequence originated. PP support values are indicated above the branch for the BI and corresponding aLRT support values below the branch for the ML analysis. *: PP > 0.5 or aLRT > 0.7. **: PP > 0.7 or aLRT > 0.9. The corresponding tree using the ML method is presented in Supplementary Figure S5A.

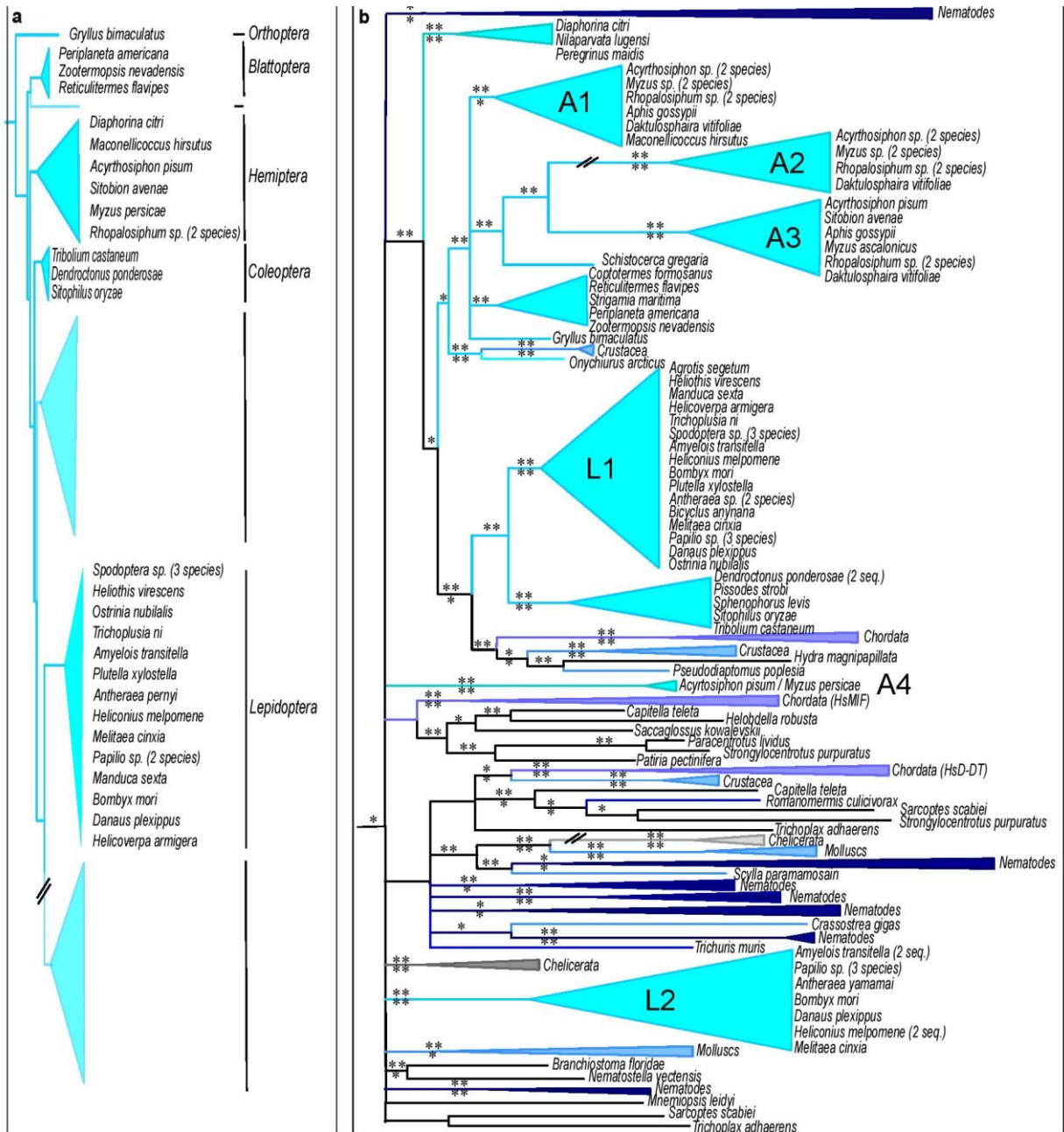


Figure 5: Insect species tree and corresponding MIF phylogeny.

(a) BI tree topology based on insect 18S rRNA sequences. (b) BI tree topology for MIF sequences of the same species. The symbol // indicates that the respective branch does not follow the scale of the figure. Species and orders (*Diptera*, *Hymenoptera*, and *Phthiraptera*) lacking MIF sequences appear in grey. Colours of MIF clusters correspond to the species from which the MIF sequence originated. PP support values are indicated above the branch for the BI and corresponding aLRT support values below the branch for the ML analysis. *: PP > 0.5 or aLRT > 0.7. **: PP > 0.7 or aLRT > 0.9. The corresponding tree using the ML method is presented in Supplementary Figure S5b.

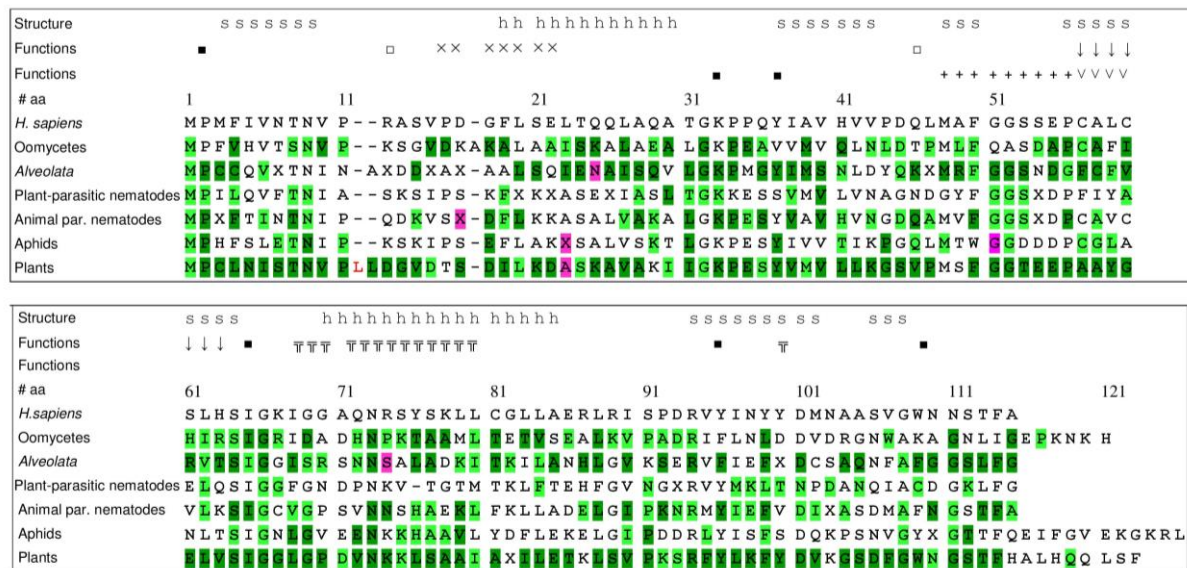


Figure 6: Conservation of amino acid sites and motifs in consensus MIF sequences

Alignment (amino acid one letter code) of human MIF sequences and consensus sequences of oomycetes, *Alveolata*, plant-parasitic nematodes, animal-parasitic nematodes, aphids and plants. Amino acid numbers are given above the human sequence. Hyphens (-) indicate a gap in the respective sequence. Consensus sequences, obtained using Jalview version 2.9, show the most frequent amino acid at each position, or an X when the frequency of the most prominent amino acid is below 40%. Sites under purifying or diversifying selection are coloured in green and pink, respectively. Light and dark green amino acids indicate purifying selection confirmed by two to three, or by four to five different algorithms, respectively. Alpha helices and beta sheets of human MIF are indicated by the letters “h” and “s”, respectively. Amino acids shown to determine specific functions or activities in human MIF are specified by symbols above the sequence. They are involved in the tautomerase activity (■ symbol), oxidoreductase activity (CXXC motif; √ symbol), binding to the CXCR4 receptor (⌋ symbol), binding to the CXCR2 receptor (pseudo ELR motif; □ symbol), interaction between MIF and the CXCR2 receptor (N like loop pattern; + symbol), nuclease activity (PE/DxxxxE motif; × symbol) and DNA damage response protein. (CxxCxxHx(n) zinc finger domain; ↓ symbol).

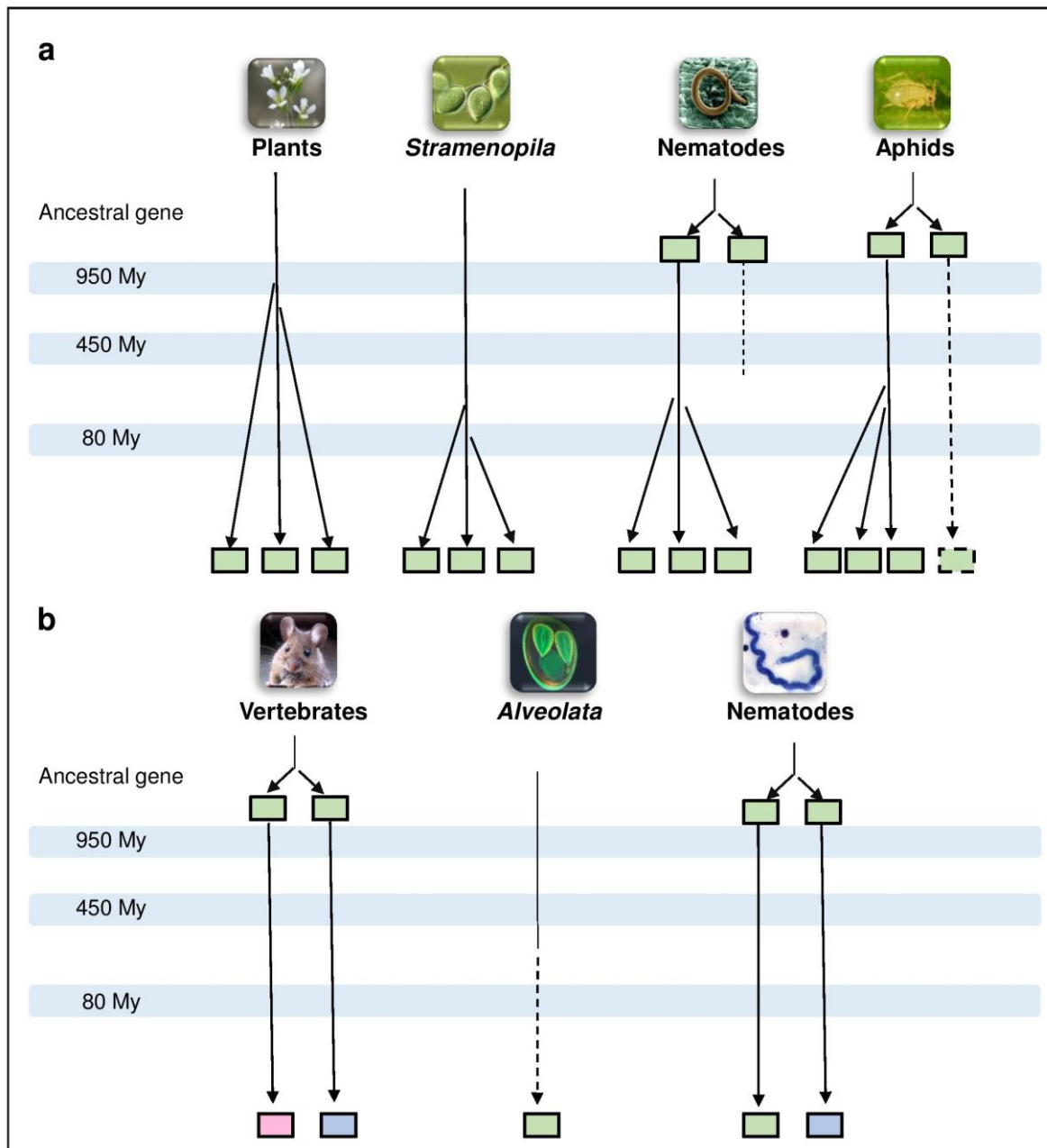


Figure 7: Evolutionary model showing MIF duplication events in phyla from selected eukaryotic kingdoms.

The hypothesized duplication events are indicated on simplified lineage trees reconciling the median MIF number analysis and MIF phylogeny, for plants and plant parasites belonging to protists or metazoans (**a**), and for vertebrate and vertebrate parasites belonging to protists or metazoans (**b**). Light green boxes represent MIF members that cannot be related to the chordate MIF or D-DT variants in our study. Light pink and light blue boxes represent sequences associated with the chordate MIF and D-DT, respectively. The approximate time scale is based on median estimate from TimeTree. For example, 950 MYA represents the period between 757 and 1147 MYA.

Figures supplémentaires

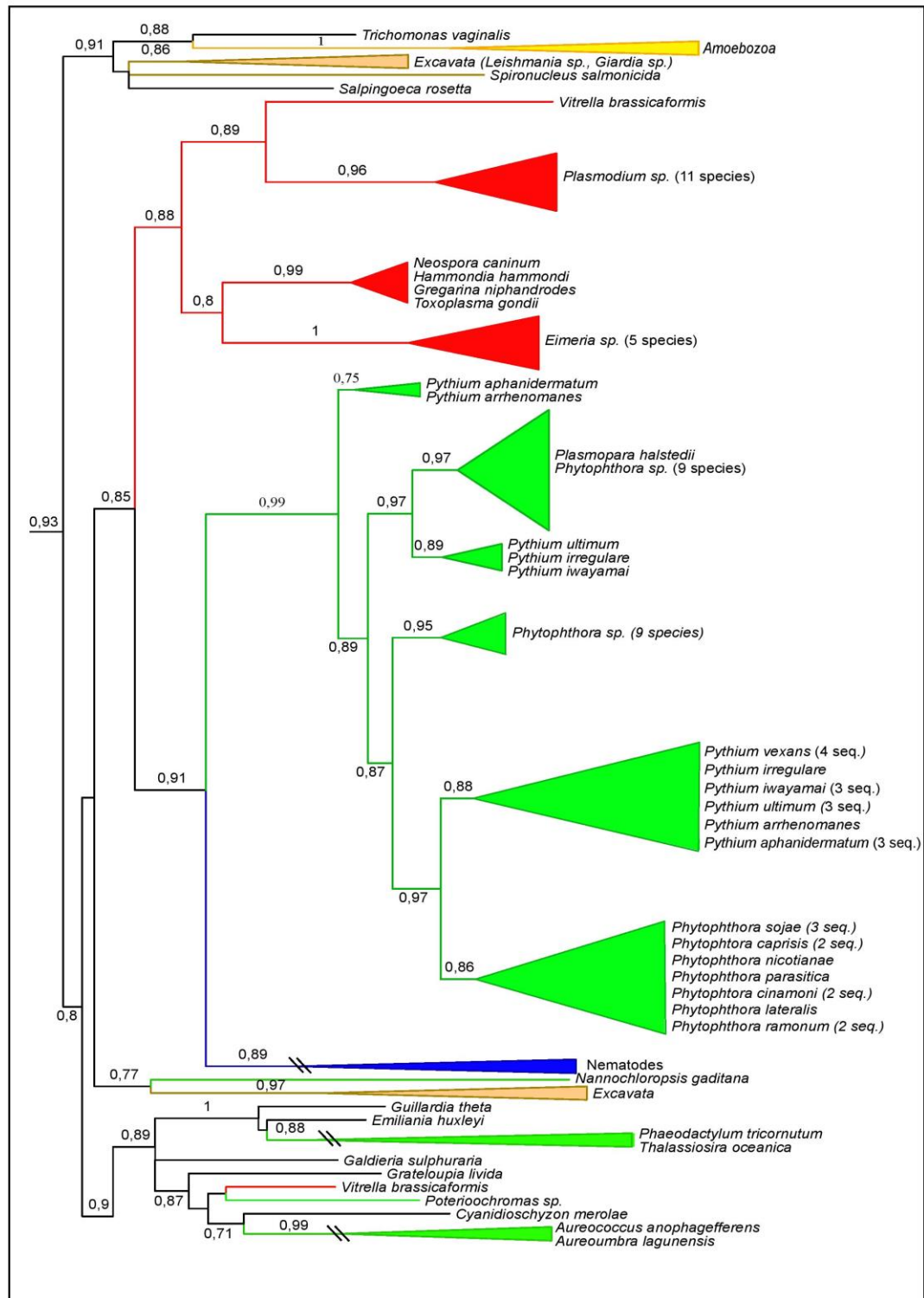


Figure 32 – Phylogénie (Maximum de vraisemblance) des Straménopiles et des Alvéolates

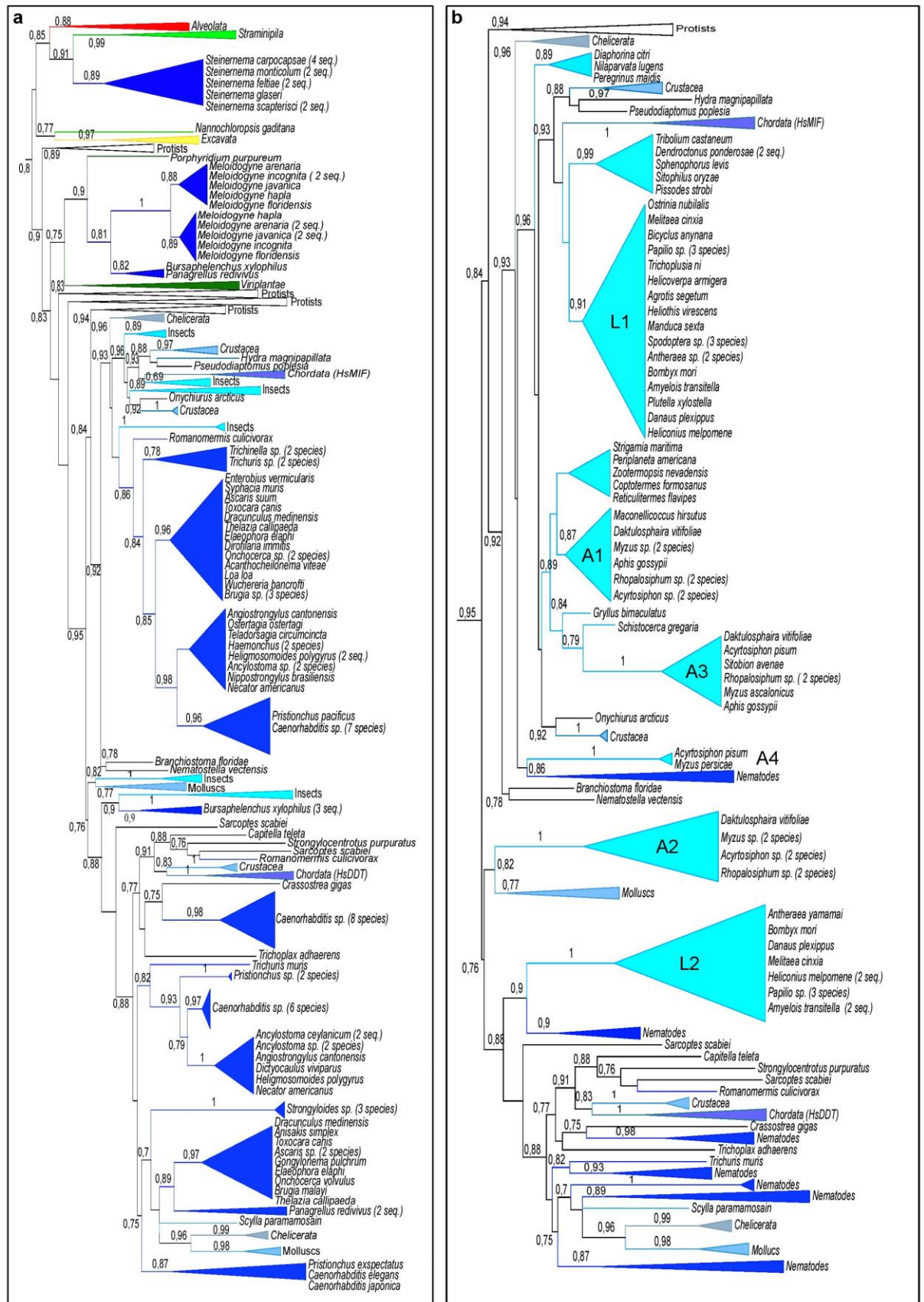


Figure 33 – Phylogénie (Maximum de vraisemblance) des nématodes (A) et des insectes (B)

Phyla	Species	MIF number	Environment	Life-style	Description	Sequenced Genome	Selected for phylogeny reconstruction	ESTs (Number)
Cryptophyta								
	Chroomonas mesostigmatica	0	marine	autotrophic species	algae	present	absent	
	Cryptomonas paramecium	0	marine	autotrophic species	algae	present	absent	
	Guillardia theta	2	marine	autotrophic species	algae	present	yes	
Apusozoa								
	Thecamonas trahens	0	marine	heterotrophic species	free living, biflagellate	present	absent	
Parabasalida								
	Trichomonas vaginalis	1	terrestrial	animal parasites	causes uro genital trichomonoz	present	yes	
Haptophyta								
	Emiliania huxleyi	1	marine	autotrophic species	algae, phytoplankton, causes blooms	present	yes	
Rhizaria --> Cercomonads								
	Bigelowiella natans	0	marine	autotrophic species	model organisms, algae	present	absent	
Rhizaria --> Cercozoa								
	Lotharella oceanica	0	marine	autotrophic species	chlorarachniophyte algae, plankton	present	absent	
	Paulinella chromatophora	0	freshwater	heterotrophic species	amoeba - free living - photosynthetic	present	absent	
	Plasmodiophora brassicae	0	terrestrial	plant parasites	causes clubroot	present	absent	
Rhizaria --> Foraminifera								
	Reticulomyxa filosa	0	freshwater	heterotrophic species	foraminifers	present	absent	
Excavates --> Diplomonads								
	Giardia intestinalis	2	terrestrial	animal parasites	causes giardiasis, flagellated parasite	present	yes	
	Giardia lamblia	2	terrestrial	animal parasites	causes giardiasis, flagellated parasite	present	yes	
	Spironucleus salmonicida	2	terrestrial	animal parasites	fish parasite, systemic infection in salmonid fish	present	yes	
Excavates --> Kinetoplastids								
	Leishmania braziliensis	1	terrestrial	animal parasites	causes leishmaniasis	present	no	
	Leishmania donovani	2	terrestrial	animal parasites	causes leishmaniasis	present	yes	
	Leishmania infantum	2	terrestrial	animal parasites	causes leishmaniasis	present	yes	
	Leishmania major	2	terrestrial	animal parasites	causes leishmaniasis	present	yes	
	Leishmania mexicana	2	terrestrial	animal parasites	causes leishmaniasis	present	yes	
	Leishmania panamensis	1	terrestrial	animal parasites	causes leishmaniasis	present	no	
	Phytomonas sp	0	terrestrial	plant parasites	conveyed by bugs	present	absent	

Partie V : Annexes

	Strigomonas culicis	1	terrestrial	animal parasites	infects and colonizes the Aedes aegypti	present	yes	
	Trypanosoma brucei	0	terrestrial	animal parasites	causes trypanosomiasis, flagellate parasite	present	absent	
	Trypanosoma cruzi	0	terrestrial	animal parasites	causes trypanosomiasis, flagellate parasite	present	absent	
	Trypanosoma rangeli	0	terrestrial	animal parasites	causes trypanosomiasis, flagellate parasite	present	absent	
Excavates --> Heterolobozoa								
	Naegleria gruberi	0	fresh water	heterorophic species	amoeba, free - living organism	present	absent	
Amoebozoa --> Mycetozoa								
	Dictyostelium discoideum	0	terrestrial	heterorophic species	soil - living amoeba	present	absent	
	Dictyostelium fasciculatum	0	terrestrial	heterorophic species	soil - living amoeba	present	absent	
	Dictyostelium purpureum	0	terrestrial	heterorophic species	soil - living amoeba	present	absent	
	Polysphondylium pallidum	0	terrestrial	heterorophic species	amoeba	present	absent	
Amoebozoa --> Archamoebae								
	Entamoeba dispar	0	terrestrial	animal parasites	pathogenic amoeba, causes amoebiasis, dysentery	present	absent	
	Entamoeba histolytica	1	terrestrial	animal parasites	pathogenic amoeba, causes amoebiasis, dysentery	present	yes	
	Entamoeba invadens	2	terrestrial	animal parasites	pathogenic amoeba, causes amoebiasis, dysentery	present	yes	
	Entamoeba nuttalli	0	terrestrial	animal parasites	pathogenic amoeba, causes amoebiasis, dysentery	present	absent	
Amoebozoa --> Discosea								
	Acanthamoeba castellanii	0	fresh water	animal parasites	amoeba, bacterivores, cause acanthamebiasis	present	absent	
Straminipila								
	Blastocystis hominis	0	terrestrial	animal parasites	causess stomach problem in humans	present	absent	
Straminipila --> Synurophyceae								
	Poterioochromonas sp.	1	freshwater	autotrophic species	heterokont algae, photosynthetic	absent	yes	0
	Synura sp.	1	freshwater	autotrophic species	heterokont algae, photosynthetic	absent	no	0
Straminipila --> Bacillariophyta								
	Chaetoceros affinis	1	marine	autotrophic species	diatom (brown algae), phytoplankton	absent	no	0
	Leptocylindrus danicus	1	marine	autotrophic species	diatom (brown algae), phytoplankton	absent	no	0

Partie V : Annexes

	Nitzschia sp.	1	marine	autotrophic species	diatom (brown algae), phytoplankton, produces neurotoxin, grow at cold temperature	absent	yes	0
	Phaeodactylum tricornutum	1	marine	autotrophic species	diatom (brown algae) phytoplankton	present	yes	
	Thalassiosira oceanica	1	marine	autotrophic species	phytoplankton, diatom (brown algae), symbiosis with bacteria	present	yes	
Straminipila --> Oomycetes								
	Albugo candida	0	terrestrial	plant parasites	obligatory biotrophic plant parasite, cause white blister rust	present	absent	
	Aphanomyces astaci	0	terrestrial	animal parasites	crayfish parasites,	present	absent	
	Aphanomyces invadans	0	terrestrial	animal parasites	fish parasite	present	absent	
	Hyaloperonospora arabidopsidis	0	terrestrial	plant parasites	causes mildew	present	absent	
	Phytophthora capsici	4	terrestrial	plant parasites	mildew	absent	yes	58.277
	Phytophthora cinamoni	4	terrestrial	plant parasites	mildew	absent	yes	330.482
	Phytophthora infestans	2	terrestrial	plant parasites	mildew	present	yes	
	Phytophthora kernoviae	2	terrestrial	plant parasites	mildew	present	yes	
	Phytophthora lateralis	3	terrestrial	plant parasites	mildew	present	yes	
	Phytophthora nicotianae	3	terrestrial	plant parasites	cause root rot, crown rot, leaf infection	present	yes	
	Phytophthora parasitica	3	terrestrial	plant parasites	soilborne pathogen	present	yes	
	Phytophthora ramorum	4	terrestrial	plant parasites	causes the disease sudden oak death (SOD)	present	yes	
	Phytophthora sojae	5	terrestrial	plant parasites	causes damping - off	present	yes	
	Plasmopara halstedii	1	terrestrial	plant parasites	causes sunflowers mildew	present	yes	
	Pythium aphanidermatum	4	terrestrial	plant parasites	mildew	present	yes	
	Pythium arrhenomanes	2	terrestrial	plant parasites	mildew	present	yes	
	Pythium irregulare	2	terrestrial	plant parasites	mildew	present	yes	
	Pythium iwayamai	4	terrestrial	plant parasites	mildew	present	yes	
	Pythium ultimum	4	terrestrial	plant parasites	mildew	present	yes	
	Pythium vexans	4	terrestrial	plant parasites	mildew	present	yes	

Partie V : Annexes

	<i>Saprolegnia diclina</i>	0	terrestrial	animal parasites	amphibian and fish parasites, causes oomycosis, secondary pathogen	present	absent	
Straminipila --> Phaeophyceae								
	<i>Ectocarpus siliculosus</i>	0	marine	autotrophic species	brown algae	present	absent	
	<i>Sargassum vulgare</i>	1	marine	autotrophic species	brown algae	absent	no	0
Straminipila --> Pelagophyceae								
	<i>Aureococcus anophagefferens</i>	1	marine	autotrophic species	algae, causes algal blooms	present	yes	
	<i>Aureoumbra lagunensis</i>	1	marine	heterotrophic species	picoplankton	absent	yes	0
Straminipila --> Eustigmatophyceae								
	<i>Nannochloropsis gaditana</i>	2	fresh water	autotrophic species	microalgae, phytoplankton	present	yes	
Alveolata --> Dinophyceae								
	<i>Alexandrium tamarense</i>	1	marine	autotrophic species	phytoplankton, algae	absent	no	10.887
	<i>Gambierdiscus excentricus</i>	1	marine	autotrophic species	brown macroalgae, produce toxins, algae	absent	no	0
	<i>Lingulodinium polyedrum</i>	1	marine	autotrophic species	algae, motile dinoflagellates, production of Yessotoxins (YTXs)	absent	no	3.640
	<i>Perkinsus marinus</i>	0	marine	animal parasites	pathogen of oysters, , algae	present	absent	
	<i>Symbiodinium minutum</i>	2	marine	autotrophic species	endosymbiotic dinoflagellates, algae	present	yes	
Alveolata --> Apicomplexa								
	<i>Babesia bigemina</i>	0	terrestrial	animal parasites	causes babesiosis	present	absent	
	<i>Babesia bovis</i>	0	terrestrial	animal parasites	causes babesiosis	present	absent	
	<i>Cryptosporidium hominis</i>	0	terrestrial	animal parasites	obligate humans parasite, causes cryptosporidiosis	present	absent	
	<i>Eimeria acervulina</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria brunetti</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria maxima</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria mitis</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria necatrix</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria praecox</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Eimeria tenella</i>	1	terrestrial	animal parasites	causes coccidiosis	present	yes	
	<i>Gregarina niphandrodes</i>	1	terrestrial	animal parasites	invertebrate parasite	present	yes	

Partie V : Annexes

	Hammondia hammondi	1	terrestrial	animal parasites	dogs parasite	present	yes	
	Neospora caninum	1	terrestrial	animal parasites	causes neosporosis	present	yes	
	Plasmodium berghei	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium chabaudi	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium cynomolgi	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium falciparum	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium fragile	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium inui	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium knowlesi	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium reichenowi	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium vinckei	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium vivax	1	terrestrial	animal parasites	causes malaria	present	yes	
	Plasmodium yoelii	1	terrestrial	animal parasites	causes malaria	present	yes	
	Theileria annulata	0	terrestrial	animal parasites	causes theileriosis (east coast fever)	present	absent	
	Theileria orientalis	0	terrestrial	animal parasites	causes theileriosis (east coast fever)	present	absent	
	Theileria parva	0	terrestrial	animal parasites	causes theileriosis (east coast fever)	present	absent	
	Toxoplasma gondii	1	terrestrial	animal parasites	causes toxoplasmosis	present	yes	
Alveolata --> Chromerida								
	Vitrella brassicaformis	2	marine	autotrophic species	isolated from the Great Barrier Reef, mixotrophe	present	yes	
Alveolata --> Ciliophora								
	Ichthyophthirius multifiliis	0	fresh water	animal parasites	fish parasite (causes ichthyophthiriosis)	present	absent	
	Oxytricha trifallax	0	fresh water	heterorophic species	free living, very mobile	present	absent	
	Paramecium tetraurelia	0	fresh water	heterorophic species	phagocytose bacteria	present	absent	
	Stylonichia lemnae	0	fresh water	heterorophic species	free living, carnivorous (eat other protozoans and bacteria)	present	absent	
	Tetrahymena thermophila	0	fresh water	heterorophic species	commensalistic	present	absent	
Opithokonts --> Ichthyospora								
	Capsaspora owczarzaki	0	fresh water	animal parasites	amoeboid, snail parasite and kill the sporocysts of flatworm	present	absent	
Opithokonts --> Nucleariids								
	Fonticula alba	0	terrestrial	heterorophic species	amoeboid - free living	present	absent	

Partie V : Annexes

Opithokonts --> Choanoflagellata								
	<i>Monosiga brevicollis</i>	0	marine	heterorophic species	free living, nanoplankton	present	absent	
	<i>Salpingoeca rosetta</i>	1	fresh water	heterorophic species	free living	present	yes	
Opisthokonts --> Dikarya --> Basidiomycota --> Pucciniomycotina								
	<i>Agaricostilbum hyphaenes</i>	0	terrestrial	heterorophic species	occurs commonly on palm litter, saprophytic fungi	present	absent	
	<i>Chionosphaera apobasidialis</i>	0	terrestrial	plant parasites	parasite of lichens	present	absent	
	<i>Cronartium quercuum</i>	0	terrestrial	plant parasites	causes plant diseases of pinus	present	absent	
	<i>Erythrobasidium hasegawianum</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Heterogastridium pycnidioideum</i>	0	terrestrial	animal parasites	fungi parasites	present	absent	
	<i>Hyalopycnis blepharistoma</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Leucosporidiella creatinivora</i>	0	terrestrial	heterorophic species	saprophytic fungi, isolated from permafrost soil	present	absent	
	<i>Melampsora laricis</i>	0	terrestrial	plant parasites	most devastating and widespread pathogen of poplars	present	absent	
	<i>Meredithblackwellia eburnea</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Mixia osmundae</i>	0	terrestrial	plant parasites	intracellular parasite of ferns, obligatory plant pathogens	present	absent	
	<i>Naiadella fluitans</i>	0	fresh water	animal parasites	mycoparasitic	present	absent	
	<i>Naohidea sebacea</i>	0	terrestrial	animal parasites	mycoparasitic	present	absent	
	<i>Puccinia graminis</i>	0	terrestrial	plant parasites	causes stem rusts, obligatory plant pathogens	present	absent	
	<i>Puccinia triticina</i>	0	terrestrial	plant parasites	causes wheat leaf rust, obligatory plant pathogens	present	absent	
	<i>Rhodospodium toruloides</i>	0	terrestrial	animal parasites	isolated from wood pulp from conifers, link to bovine mastitis	present	absent	
	<i>Rhodotorula graminis</i>	0	terrestrial	animal parasites	causes infections in animals	present	absent	
	<i>Rhodotorula minuta</i>	0	terrestrial	animal parasites	human parasite	present	absent	
	<i>Septobasidium sp.</i>	0	terrestrial	animal parasites	entomopathogens	present	absent	
	<i>Sporobolomyces paraseus</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
Dikarya --> Basidiomycota --> Ustilaginomycotina								
	<i>Acaromyces ingoldii</i>	0	terrestrial	heterorophic species	causes mites mortality, saprophytic fungi	present	absent	
	<i>Ceraceosorus guamensis</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	

Partie V : Annexes

	<i>Exobasidium vaccinii</i>	0	terrestrial	plant parasites	plant pathogen that grow on Ericaceae	present	absent	
	<i>Jaminalia</i> sp.	0	terrestrial	heterotrophic species	saprophytic fungi	present	absent	
	<i>Malassezia globosa</i>	0	terrestrial	animal parasites	opportunistic pathogens of humans, causes skin disorders	present	absent	
	<i>Malassezia sympodialis</i>	0	terrestrial	animal parasites	opportunistic pathogens of humans, causes skin disorders	present	absent	
	<i>Meira mitonrushii</i>	0	terrestrial	heterotrophic species	saprophytic fungi, endophyte	present	absent	
	<i>Melanotaenium endogenum</i>	0	terrestrial	plant parasites	Bedstraw smut, pathogen of <i>Galium mollugo</i>	present	absent	
	<i>Pseudozyma antarctica</i>	0	terrestrial	heterotrophic species	saprophytic fungi	present	absent	
	<i>Pseudozyma hubeiensis</i>	0	terrestrial	plant parasites	epiphytic fungus	present	absent	
	<i>Sporisorium reilianum</i>	0	terrestrial	plant parasites	Head Smut pathogen of Sorghum	present	absent	
	<i>Testicularia cyperi</i>	0	terrestrial	plant parasites	colonizes <i>Rhynchospora</i> sp, plant pathogens	present	absent	
	<i>Tilletiaria anomala</i>	0	terrestrial	heterotrophic species	saprophytic fungi - decayed wood	present	absent	
	<i>Tilletiopsis washingtonensis</i>	0	terrestrial	heterotrophic species	saprophytic fungi	present	absent	
	<i>Ustilago maydis</i>	0	terrestrial	plant parasites	plant pathogen, causes corn smut	present	absent	
Dikarya --> Basidiomycota --> Agaricomycotina --> Agaricales								
	<i>Agaricus bisporus</i>	1	terrestrial	heterotrophic species	portobello mushroom, saprophytic fungi	present	yes	
	<i>Amanita muscaria</i>	1	terrestrial	heterotrophic species	fly agaric or fly amanita, mycorrhizal	present	yes	
	<i>Armillaria mellea</i>	1	terrestrial	plant parasites	honey fungus, plant pathogen	absent	yes	30
	<i>Coprinopsis cinerea</i>	1	terrestrial	heterotrophic species	saprophytic fungi, gray shap	present	yes	
	<i>Cortinarius glaucopus</i>	0	terrestrial	plant parasites	saprophytic fungi, blue-foot webcap, ectomycorrhizal	present	absent	
	<i>Fistulina hepatica</i>	1	terrestrial	plant parasites	beefsteack fungus, parasitic	present	yes	
	<i>Galerina marginata</i>	1	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal - produces amatoxins	present	yes	
	<i>Gymnopus luxurians</i>	1	terrestrial	heterotrophic species	saprophytic fungi	present	yes	
	<i>Hebeloma cylindrosporum</i>	1	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal	present	yes	
	<i>Hypholoma sublateritium</i>	1	terrestrial	heterotrophic species	saprophytic fungi, brick cap	present	yes	

Partie V : Annexes

	<i>Hypsizygus marmoreus</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Laccaria amethystina</i>	1	terrestrial	plant parasites	amethyst deceiver, mycorrhizal	present	yes	
	<i>Laccaria bicolor</i>	1	terrestrial	plant parasites	ectomycorrhizal	present	yes	
	<i>Leucoagaricus gongylophorus</i>	1	terrestrial	heterorophic species	saprophytic fungi - associates with ants	present	yes	
	<i>Panellus stipticus</i>	0	terrestrial	heterorophic species	saprophytic fungi, bitter oyster	present	absent	
	<i>Pleurotus ostreatus</i>	1	terrestrial	heterorophic species	saprophytic fungi - white -rot wood decay fungus	present	yes	
	<i>Rhodocollybia butyracea</i>	0	terrestrial	heterorophic species	saprophytic fungi, butterfly collybia	present	absent	
	<i>Schizophyllum commune</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Termitomyces</i> sp. J132	1	terrestrial	heterorophic species	saprophytic fungi - food source for termites	present	no	
Dikarya --> Basidiomycota --> Agaricomycotina --> Atheliales								
	<i>Piloderma croceum</i>	1	terrestrial	plant parasites	ectomycorrhizal	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Auriculariales								
	<i>Auricularia subglabra</i>	1	terrestrial	heterorophic species	saprophytic fungi, produce a white rot on dead wood	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Polyporales								
	<i>Bjerkandera adusta</i>	1	terrestrial	plant parasites	saprophytic fungi, smoky polypore, plant pathogen, cause white rot in live trees	present	yes	
	<i>Ceriporiopsis subvermispora</i>	1	terrestrial	heterorophic species	saprophytic fungi, crust fungi that cause a white rot	present	yes	
	<i>Daedalea quercina</i>	1	terrestrial	heterorophic species	saprophytic fungi, oak mazegill, causes brown rot	present	yes	
	<i>Dichomitus squalens</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Exidia glandulosa</i>	1	terrestrial	heterorophic species	black jelly roll, saprophytic fungi	present	yes	
	<i>Fibroporia radiculosa</i>	0	terrestrial	heterorophic species	saprophytic fungi - brown rot fungus	present	absent	
	<i>Fibulorhizoctonia</i> sp.	0	terrestrial	heterorophic species	cuckoo fungus, saprophytic fungis, use by termite	present	absent	
	<i>Fomitopsis pinicola</i>	1	terrestrial	heterorophic species	stem decay fungus, red belt conk, saprophytic fungi	present	yes	
	<i>Ganoderma lucidum</i>	1	terrestrial	heterorophic species	saprophytic fungi, grow on both coniferous and hardwood species	present	yes	
	<i>Gelatoporia subvermispora</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	

Partie V : Annexes

	<i>Laetiporus sulphureus</i>	1	terrestrial	heterorophic species	sulphur polyphore, saprophytic fungi	present	yes	
	<i>Lentinus tigrinus</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Phanerochaete carnosae</i>	1	terrestrial	plant parasites	plant pathogen that infects plane trees	present	yes	
	<i>Phlebiopsis gigantea</i>	1	terrestrial	heterorophic species	saprophytic fungi, causes white rot on conifer	present	yes	
	<i>Plicaturopsis crispa</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Polyporus arcularius</i>	1	terrestrial	heterorophic species	causes white rot, saprophytic fungi	present	yes	
	<i>Postia placenta</i>	1	terrestrial	heterorophic species	saprophytic fungi, causes brown rot	present	yes	
	<i>Pycnoporus cinnabarinus</i>	1	terrestrial	plant parasites	white rot saprobic fungus, plant pathogen infecting plane and mango trees	present	yes	
	<i>Trametes cinnabarina</i>	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	
	<i>Trametes versicolor</i>	1	terrestrial	heterorophic species	turkey tail, saprophytic fungi	present	yes	
	<i>Wolfiporia cocos</i>	2	terrestrial	heterorophic species	hoelen, saprophytic fungi	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Boletales								
	<i>Boletus edulis</i>	0	terrestrial	plant parasites	penny bun, formes ectomycorrhizal associations with living trees	present	absent	
	<i>Coniophora puteana</i>	1	terrestrial	heterorophic species	saprophytic fungi - causes wet rot in wood	present	yes	
	<i>Gyrodon lividus</i>	0	terrestrial	plant parasites	the alder bolete, mycorrhizal association with alder	present	absent	
	<i>Hydnomerulius pinastri</i>	1	terrestrial	heterorophic species	saprophytic fungi, causes a brown - rot, called the spiny dry rot fungus	present	yes	
	<i>Paxillus involutus</i>	1	terrestrial	plant parasites	saprophytic fungicommonly known as the brown roll-rim, ectomycorrhizal	present	yes	
	<i>Paxillus rubicundulus</i>	1	terrestrial	plant parasites	ectomycorrhizal	present	yes	
	<i>Pisolithus tinctorius</i>	1	terrestrial	plant parasites	mycchorrhizal, also know as the dyemaker's puffball	present	yes	
	<i>Rhizopogon saebris</i>	0	terrestrial	plant parasites	false truffles, ectomycorrhizal	present	absent	
	<i>Serpula lacrymans</i>	1	terrestrial	heterorophic species	saprophytic fungi - causes dry rot - most damaging destroyer of indoor wood	present	yes	

Partie V : Annexes

	Suillus granulatus	0	terrestrial	plant parasites	ectomycorrhizal, weeping or granulated bolete	present	absent	
	Suillus hirtellus	0	terrestrial	plant parasites	ectomycorrhizal	present	absent	
	Suillus luteus	0	terrestrial	plant parasites	saprophytic fungi, slippery jack or sticky bun, ectomycorrhizal	present	absent	
	Suillus pictus - spraguei	0	terrestrial	plant parasites	painted slippery cap, mycorrhizal	present	absent	
Dikarya --> Basidiomycota --> Agaricomycotina --> Corticiales								
	Cylindrobasidium torrendii	1	terrestrial	heterotrophic species	saprophytic fungi	present	yes	
	Punctularia strigosozonata	1	terrestrial	heterotrophic species	saprophytic fungi, causes a white rot	present	yes	
	Schizopora paradoxa	1	terrestrial	heterotrophic species	saprophytic fungi - decayed frondose wood	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Hymenochaetales								
	Fomitiporia mediterranea	1	terrestrial	plant parasites	cause esca of grapevine	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Gloeophyllales								
	Gloeophyllum trabeum	1	terrestrial	heterotrophic species	brown rot fungus, saprophytic fungi	present	yes	
	Neolentinus lepideus	1	terrestrial	heterotrophic species	scaly lentinus, saprophytic fungi	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Russulales								
	Heterobasidion irregulare	1	terrestrial	plant parasites	tree root rotting pathogenic fungus	present	yes	
	Peniophora sp.	1	terrestrial	plant parasites	plant pathogens	present	yes	
	Stereum hirsutum	1	terrestrial	plant parasites	false turkey tail, plant pathogen infecting peach trees	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Jaapiales								
	Jaapia argillacea	2	terrestrial	heterotrophic species	saprophytic fungi, produces a type of brown rot	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Gomphales								
	Ramaria rubella	1	terrestrial	plant parasites	mycorrhizal, red coral	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Geastrales								
	Sphaerobolus stellatus	1	terrestrial	heterotrophic species	saprophytic fungi, shotgun or cannonball fungus	present	yes	
Dikarya --> Basidiomycota --> Agaricomycotina --> Tremellomycetes								
	Cryptococcus gattii	0	terrestrial	animal parasites	cause cryptococcosis	present	absent	
	Cryptococcus neoformans	0	terrestrial	animal parasites	cause cryptococcosis	present	absent	
	Dioszegia cryoxerica	0	terrestrial	heterotrophic species	psychrophilic yeasts from polar desert soils	present	absent	
	Fellomyces penicillatus	0	terrestrial	heterotrophic species	saprophytic fungi	present	absent	
	Fibulobasidium inconspicuum	0	terrestrial	heterotrophic species	saprophytic fungi	present	absent	

Partie V : Annexes

	Sistotremastrum suecicum	0	terrestrial	heterorophic species	saprophytic fungi, cause a brown rot of various tree species	present	absent	
	Tremella enphala	0	terrestrial	animal parasites	parasitic on another species of fungus	present	absent	
	Tremella mesenterica	0	terrestrial	animal parasites	yellow brain, parasite of wood decay fungi	present	absent	
	Trichosporon oleaginosus	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
Dikarya --> Basidiomycota --> Agaricomycotina --> Cantharellales								
	Botryobasidium botryosum	0	terrestrial	heterorophic species	saprophytic fungis	present	absent	
	Cantharellus anzutake	0	terrestrial	plant parasites	golden chanterelle, ectomycorrhizal	present	absent	
	Clavulina sp.	0	terrestrial	plant parasites	saprophytic fungis - ectomycorrhizal	present	absent	
	Hydnum rufescens	0	terrestrial	plant parasites	symbiotic fungi identified in orchids	present	absent	
	Rhizoctonia solani	1	terrestrial	plant parasites	plant pathogenus fungi, causes collar and root rot	present	yes	
	Tulasnella calospora	0	terrestrial	plant parasites	mychorhizien	present	absent	
Dikarya --> Basidiomycota --> Agaricomycotina --> Wallemiomycetes								
	Wallemia ichthyophaga	0	terrestrial	heterorophic species	halophilic - saprophytic fungi	present	absent	
	Wallemia sebi	0	terrestrial	heterorophic species	xerophilic - saprophytic fungi	present	absent	
Dikarya --> Basidiomycota --> Agaricomycotina --> Dacrymycetales								
	Calocera cornea	0	terrestrial	heterorophic species	saprophytic fungi - grows on decaying wood	present	absent	
	Calocera viscosa	0	terrestrial	heterorophic species	saprophytic fungi - yellow stagshorn	present	absent	
	Dacryopinax primogenitus	0	terrestrial	heterorophic species	saprophytic fungi - causes extensive decay in dead wood	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Pezizomycetes								
	Ascobolus immersus	0	terrestrial	heterorophic species	coprophilous	present	absent	
	Ascodesmis nigricans	0	terrestrial	heterorophic species	coprophilous	present	absent	
	Caloscypha fulgens	0	terrestrial	plant parasites	saprophytic fungi - infect dormant seeds of Picea sitchensis - spring orange peel fungus	present	absent	
	Choiromyces venosus	0	terrestrial	plant parasites	ectomycorrhizal - pig truffle	present	absent	
	Gyromitra esculenta	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Kalaharituber pfeilii	0	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal	present	absent	
	Morchella importuna	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	

Partie V : Annexes

	Plectania melastoma	0	terrestrial	heterorophic species	saprophytic fungi (neither mycorrhizal nor edible)	present	absent	
	Pyronema confluens	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Sarcoscypha coccinea	0	terrestrial	heterorophic species	saprophytic fungi - scarlet elf cup	present	absent	
	Terfezia boudieri	0	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal - desert truffle	present	absent	
	Tuber melanosporum	0	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal- black truffle	present	absent	
	Wilcoxina mikolae	0	terrestrial	plant parasites	ectomycorrhizal	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Orbiliomycetes								
	Arthrobotrys oligospora	0	terrestrial	heterorophic species	carnivorous fungus (nematode-trapping)	present	absent	
	Dactylellina haptotyla	0	terrestrial	heterorophic species	carnivorous fungus (nematode-trapping)	present	absent	
	Monacrosporium haptotylum	0	terrestrial	heterorophic species	carnivorous fungus (nematode-trapping)	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Eurotiomycetes								
	Ajellomyces capsulatus	1	terrestrial	animal parasites	cause histoplasmosis	present	yes	
	Ajellomyces dermatitidis	1	terrestrial	animal parasites	cause histoplasmosis	present	yes	
	Arthroderma otae	0	terrestrial	animal parasites	anthropophilic saprotrop, colonizes the dead skin, fungal infection of nail	present	absent	
	Aspergillus sp.	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Aspergillus sydowii	0	terrestrial	animal parasites	cause of aspergillosis in sea fan corals	present	absent	
	Blastomyces gilchristii	1	terrestrial	animal parasites	human parasites, causes blastomycosis	present	yes	
	Emmonsia crescens	1	terrestrial	animal parasites	causes adiaspiromycosis	present	yes	
	Emmonsia parva	1	terrestrial	animal parasites	causes adiaspiromycosis	present	yes	
	Eurotium rubrum	0	terrestrial	heterorophic species	saprophytic fungi - live in Dead Sea water	present	absent	
	Gymnascella aurantiaca	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Gymnascella citrina	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Histoplasma capsulatum	1	terrestrial	animal parasites	causes histoplasmosis - opportunist human pathogen	present	yes	
	Monascus ruber	0	terrestrial	heterorophic species	corn silage mold	present	absent	

Partie V : Annexes

	Paracoccidioides brasiliensis	1	terrestrial	animal parasites	causes paracoccidioidomycosis	present	yes	
	Paracoccidioides lutzii	1	terrestrial	animal parasites	causes paracoccidioidomycosis	present	yes	
	Penicilium oxalicum	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	Talaromyces marneffei	0	terrestrial	animal parasites	opportunistic human pathogen	present	absent	
	Thermoascus aurantiacus	0	terrestrial	heterorophic species	thermophilic fungus, saprophytic fungi	present	absent	
	Trichophyton rubrum	0	terrestrial	animal parasites	anthropophilic saprotrop, colonizes the dead skin, fungal infection of nail	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Dothideomycetes								
	Acidomyces richmondensis	0	terrestrial	heterorophic species	acidophilic filamentous fungi	present	absent	
	Ascomycota	0	terrestrial	plant parasites	pathogene of plants (particularly cereal crops)	present	absent	
	Aureobasidium pullulans	0	terrestrial	plant parasites	epiphyte, endophyte, pathogen opportunist of humans	present	absent	
	Bipolaris oryzae	0	terrestrial	plant parasites	causes brown spot disease in rice	present	absent	
	Dothistroma septosporum	0	terrestrial	plant parasites	causes red band needle blight	present	absent	
	Leptosphaeria maculans	0	terrestrial	plant parasites	causal agent of blackleg disease	present	absent	
	Macrophomina phaseolina	0	terrestrial	plant parasites	causes damping off, seedling blight and collar rot	present	absent	
	Neofusicoccum parvum	0	terrestrial	plant parasites	causing serious damage to Eucalyptus	present	absent	
	Parastagonospora nodorum	0	terrestrial	plant parasites	major fungal pathogen of wheat	absent	absent	15.973
	Pseudocercospora fijiensis	0	terrestrial	plant parasites	plant pathogens	present	absent	
	Verruconis gallopava	0	terrestrial	animal parasites	Human parasitic phaeohyphomycosis	present	absent	
	Zymoseptoria brevis	1	terrestrial	plant parasites	causes septoria leaf blotch	present	no	
Dikarya --> Ascomycota --> Pezizomycotina --> Lecanoromycetes								
	Xanthoria parietina	0	terrestrial	heterorophic species	lichens	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Leotiomyces								
	Acephala macrosclerotiorum	0	terrestrial	heterorophic species	saprophytic fungi - forms ectomycorrhizas on Pinus sylvestris	present	absent	

Partie V : Annexes

	<i>Amorphotheca resinae</i>	0	terrestrial	heterorophic species	saprophytic fungi - can live in tanks of diesel	present	absent	
	<i>Blumeria graminis</i>	0	terrestrial	plant parasites	causes barley powdery mildew on grasses including cereals	present	absent	
	<i>Botrytis cinerea</i>	0	terrestrial	plant parasites	necrotrophic fungus, botrytis bunch rot	present	absent	
	<i>Bulgaria inquinans</i>	0	terrestrial	heterorophic species	saprophytic fungi - black bulgar	present	absent	
	<i>Cadophora malorum</i>	0	terrestrial	plant parasites	plant pathogen, causes side rot in apple and pear	present	absent	
	<i>Chalara longipes</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Coccomyces strobili</i>	0	terrestrial	plant parasites	plant pathogen - causes Cherry leaf spot	present	absent	
	<i>Erysiphe necator</i>	0	terrestrial	plant parasites	causes powdery mildew of grape	present	absent	
	<i>Leptodontium</i> sp.	0	terrestrial	heterorophic species	saprophytic fungi - responsible for Sooty Blotch disease	present	absent	
	<i>Loramyces juncicola</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Loramyces macrosporus</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Meliniomyces bicolor</i>	0	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal - root-associated fungi	present	absent	
	<i>Meliniomyces variabilis</i>	0	terrestrial	plant parasites	saprophytic fungi - ectomycorrhizal - root-associated fungi	present	absent	
	<i>Oidiodendron maius</i>	0	terrestrial	heterorophic species	mycorrhizal fungus	present	absent	
	<i>Phialocephala scopiformis</i>	0	terrestrial	plant parasites	endophyte	present	absent	
	<i>Rhizoscyphus ericae</i>	0	terrestrial	plant parasites	mycorrhizal fungus	present	absent	
	<i>Rutstroemia firma</i>	0	terrestrial	heterorophic species	saprophytic fungi - grows on wood	present	absent	
	<i>Sclerotinia borealis</i>	0	terrestrial	plant parasites	cause a plant pathogen infecting barley	present	absent	
	<i>Sclerotinia sclerotiorum</i>	0	terrestrial	plant parasites	cause a disease called white mold	present	absent	
	<i>Sporothrix insectorum</i>	1	terrestrial	heterorophic species	decomposing plant material and pathogen opportunist (causative agent of sporotrichosis)	present	yes	
	<i>Thelebolus microsporus</i>	0	terrestrial	heterorophic species	saprobic fungus	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Sordariomycetes								

Partie V : Annexes

	<i>Acremonium alcalophilum</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Anthostoma avocetta</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Apiospora montagnei</i>	0	terrestrial	plant parasites	causes kernel blight on barley	present	absent	
	<i>Ceratocystis</i> sp.	0	terrestrial	plant parasites	plant pathogens, causing diseases such as oak wilt and pineapple black rot	absent	absent	3.432
	<i>Clonostachys rosea</i>	0	terrestrial	animal parasites	endophyte, saprophytic fungi, parasite of other fungi and of nematodes	present	absent	
	<i>Colletotrichum caudatum</i>	0	terrestrial	plant parasites	phytopathogens	present	absent	
	<i>Colletotrichum zoysiae</i>	0	terrestrial	plant parasites	most common and destructive of plant pathogenic fungi	present	absent	
	<i>Coniochaeta</i> sp.	0	terrestrial	plant parasites	saprotrophs, endophytes and pathogens of trees	present	absent	
	<i>Corollospora maritima</i>	0	marine	heterorophic species	saprophytic fungi	present	absent	
	<i>Cryphonectria parasitica</i>	0	terrestrial	plant parasites	causes chestnut blight	present	absent	
	<i>Daldinia eschscholzii</i>	0	terrestrial	heterorophic species	saprotrophic	present	absent	
	<i>Drechmeria coniospora</i>	0	terrestrial	animal parasites	entomopathons	absent	absent	0
	<i>Fusarium</i> sp.	0	terrestrial	heterorophic species	saprophytic fungus - produces mycotoxins in cereal crops	present	absent	
	<i>Glomerella graminicola</i>	0	terrestrial	plant parasites	causes anthracnose in many cereal	present	absent	
	<i>Hypoxyton</i> sp.	0	terrestrial	heterorophic species	saprophytic fungi, found on dead wood	present	absent	
	<i>Ilyonectria</i> sp.	0	terrestrial	plant parasites	symptoms of root rot on herbaceous plants	present	absent	
	<i>Lindra thalassiae</i>	0	marine	plant parasites	parasites of various sea grasses	present	absent	
	<i>Lolliopopaia minuta</i>	0	terrestrial	heterorophic species	saprophytic fungi	present	absent	
	<i>Magnaporthe oryzae</i>	0	terrestrial	plant parasites	also known as rice blast fungus, rice rotten neck	present	absent	
	<i>Melanconium juglandinum</i>	0	terrestrial	plant parasites	plant pathogen	present	absent	
	<i>Melanospora tiffanyae</i>	0	terrestrial	animal parasites	mycoparasitic fungus	present	absent	
	<i>Microascus trigonosporus</i>	0	terrestrial	animal parasites	saprotrophic, opportunistic human pathogen	present	absent	

Partie V : Annexes

	<i>Microdochium bolleyi</i>	0	terrestrial	plant parasites	plant pathogen - causes root rot in flax and wheat	present	absent	
	<i>Myceliophthora thermophila</i>	0	terrestrial	heterorrophic species	saprophytic fungi	present	absent	
	<i>Myrothecium inundatum</i>	0	terrestrial	heterorrophic species	saprotrophic, endophyte	present	absent	
	<i>Neurospora crassa</i>	0	terrestrial	heterorrophic species	model organism, can survive on minimal media	present	absent	
	<i>Niesslia exilis</i>	0	terrestrial	heterorrophic species	saprotrophic	present	absent	
	<i>Periconia macrospina</i>	0	terrestrial	heterorrophic species	plant symbiotic species	present	absent	
	<i>Plectosphaerella cucumerina</i>	0	terrestrial	plant parasites	pathogen of several plant causing fruit, root and collar rot	absent	absent	0
	<i>Podospora curvicolle</i>	0	terrestrial	heterorrophic species	coprophilous	present	absent	
	<i>Pseudogymnoascus</i>	0	terrestrial	heterorrophic species	saprophytic fungi	present	absent	
	<i>Sodiomyces alkalinus</i>	0	terrestrial	heterorrophic species	saprophytic fungi	present	absent	
	<i>Sordaria macrospora</i>	0	terrestrial	heterorrophic species	saprophytic fungi - coprophilous species	absent	absent	27
	<i>Thielavia antarctica</i>	0	terrestrial	heterorrophic species	coprophilous	present	absent	
	<i>Thielavia terrestris</i>	0	terrestrial	animal parasites	can cause infections in humans	present	absent	
	<i>Thozetella</i> sp.	0	terrestrial	plant parasites	root endophyte of <i>Populus</i>	present	absent	
	<i>Torpedospora radiata</i>	0	terrestrial	heterorrophic species	saprobic and cellulolytic fungi	present	absent	
	<i>Trichoderma reesei</i>	0	terrestrial	animal parasites	saprotrophic, pathogens of fungi	present	absent	
	<i>Valetoniellopsis laxa</i>	0	terrestrial	heterorrophic species	saprotrophic	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Xylonomycetes								
	<i>Symbiotaphrina buchneri</i>	0	terrestrial	animal parasites	endosymbiote of beetles	present	absent	
	<i>Trinosporium guianense</i>	0	terrestrial	heterorrophic species	live in association with a tropical tree	present	absent	
	<i>Xylona heveae</i>	0	terrestrial	plant parasites	endophyte of rubber trees	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Saccharomycotina								
	<i>Candida albicans</i>	0	terrestrial	animal parasites	cause candidiasis or thrush in humans	present	absent	
	<i>Candida tenuis</i>	0	terrestrial	heterorrophic species	saprophytic fungi - ferment xylose	present	absent	
	<i>Eremothecium gossypii</i>	0	terrestrial	plant parasites	cause stigmatomycosis	present	absent	
	<i>Lachancea quebecensis</i>	0	terrestrial	heterorrophic species	associated with maple, oak and birch	present	absent	
	<i>Lachancea thermotolerans</i>	0	terrestrial	heterorrophic species	usually associated with fruits and <i>Drosophila</i>	present	absent	

Partie V : Annexes

	Millerozyma farinosa	0	terrestrial	animal parasites	rare cases of pathogenic on humans	absent	absent	0
	Scheffersomyces stipitis	0	terrestrial	heterorophic species	saprophytic fungi - ferment xylose to ethanol	present	absent	
	Spathaspora passalidatum	0	terrestrial	heterorophic species	study for this ability to assimilate xylose	present	absent	
	Yarrowia lipolytica	0	terrestrial	heterorophic species	saprophytic fungi - can use hydrocarbons	present	absent	
Dikarya --> Ascomycota --> Pezizomycotina --> Taphrinomycotina								
	Protomyces lactucaedelis	0	terrestrial	plant parasites	forms galls on different plants	present	absent	
	Saitoella complicata	2	terrestrial	heterorophic species	saprophytic fungi, feeds on decaying matter	present	no	
	Schizosaccharomyces pombe	0	terrestrial	heterorophic species	model organism	present	absent	
Glomeromycota								
	Rhizophagus irregularis	1	terrestrial	plant parasites	mycorrhizal fungus	present	yes	
Mortierellomycotina								
	Mortierella sp.	0	terrestrial	heterorophic species	saprotrophs in soil	present	absent	
Mucoromycotina								
	Absidia glauca	1	terrestrial	animal parasites	causes zygomycosis and mucormycosis	absent	yes	3
	Absidia repens	1	terrestrial	animal parasites	causes zygomycosis and mucormycosis	present	yes	
	Backusella circina	0	terrestrial	heterorophic species	thermophilic fungus	present	absent	
	Blakeslea trispora	1	terrestrial	plant parasites	pathogen autotrophic species, source of beta carotene	present	yes	
	Hesseltinella vesiculosa	0	terrestrial	plant parasites	saprophytic fungi or facultative plant parasites or facultative pathogens of humans (mucoromycotina)	present	absent	
	Lichtheimia corymbifera	1	terrestrial	animal parasites	saprotrophic mold and opportunistic animals pathogen	present	yes	
	Lichtheimia ramosa	1	terrestrial	animal parasites	saprotrophic mold and opportunistic animals pathogen	absent	yes	0
	Mucor ambiguus	1	terrestrial	plant parasites	plant pathogen	absent	yes	0
	Mucor circinelloides	1	terrestrial	plant parasites	plant pathogen	present	yes	
	Parasitella parasitica	1	terrestrial	animal parasites	mycoparasite	present	yes	
	Phycomyces blakesleeanus	1	terrestrial	heterorophic species	saprophytic fungi	present	yes	

Partie V : Annexes

	Rhizopus deleamar	1	terrestrial	plant parasites	infecting carrots, pineapples, ...	present	yes	
	Rhizopus microsporus	1	terrestrial	plant parasites	infecting maize, sunflower, ...	present	yes	
	Syncephalastrum racemosum	1	terrestrial	heterotrophic species	nonpathogenic organism, common species in soil	present	yes	
	Thamnidium elegans	0	terrestrial	animal parasites	saprophytic fungi mucuromycete	present	absent	
	Umbelopsis ramanniana	1	terrestrial	plant parasites	cosmopolitan saprotroph and endophyte of healthy and declining conifers	present	yes	
Entomophthoromycotina								
	Basidiobolus meristosporus	1	terrestrial	animal parasites	human pathogens - commensal in frogs	present	no	
Blastocladiomycota								
	Allomyces magrogynus	0	terrestrial	heterotrophic species	-	present	absent	
	Catenaria anguillulae	0	terrestrial	animal parasites	facultative parasite of nematode parasites	present	absent	
Chytridiomycota								
	Batrachochytrium dendrobatidis	0	terrestrial	animal parasites	amphibian chytrid fungus	present	absent	
	Gonapodya prolifera	1	fresh-water	heterotrophic species	carnivorous	present	no	
	Spizellomyces punctatus	0	terrestrial	animal parasites	parasites of organisms that attack nematodes	present	absent	
Neocallimastigomycota								
	Piromyces finnis	1	terrestrial	heterotrophic species	obligate anaerobes	present	yes	
Microsporidia								
	Anncaliia algerae	0	fresh-water	animal parasites	infects mosquitoes	present	absent	
	Anncaliia algerae	0	terrestrial	animal parasites	infects mosquitoes	present	absent	
	Antonospora locustae	0	terrestrial	animal parasites	obligate intracellular parasites	present	absent	
	Encephalitozoon intestinalis	0	terrestrial	animal parasites	cause microsporidiosis	present	absent	
	Encephalitozoon romaleae	0	terrestrial	animal parasites	obligate intracellular parasites	present	absent	
	Nematocida parisii	0	terrestrial	animal parasites	nematode-killer from Paris	present	absent	
	Nosema ceranae	0	terrestrial	animal parasites	parasite of Apis cerana, the Asiatic honey bee	present	absent	
	Ordospora colligata	0	terrestrial	animal parasites	obligatory intracellular parasite	absent	absent	0
Cryptomycota								
	Rozella allomyces	0	terrestrial	animal parasites	obligate endoparasite of the water mold	present	absent	

Partie V : Annexes

Opisthokonts --> Metazoa --> Diploblasta --> Placozoa								
	Trichoplax adhaerens	2	terrestrial	heterorophic species	very flat organisms	present	yes	
Diploblasta --> Porifera								
	Amphimedon queenslandica	0	marine	heterorophic species	sponge	present	absent	
Diploblasta --> Ctenophora								
	Mnemiopsis leidyi	1	marine	heterorophic species	warty comb jelly or sea walnut	present	yes	
Diploblasta --> Cnidaria								
	Hydra magnipapillata	1	fresh water	heterorophic species	hydra	present	yes	
	Nematostella vectensis	1	marine	heterorophic species	Starlet sea anemone	present	yes	
Opisthokonts --> Metazoa --> Bilateria --> Deuterostomia --> Hemichordata								
	Saccaglossus kowalevskii	1	marine	heterorophic species	Acorn worm	present	yes	
Bilateria --> Deuterostomia --> Echinodermata								
	Paracentrotus lividus	1	marine	heterorophic species	Sea urchin	absent	yes	141217
	Patiria miniata	1	marine	heterorophic species	Blue bat star	present	yes	
	Strongylocentrotus purpuratus	2	marine	heterorophic species	Sea urchin	present	yes	
Bilateria --> Deuterostomia --> Chordata								
Urochordates								
	Ciona intestinalis	2	marine	heterorophic species	Vase tunicate	present	yes	
Cephalochordates								
	Branchiostoma floridae	1	marine	heterorophic species	Florida lancelet	present	yes	
Cyclostomes								
	Myxine glutinosa	1	marine	heterorophic species	Atlantic hagfish	present	yes	
	Petromyzon marinus	1	marine	heterorophic species	Sea lamprey	present	yes	
Cartilaginous fish								
	Callorhynchus milli	2	marine	heterorophic species	Australian ghostshark	present	yes	
Bony fish								
	Cyprinus carpio	2	marine	heterorophic species	Common carp	present	yes	
	Dicentrarchus labrax	1	marine	heterorophic species	European bass	present	yes	
	Fundulus heteroclitus	2	marine	heterorophic species	Mummichog	present	yes	
	Haplochromis chilotes	1	fresh-water	heterorophic species	cichlid (fish)	absent	no	35763
	Osmerus mordax	2	marine	heterorophic species	Rainbow smelt	present	yes	
	Salmo salar	2	marine	heterorophic species	Atlantic salmon	present	yes	
	Takifugu rubripes	2	marine	heterorophic species	Japanese puffer	present	yes	
	Xiphophorus maculatus	2	marine	heterorophic species	Southern platyfish	present	yes	
Amphibians								
	Andrias davidianus	1	terrestrial	heterorophic species	Chinese giant salamander	absent	yes	1504

Partie V : Annexes

	Nanorana parkeri	2	terrestrial	heterorophic species	High Himalaya frog	present	yes	
	Rana catesbeiana	2	terrestrial	heterorophic species	American bullfrog	absent	yes	29138
	Xenopus laevis	2	terrestrial	heterorophic species	African clawed frog	present	yes	
Reptiles --> Crocodylia								
	Alligator mississippiensis	2	terrestrial	heterorophic species	Crocodylian	present	yes	
	Alligator sinensis	2	terrestrial	heterorophic species	Crocodylian	absent	yes	1
Reptiles --> Squamata								
	Gekko japonicus	2	terrestrial	heterorophic species	Geckos	absent	yes	6957
	Protobothrops mucrosquamatus	2	terrestrial	heterorophic species	Snakes	absent	yes	0
Reptiles --> Chelonii								
	Chelonia mydas	2	terrestrial	heterorophic species	Green sea turtle	present	yes	
	Chrysemys picta	2	terrestrial	heterorophic species	Painted turtle	present	yes	
Birds								
	Anas platyrhynchos	2	terrestrial	heterorophic species	Mallard (duck)	present	yes	
	Calidris pugnax	2	terrestrial	heterorophic species	Ruff	absent	yes	0
	Ficedula albicollis	2	terrestrial	heterorophic species	Collared flycatcher	present	yes	
	Gallus gallus	2	terrestrial	heterorophic species	Red junglefowl	present	yes	
	Lepidothrix coronata	2	terrestrial	heterorophic species	blue-crowned manakin	absent	yes	0
Mammalian								
	Callithrix jacchus	2	terrestrial	heterorophic species	Common marmoset	present	yes	
	Camelus dromedarius	2	terrestrial	heterorophic species	Dromedary	absent	yes	7
	Canis lupus	2	terrestrial	heterorophic species	Gray wolf	present	no	
	Ceratotherium simum	2	terrestrial	heterorophic species	White rhinoceros	present	yes	
	Chinchilla lanigera	2	terrestrial	heterorophic species	Long-tailed chinchilla	present	yes	
	Chrysochloris asiatica	2	terrestrial	heterorophic species	Cape golden mole	absent	yes	0
	Colobus angolensis	2	terrestrial	heterorophic species	Angola colobus	absent	yes	0
	Condylura cristata	2	terrestrial	heterorophic species	Star-nosed mole	absent	yes	0
	Dasypus novemcinctus	2	terrestrial	heterorophic species	Nine-banded armadillo	present	yes	
	Dipodomys ordii	2	terrestrial	heterorophic species	Kangaroo rats	present	yes	
	Elephantulus edwardii	2	terrestrial	heterorophic species	Cape elephant shrew	absent	yes	0
	Eptesicus fuscus	2	terrestrial	heterorophic species	Big brown bats	absent	yes	0
	Erinaceus europaeus	2	terrestrial	heterorophic species	European hedgehog	present	yes	

Partie V : Annexes

	<i>Felis catus</i>	2	terrestrial	heterorophic species	Cats	present	yes	
	<i>Fukomys damarensis</i>	2	terrestrial	heterorophic species	Damaraland mole-rat	present	yes	
	<i>Galeopterus variegatus</i>	2	terrestrial	heterorophic species	Sunda flying lemur	absent	yes	0
	<i>Homo sapiens</i>	2	terrestrial	heterorophic species	human	present	yes	
	<i>Ictidomys tridecemlineatus</i>	2	terrestrial	heterorophic species	Thirteen-lined ground squirrel	present	yes	
	<i>Jaculus jaculus</i>	2	terrestrial	heterorophic species	Jerboas	present	yes	
	<i>Leptonychotes weddellii</i>	2	marine	heterorophic species	Weddell seal	absent	yes	0
	<i>Lipotes vexillifer</i>	2	fresh-water	heterorophic species	Baiji	absent	yes	0
	<i>Loxodonta africana</i>	2	terrestrial	heterorophic species	African bush elephant	present	no	
	<i>Melopsittacus undulatus</i>	1	terrestrial	heterorophic species	Budgerigar	present	yes	
	<i>Meriones unguiculatus</i>	2	terrestrial	heterorophic species	Mongolian gerbil	absent	yes	1442
	<i>Mesocricetus auratus</i>	2	terrestrial	heterorophic species	Golden hamster	present	yes	
	<i>Microcebus murinus</i>	2	terrestrial	heterorophic species	Gray mouse lemur	present	yes	
	<i>Microtus ochrogaster</i>	2	terrestrial	heterorophic species	Prairie vole	present	yes	
	<i>Monodelphis domestica</i>	2	terrestrial	heterorophic species	Gray short-tailed opossum	present	yes	
	<i>Mus musculus</i>	2	terrestrial	heterorophic species	house mouse	present	yes	
	<i>Mustela putorius furo</i>	2	terrestrial	heterorophic species	Ferret	present	yes	
	<i>Nomascus leucogenys</i>	2	terrestrial	heterorophic species	Northern white-cheeked gibbon	present	yes	
	<i>Octodon degus</i>	2	terrestrial	heterorophic species	octodontid rodents	present	yes	
	<i>Odobenus rosmarus</i>	2	marine	heterorophic species	Walrus	present	yes	
	<i>Orcinus orca</i>	2	marine	heterorophic species	Killer whale	present	yes	
	<i>Orycteropus afer</i>	2	terrestrial	heterorophic species	Aardvark	present	yes	
	<i>Otolemur garnettii</i>	2	terrestrial	heterorophic species	Northern greater galago	present	yes	
	<i>Ovis aries</i>	2	terrestrial	heterorophic species	Sheep	present	yes	
	<i>Pan paniscus</i>	2	terrestrial	heterorophic species	Bonobo	present	yes	
	<i>Pongo abelii</i>	2	terrestrial	heterorophic species	Sumatran orangutan	present	yes	
	<i>Propithecus coquereli</i>	2	terrestrial	heterorophic species	Coquerel sifaka (lemur)	present	yes	
	<i>Pteropus vampyrus</i>	2	terrestrial	heterorophic species	Megabat	present	yes	
	<i>Rattus norvegicus</i>	2	terrestrial	heterorophic species	Brown rat	present	yes	
	<i>Saimiri boliviensis</i>	2	terrestrial	heterorophic species	Squirrel monkey	present	yes	

Partie V : Annexes

	<i>Sus scrofa</i>	2	terrestrial	heterorophic species	Wild boar	present	yes	
	<i>Trichechus manatus</i>	2	marine	heterorophic species	West Indian manatee	present	yes	
	<i>Tupaia chinensis</i>	2	terrestrial	heterorophic species	Tree shrew	present	yes	
Bilateria --> Protostomia --> Molluscs								
	<i>Aplysia californica</i>	2	marine	heterorophic species	California sea slug	present	yes	
	<i>Biomphalaria glabrata</i>	2	fresh water	heterorophic species	Snail	present	yes	
	<i>Crassostrea gigas</i>	2	marine	heterorophic species	Pacific oyster	present	yes	
	<i>Haliotis discus</i>	1	marine	heterorophic species	Sea snail	absent	yes	8154
	<i>Haliotis diversicolor</i>	1	marine	heterorophic species	Sea snail	absent	no	7435
	<i>Meretrix meretrix</i>	1	marine	heterorophic species	Saltwater clams	absent	yes	3131
	<i>Mytilus galloprovincialis</i>	1	marine	heterorophic species	Mediterranean mussel	absent	yes	19756
	<i>Octopus bimaculoides</i>	1	marine	heterorophic species	Octopus	present	yes	
	<i>Pinctada fucata</i>	1	marine	heterorophic species	Akoya pearl oyster	absent	yes	1380
Bilateria --> Protostomia --> Platyhelminthes								
	<i>Clonorchis sinensis</i>	0	terrestrial	animal parasites	human liver fluke	present	absent	
	<i>Echinococcus granulosus</i>	0	terrestrial	animal parasites	cyclophyllid cestode	present	absent	
	<i>Echinococcus multilocularis</i>	0	terrestrial	animal parasites	cyclophyllid tapeworm	present	absent	
	<i>Hymenolepis microstoma</i>	0	terrestrial	animal parasites	rodent tapeworm	present	absent	
	<i>Schistosoma haematobium</i>	0	terrestrial	animal parasites	trematodes	present	absent	
	<i>Schistosoma japonicum</i>	0	terrestrial	animal parasites	trematodes	present	absent	
	<i>Schistosoma mansoni</i>	0	terrestrial	animal parasites	trematodes	present	absent	
	<i>Schmidtea mediterranea</i>	0	fresh water	heterorophic species	triclad	present	absent	
	<i>Taenia solium</i>	0	terrestrial	animal parasites	pork tapeworm	present	absent	
Bilateria --> Protostomia --> Annelids								
	<i>Capitella teleta</i>	2	marine	heterorophic species	annelid worm	present	yes	
	<i>Helobdella robusta</i>	1	fresh water	heterorophic species	leech	present	yes	
Bilateria --> Protostomia --> Brachiopoda								
	<i>Lingula anatina</i>	2	marine	heterorophic species		present	no	
Bilateria --> Protostomia --> Rotifera								
	<i>Adineta vaga</i>	0	fresh water	heterorophic species	bdelloid rotifer	present	absent	
Bilateria --> Protostomia --> Tardigrades								
	<i>Hypsibius dujardini</i>	0	terrestrial	heterorophic species	tardigrade	present	absent	
Bilateria --> Protostomia --> Nematodes								
Classe I								

Partie V : Annexes

	Trichuris muris	2	terrestrial	animal parasites	parasite of mice	present	yes	
	Trichuris suis	2	terrestrial	animal parasites	whipworm	present	yes	
Classe III								
	Acanthocheilonema viteae	2	terrestrial	animal parasites	rodent filarial worm	present	yes	
	Anisakis simplex	2	terrestrial	animal parasites	infect fish and marine mammals	present	yes	
	Ascaris lumbricoides	2	terrestrial	animal parasites	causes ascariasis	present	yes	
	Ascaris suum	2	terrestrial	animal parasites	causes ascariasis	present	yes	
	Brugia malayi	2	terrestrial	animal parasites	causes elephantiasis	present	yes	
	Brugia pahangi	2	terrestrial	animal parasites	nematode that infects domestic cats	present	yes	
	Brugia timori	3	terrestrial	animal parasites	human filarial parasitic nematode	present	yes	
	Dracunculus medinensis	2	terrestrial	animal parasites	causes dracunculiasis	present	yes	
	Elaeophora elaphi	2	terrestrial	animal parasites	parasitic roundworm	present	yes	
	Gongylonema pulchrum	3	terrestrial	animal parasites	human parasitic roundworms	present	yes	
	Loa loa	2	terrestrial	animal parasites	parasitic roundworm	present	yes	
	Onchocerca flexuosa	2	terrestrial	animal parasites	parasitic roundworm	present	no	
	Onchocerca ochengi	2	terrestrial	animal parasites	parasitic roundworm	present	yes	
	Onchocerca volvulus	2	terrestrial	animal parasites	cause of river blindness in humans	present	yes	
	Ostertagia ostertagi	1	terrestrial	animal parasites	brown stomach worm	present	yes	
	Parascaris equorum	0	terrestrial	animal parasites	equine roundworm	absent	absent	17
	Thelazia callipaeda	2	terrestrial	animal parasites	cause thelaziasis	present	yes	
	Toxocara canis	2	terrestrial	animal parasites	dog roundworms	present	yes	
	Wuchereria bancrofti	2	terrestrial	animal parasites	human parasitic roundworms	present	yes	
Classe IV								
	Bursaphelenchus xylophilus	5	terrestrial	plant parasites	pine wood nematode	present	yes	
	Globodera pallida	0	terrestrial	plant parasites	plant pathogen (especially potatoes)	present	absent	
	Meloidogyne arenaria	3	terrestrial	plant parasites	peanut root knot nematode	present	yes	
	Meloidogyne floridensis	2	terrestrial	plant parasites	root-knot nematodes	present	yes	
	Meloidogyne hapla	3	terrestrial	plant parasites	root-knot nematodes	present	yes	
	Meloidogyne incognita	3	terrestrial	plant parasites	southern root-knot nematodes	present	yes	

Partie V : Annexes

	Meloidogyne javanica	4	terrestrial	plant parasites	tropical root-knot nematodes	present	yes	
	Panagrellus redivivus	3	terrestrial	heterorophic species	beer mat nematode	present	yes	
	Strongyloides papillosus	1	terrestrial	animal parasites	human pathogenic parasitic roundworm	present	yes	
	Strongyloides ratti	1	terrestrial	animal parasites	causes strongyloidiasis	present	yes	
	Strongyloides venezuelensis	1	terrestrial	animal parasites	threadworm, gastro-intestinal parasite in mouse	present	yes	
Classe V								
	Ancylostoma caninum	2	terrestrial	animal parasites	infects the small intestine of dogs	present	yes	
	Ancylostoma ceylanicum	3	terrestrial	animal parasites	zoonotic hookworm	present	yes	
	Ancylostoma duodenale	3	terrestrial	animal parasites	Old World hookworm	present	yes	
	Angiostrongylus cantonensis	2	terrestrial	animal parasites	roundworms that causes angiostrongyliasis	present	yes	
	Caenorhabditis angaria	3	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis breunneri	3	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis briggsae	4	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis elegans	4	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis japonica	3	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis remanei	3	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis sinica	3	terrestrial	heterorophic species	roundworms	present	yes	
	Caenorhabditis tropicalis	2	terrestrial	heterorophic species	roundworms	present	yes	
	Cylicostephanus goldi	2	terrestrial	animal parasites	animal parasites nematodes	present	no	
	Dictyocaulus viviparus	2	terrestrial	animal parasites	parasites of horses	present	yes	
	Dirofilaria immitis	2	terrestrial	animal parasites	heartworm or dog heartworm	present	yes	
	Haemonchus contortus	1	terrestrial	animal parasites	pathogenic nematodes of ruminants	present	yes	
	Haemonchus placei	2	terrestrial	animal parasites	parasitic roundworms	present	yes	
	Heligmosomoides polygyrus	3	terrestrial	animal parasites	intestinal roundworm of rodents	present	yes	
	Necator americanus	2	terrestrial	animal parasites	New World hookworm	present	yes	
	Nippostrongylus brasiliensis	1	terrestrial	animal parasites	gastrointestinal roundworm	present	yes	
	Oesophagostomum dentatum	2	terrestrial	animal parasites	human parasitic nematodes	present	no	

Partie V : Annexes

	<i>Parastrongyloides trichosuri</i>	0	terrestrial	animal parasites	an intestinal parasite of Australian Brush-tailed possums	present	absent	
	<i>Pristionchus exspectatus</i>	2	terrestrial	animal parasites	roundworms	present	yes	
	<i>Pristionchus pacificus</i>	2	terrestrial	animal parasites	roundworms	present	yes	
	<i>Steinernema carpocapsae</i>	4	terrestrial	animal parasites	entomopathogenic nematodes	present	yes	
	<i>Steinernema feltiae</i>	2	terrestrial	animal parasites	entomopathogenic nematodes	present	yes	
	<i>Steinernema glaseri</i>	1	terrestrial	animal parasites	entomopathogenic nematodes	present	yes	
	<i>Steinernema monticolum</i>	2	terrestrial	animal parasites	insect-parasitic nematode	present	yes	
	<i>Steinernema scapterisci</i>	2	terrestrial	animal parasites	mole cricket nematode	present	yes	
	<i>Strongylus vulgaris</i>	1	terrestrial	animal parasites	blood worm, common horse parasite	present	no	
	<i>Teladorsagia circumcincta</i>	2	terrestrial	animal parasites	nematodes that infect sheep and goats	present	yes	
Classe Other chromadoera								
	<i>Enterobius vermicularis</i>	1	terrestrial	animal parasites	Pinworm (enterobiasis)	present	yes	
	<i>Romanomermis culicivorax</i>	2	terrestrial	animal parasites	endoparasites in arthropods	present	yes	
	<i>Syphacia muris</i>	1	terrestrial	animal parasites	rat pinworm	present	yes	
	<i>Trichinella nativa</i>	1	terrestrial	animal parasites	nematode	absent	yes	16
	<i>Trichinella spiralis</i>	1	terrestrial	animal parasites	nematode	present	yes	
Opisthokonts --> Metazoa --> Protostomia --> Arthropoda								
Crustacea								
	<i>Daphnia magna</i>	2	fresh water	heterorophic species	small planctonic crustacean	absent	yes	15380
	<i>Daphnia pulex</i>	2	fresh water	heterorophic species	small planctonic crustacean	present	yes	
	<i>Eriocheir sinensis</i>	1	marine	heterorophic species	Chinese mitten crab	present	yes	
	<i>Litopenaeus vannamei</i>	1	marine	heterorophic species	Whiteleg shrimp	present	yes	
	<i>Penaeus monodon</i>	1	marine	heterorophic species	Asian tiger shrimp	present	yes	
	<i>Pseudodiaptomus poplesia</i>	1	marine	heterorophic species	a calanoid copepod	absent	yes	0
	<i>Scylla paramamosain</i>	2	marine	heterorophic species	mud and mangrove crab	present	yes	
Chelicerata								
	<i>Amblyomma americanum</i>	2	terrestrial	animal parasites	Lone Star tick	present	yes	
	<i>Dermacentor variabilis</i>	1	terrestrial	animal parasites	American dog ticks	present	yes	
	<i>Haemaphysalis longicornis</i>	1	terrestrial	animal parasites	ticks	present	yes	
	<i>Ixodes scapularis</i>	1	terrestrial	animal parasites	deer tick or black-legged tick	present	yes	

Partie V : Annexes

	Limulus polyphemus	1	marine	heterorophic species	Atlantic horseshoe crab	present	yes	
	Onychiurus arcticus	1	terrestrial	heterorophic species	Arctic springtail	present	yes	
	Rhipicephalus microplus	1	terrestrial	animal parasites	Asian blue tick	present	no	
	Sarcoptes scabiei	2	terrestrial	animal parasites	itch mites	present	yes	
	Stegodyphus mimosarum	1	terrestrial	heterorophic species	spiders	present	yes	
	Tetranychus urticae	1	terrestrial	plant parasites	red spider mite	present	yes	
Myriapoda								
	Strigamia maritima	2	marine	heterorophic species	European centipede	present	yes	
Insecta --> Hymenoptera								
	Acromyrmex echinatior	0	terrestrial	heterorophic species	New World ants	present	absent	
	Apis florea	0	terrestrial	heterorophic species	dwarf honey bee	present	absent	
	Apis mellifera	0	terrestrial	heterorophic species	western honey bee	present	absent	
	Atta cephalotes	0	terrestrial	heterorophic species	leafcutter ant	present	absent	
	Bombus impatiens	0	terrestrial	heterorophic species	common eastern bumble bee	present	absent	
	Bombus terrestris	0	terrestrial	heterorophic species	buff-tailed bumblebee	present	absent	
	Camponotus floridanus	0	terrestrial	heterorophic species	Florida capenter ant	present	absent	
	Harpegnathos saltator	0	terrestrial	heterorophic species	Indian jumping ant	present	absent	
	Linepithema humile	0	terrestrial	heterorophic species	Argentine ant	present	absent	
	Megachile rotundata	0	terrestrial	heterorophic species	alfalfa leafcutter bee	present	absent	
	Nasonia giraulti	0	terrestrial	heterorophic species	parasitoid	present	absent	
	Nasonia longicornis	0	terrestrial	heterorophic species	parasitoid	present	absent	
	Nasonia vitripennis	0	terrestrial	heterorophic species	parasitoid	present	absent	
	Pogonomyrmex barbatus	0	terrestrial	heterorophic species	red harvester ant	present	absent	
	Solenopsis invicta	0	terrestrial	heterorophic species	red imported fire ant	present	absent	
	Wasmannia auropunctata	0	terrestrial	heterorophic species	little fire ant	present	absent	
Diptera								
	Aedes aegypti	0	terrestrial	heterorophic species	yellow fever mosquito	present	absent	
	Aedes albopictus	0	terrestrial	heterorophic species	Asian tiger mosquito	present	absent	
	Anopheles albimanus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles aquasalis	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles arabiensis	0	terrestrial	heterorophic species	mosquito	present	absent	

Partie V : Annexes

	Anopheles atroparvus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles christyi	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles coluzzii	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles culicifacies	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles darlingi	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles dirus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles epiroticus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles farauti	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles funestus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles gambiae	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles maculatus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles melas	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles merus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles minimus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles quadriannulatus	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles sinensis	0	terrestrial	heterorophic species	mosquito	present	absent	
	Anopheles stephensi	0	terrestrial	heterorophic species	mosquito	present	absent	
	Clogmia albupunctata	0	terrestrial	heterorophic species	drain flies	present	absent	
	Culex quinquefasciatus	0	terrestrial	heterorophic species	southern house mosquito	present	absent	
	Drosophila albomicans	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila ananassae	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila biarmipes	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila bipectinata	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila elegans	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila erecta	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila eugracilis	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila ficusphila	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila grimshawi	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila kikkawai	0	terrestrial	heterorophic species	fly	present	absent	
	Drosophila melanogaster	0	terrestrial	heterorophic species	fly	present	absent	

Partie V : Annexes

	<i>Drosophila miranda</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila mojavensis</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila neotestacea</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila persimilis</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila pseudoobscura</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila rhopaloa</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila santomea</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila sechellia</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila simulans</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila takahashii</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila virilis</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila willistoni</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Drosophila yakuba</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Episyrphus balteatus</i>	0	terrestrial	heterorophic species	marmalade hoverfly	present	absent	
	<i>Glossina austeni</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Glossina brevipalpis</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Glossina fuscipes</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Glossina morsitans morsitans</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Glossina pallidipes</i>	0	terrestrial	heterorophic species	Tsetse fly	present	absent	
	<i>Lutzomyia longipalpis</i>	0	terrestrial	heterorophic species	blood sucker	present	absent	
	<i>Mayetiola destructor</i>	0	terrestrial	heterorophic species	Hessian fly or barley midge	present	absent	
	<i>Megaselia scalaris</i>	0	terrestrial	heterorophic species	fly	present	absent	
	<i>Musca domestica</i>	0	terrestrial	heterorophic species	housefly	present	absent	
	<i>Phlebotomus papatasi</i>	0	terrestrial	heterorophic species	sand flies	present	absent	
	<i>Stomoxys calcitrans</i>	0	terrestrial	heterorophic species	stable fly, dog fly	present	absent	
	<i>Triatoma infestans</i>	0	terrestrial	heterorophic species	winchuka, "kissing bug", blood sucker	absent	absent	3488
Hemiptera								
	<i>Acyrtosiphon kondoi</i>	2	terrestrial	plant parasites	aphid	absent	yes	23.041
	<i>Acyrtosiphon pisum</i>	4	terrestrial	plant parasites	pea aphid	present	yes	
	<i>Aphis gossypii</i>	2	terrestrial	plant parasites	melon and cotton aphid	present	yes	

Partie V : Annexes

	<i>Cimex lectularius</i>	0	terrestrial	free-living species	Bed bugs - human primary hosts	present	absent	
	<i>Diaphorina citri</i>	1	terrestrial	plant parasites	Asian citrus psyllid	present	yes	
	<i>Maconellicoccus hirsutus</i>	1	terrestrial	plant parasites	pink, grape or hibiscus mealybug	absent	yes	7.696
	<i>Myzus ascalonicus</i>	3	terrestrial	plant parasites	Shallot aphid	absent	yes	22.942
	<i>Myzus persicae</i>	3	terrestrial	plant parasites	Green peach aphid	present	yes	
	<i>Nilaparvata lugens</i>	1	terrestrial	plant parasites	brown planthopper	absent	yes	118.141
	<i>Peregrinus maidis</i>	1	terrestrial	plant parasites	Corn planthopper	absent	yes	20.678
	<i>Rhodnius prolixus</i>	0	terrestrial	free-living species	principal triatomine vector of the Chagas (trypanosoma cruzi) parasite	present	absent	
	<i>Rhopalosiphum maidis</i>	3	terrestrial	plant parasites	Corn leaf aphid	present	yes	
	<i>Rhopalosiphum padi</i>	3	terrestrial	plant parasites	Bird cherry-oat aphid	present	yes	
	<i>Sitobion avenae</i>	1	terrestrial	plant parasites	English grain aphid	absent	yes	3.614
Lepidoptera								
	<i>Agrotis segetum</i>	1	terrestrial	heterorophic species	turnip moth	present	yes	
	<i>Amyelois transitella</i>	3	terrestrial	heterorophic species	navel orangeworm	present	yes	
	<i>Antheraea assama</i>	1	terrestrial	heterorophic species	muga silkworm	absent	yes	35,191
	<i>Antheraea mylitta</i>	1	terrestrial	heterorophic species	tasar silkworm	absent	yes	3.913
	<i>Antheraea yamamai</i>	1	terrestrial	heterorophic species	Japanese silk moth	absent	yes	662
	<i>Bicyclus anynana</i>	1	terrestrial	heterorophic species	squinting bush brown	absent	yes	101.595
	<i>Bombyx mori</i>	2	terrestrial	heterorophic species	domestic silkworm	present	yes	
	<i>Danaus plexippus</i>	2	terrestrial	heterorophic species	monarch butterfly	present	yes	
	<i>Heliconius melpomene</i>	3	terrestrial	heterorophic species	postman butterfly	present	yes	
	<i>Helicoverpa armigera</i>	1	terrestrial	heterorophic species	cotton bollworm, corn earworm	present	yes	
	<i>Heliothis virescens</i>	1	terrestrial	heterorophic species	tobacco budworm	present	yes	
	<i>Manduca sexta</i>	1	terrestrial	heterorophic species	Carolina sphinx moth	present	yes	
	<i>Melitaea cinxia</i>	2	terrestrial	heterorophic species	Glanville fritillary butterfly	present	yes	
	<i>Ostrinia nubilalis</i>	1	terrestrial	heterorophic species	European corn borer	absent	yes	15.040
	<i>Papilio machaon</i>	2	terrestrial	heterorophic species	Old World swallowtail butterfly	present	yes	
	<i>Papilio polytes</i>	2	terrestrial	heterorophic species	common mormon butterfly	present	yes	

Partie V : Annexes

	Papilio xuthus	2	terrestrial	heterorophic species	Asian swallowtail butterfly	present	yes	
	Plutella xylostella	1	terrestrial	heterorophic species	diamondback moth, cabbage moth	present	yes	
	Spodoptera exigua	1	terrestrial	heterorophic species	armyworms pests	absent	yes	16359
	Spodoptera frugiperda	2	terrestrial	heterorophic species	armyworms pests	present	yes	
	Spodoptera littoralis	1	terrestrial	heterorophic species	armyworms pests	absent	yes	72.261
	Trichoplusia ni	1	terrestrial	heterorophic species	cabbage looper, pests	present	yes	
Phthiraptera								
	Pediculus humanus	0	terrestrial	heterorophic species	head louse, ectoparasite of humans	present	yes	
Coleoptera								
	Dendroctonus ponderosae	2	terrestrial	heterorophic species	mountain pine beetle	present	yes	
	Pissodes strobi	1	terrestrial	plant parasites	white pine weevil	absent	yes	69.475
	Sitophilus oryzae	1	terrestrial	plant parasites	rice weevil	absent	yes	25.748
	Sphenophorus levis	1	terrestrial	plant parasites	weevil	absent	yes	3.447
	Tribolium castaneum	1	terrestrial	heterorophic species	red flour beetle	present	yes	
Orthoptera								
	Gryllus bimaculatus	1	terrestrial	heterorophic species	cricket	present	yes	
	Schistocerca gregaria	1	terrestrial	heterorophic species	locust	absent	yes	34.709
Blattoptera								
	Coptotermes formosanus	1	terrestrial	heterorophic species	formosan subterranean termite	absent	yes	142.814
	Periplaneta americana	1	terrestrial	heterorophic species	america cockroach	absent	yes	5.394
	Reticulitermes flavipes	1	terrestrial	heterorophic species	eastern subterranean termite	absent	yes	23.767
	Zootermopsis nevadensis	1	terrestrial	heterorophic species	eusocial dampwood termite	present	yes	
Archaeplastida --> Glaucophytes								
	Cyanophora paradoxa	3	fresh water	autotrophic species	algae, model organism	present	no	
Rhodophytes								
	Chondrus crispus	0	marine	autotrophic species	red algae called Irish moss or carageen moss	present	absent	
	Cyanidioschyzon merolae	1	marine	autotrophic species	red algae	present	yes	
	Galdieria sulphuraria	1	marine	autotrophic species	extermophilic red algae	present	yes	
	Grateloupia livida	1	marine	autotrophic species	red algae	present	yes	
	Porphyridium purpureum	1	marine	autotrophic species	red algae	present	yes	
Green algae - Chlorophyta								

Partie V : Annexes

	<i>Alnus glutinosa</i>	1	terrestrial	autotrophic species	black alder	absent	yes	32620
	<i>Chlamydomonas reinhardtii</i>	1	marine	autotrophic species	green algae	present	yes	
	<i>Chlorella variabilis</i>	1	marine	autotrophic species	green algae	present	yes	
	<i>Gonium pectorale</i>	1	marine	autotrophic species	colonial algae	absent	yes	0
	<i>Klebsormidium flaccidum</i>	1	marine	autotrophic species	filamentous charophyte green algae	present	yes	
	<i>Monoraphidium neglectum</i>	1	marine	autotrophic species	green algae	present	yes	
	<i>Volvox carteri</i>	1	marine	autotrophic species	colonial green algae	present	yes	
Moss - Bryophyta								
	<i>Physcomitrella patens</i>	1	terrestrial	autotrophic species	moss	present	yes	
Ferns								
	<i>Acrostichum aureum</i>	1	terrestrial	autotrophic species	golden leather fern	absent	yes	0
	<i>Acrostichum speciosum</i>	1	terrestrial	autotrophic species	Mangrove fern	absent	yes	0
	<i>Adiantum capillus</i>	1	terrestrial	autotrophic species	black maidenhair fern	absent	yes	30544
	<i>Azolla filiculoides</i>	1	fresh water	autotrophic species	floating aquatic fern	present	yes	
	<i>Ceratopteris richardii</i>	1	terrestrial	autotrophic species	fern	absent	yes	5133
	<i>Ceratopteris thalictroides</i>	1	terrestrial	autotrophic species	water sprite, Indian fern	absent	yes	0
	<i>Lygodium japonicum</i>	1	terrestrial	autotrophic species	vine-like fern or Japanese climbing fern	absent	yes	0
	<i>Pteridium aquilinum</i>	1	terrestrial	autotrophic species	eagle fern	absent	yes	424
Ferns - relative								
	<i>Selaginella moellendorffii</i>	4	terrestrial	autotrophic species	model organism	present	yes	
Gymnospermes								
	<i>Picea sitchensis</i>	2	terrestrial	autotrophic species	sitka spruce	present	yes	
Marchantiales								
	<i>Marchantia polymorpha</i>	1	terrestrial	autotrophic species	umbrella liverwort	present	yes	
Angiospermes - Dicotylédons								
	<i>Alnus glutinosa</i>	1	terrestrial	autotrophic species	black alder		yes	
	<i>Amborella trichopoda</i>	2	terrestrial	autotrophic species	understory shrubs		yes	
	<i>Arabidopsis lyrata</i>	3	terrestrial	autotrophic species	thale cress, mouse-ear cress	present	yes	
	<i>Arabidopsis thaliana</i>	3	terrestrial	autotrophic species	thale cress, mouse-ear cress	present	yes	
	<i>Artemisia annua</i>	1	terrestrial	autotrophic species	sweet wormwood		yes	
	<i>Beta vulgaris</i>	4	terrestrial	autotrophic species	sugar beet, spinach beet	present	yes	

Partie V : Annexes

	<i>Botryococcus braunii</i>	1	marine	autotrophic species	green, pyramid-shaped planktonic microalgua	present	no	
	<i>Brachypodium distachyon</i>	3	terrestrial	autotrophic species	purple false brome	present	yes	
	<i>Brassica napus</i>	3	terrestrial	autotrophic species	oilseed rape	present	yes	
	<i>Brassica oleracea</i>	4	terrestrial	autotrophic species	wild cabbage plants	present	yes	
	<i>Camelina sativa</i>	3	terrestrial	autotrophic species	gold-of-pleasure or false flax	present	yes	
	<i>Camellia sinensis</i>	1	terrestrial	autotrophic species	tea plant	present	no	
	<i>Catharanthus roseus</i>	1	terrestrial	autotrophic species	Madagascar periwinkle	present	no	
	<i>Cicer arietinum</i>	3	terrestrial	autotrophic species	chickpea	present	yes	
	<i>Citrus clementina</i>	3	terrestrial	autotrophic species	clementine	present	yes	
	<i>Citrus reticulata</i>	1	terrestrial	autotrophic species	mandarin orange	present	no	
	<i>Coffea canephora</i>	3	terrestrial	autotrophic species	robusta coffee	present	yes	
	<i>Cucumis melo</i>	3	terrestrial	autotrophic species	muskmelon	present	yes	
	<i>Dendrobium nobile</i>	1	terrestrial	autotrophic species	member of Orchidaceae	present	no	
	<i>Dioscorea alata</i>	1	terrestrial	autotrophic species	purple yam	present	no	
	<i>Erythranthe guttata</i>	3	terrestrial	autotrophic species	seep monkey flower	present	yes	
	<i>Eucalyptus grandis</i>	3	terrestrial	autotrophic species	flooded gum or rose gum	present	yes	
	<i>Euonymus alatus</i>	1	terrestrial	autotrophic species	winged spindle	present	no	
	<i>Eutrema salsugineum</i>	3	terrestrial	autotrophic species	rockcress	present	yes	
	<i>Fragaria vesca</i>	3	terrestrial	autotrophic species	strawberry	present	yes	
	<i>Glycine max</i>	3	terrestrial	autotrophic species	soybean	present	yes	
	<i>Gossypium raimondii</i>	3	terrestrial	autotrophic species	cotton plant	present	yes	
	<i>Jatropha curcas</i>	3	terrestrial	autotrophic species	physic nut, Barbados nut, poison nut, bubble bush or purging nut	present	yes	
	<i>Lotus japonicus</i>	3	terrestrial	autotrophic species	wild legume	present	yes	
	<i>Malus domestica</i>	3	terrestrial	autotrophic species	apple	present	yes	
	<i>Medicago truncatula</i>	4	terrestrial	autotrophic species	barrelclover	present	yes	
	<i>Nelumbo nucifera</i>	3	terrestrial	autotrophic species	Indian lotus	present	yes	
	<i>Nicotiana sylvestris</i>	4	terrestrial	autotrophic species	common domesticated tobacco	present	yes	

Partie V : Annexes

	Nicotiana tomentosiformis	2	terrestrial	autotrophic species	tobacco	present	no	
	Panax ginseng	1	terrestrial	autotrophic species	ginseng	present	no	
	Panicum virgatum	1	terrestrial	autotrophic species	switchgrass	present	no	
	Phaseolus vulgaris	3	terrestrial	autotrophic species	common bean	present	yes	
	Populus euphratica	4	terrestrial	autotrophic species	Euphrates poplar or desert poplar	present	yes	
	Populus trichocarpa	2	terrestrial	autotrophic species	black cottonwood	present	yes	
	Prunus mume	3	terrestrial	autotrophic species	Chinese plum or Japanese apricot	present	yes	
	Pyrus bretschneideri	3	terrestrial	autotrophic species	Chinese white pear	present	yes	
	Rehmannia glutinosa	1	terrestrial	autotrophic species	use in Chinese traditional medicine	absent	no	1301
	Ricinus communis	3	terrestrial	autotrophic species	castorbean or castor-oil-plant	present	yes	
	Sesamum indicum	3	terrestrial	autotrophic species	sesame	present	yes	
	Solanum lycopersicum	4	terrestrial	autotrophic species	tomato	present	yes	
	Solanum tuberosum	3	terrestrial	autotrophic species	potato	present	yes	
	Tarenaya hassleriana	4	terrestrial	autotrophic species	cleome hassleriana, spider flower or pink queen	present	yes	
	Theobroma cacao	3	terrestrial	autotrophic species	cacao or cocoa tree	present	yes	
	Triticum aestivum	3	terrestrial	autotrophic species	bread wheat	present	yes	
	Triticum urartu	1	terrestrial	autotrophic species	red wild einkorn	present	yes	
	Vigna radiata	4	terrestrial	autotrophic species	mung bean	present	yes	
	Vitis vinifera	4	terrestrial	autotrophic species	grape vine	present	yes	
	Wolffia arrhiza	1	terrestrial	autotrophic species	spotless watermeal and rootless duckweed	present	no	
Angiospermes - Monocotylédons Liliopsida								
	Brachypodium distachyon	3	terrestrial	autotrophic species	purple false brome or stiff brome	present	yes	
	Elaeis guineensis	2	terrestrial	autotrophic species	oil palm or macaw-fat	present	yes	
	Hordeum vulgare	2	terrestrial	autotrophic species	barley	present	yes	
	Musa acuminata	3	terrestrial	autotrophic species	banana species	present	yes	
	Oryza brachyantha	3	terrestrial	autotrophic species	grass in the rice genus	present	yes	
	Phoenix dactylifera	3	terrestrial	autotrophic species	date palm	present	yes	
	Setaria italica	3	terrestrial	autotrophic species	poales, foxtail millet	present	yes	

	Zea mays	3	terrestrial	autotrophic species	maize	present	yes	
	Zostera marina	2	marine	autotrophic species	common eelgrass, seawrack	present	yes	

Tableau 8 – Listes des espèces eucaryotes utilisées pour l'étude

Ce tableau comprend le nom des espèces, un bref descriptif de celle – ci, leurs taxons, leur modes de vie, leur environnement et le nombre de protéines MIFs trouvés par espèces. La présence d'un génome pour l'espèce donné et à défaut d'un génome le nombre d'EST présent sur la base de données NCBI est indiqué. Ces espèces pour lesquels les séquences MIF ont été sélectionnés sont indiqué.

Bibliographie

1. Akira S, Uematsu S, Takeuchi O. Pathogen Recognition and Innate Immunity. *Cell*. 2006;124:783–801.
2. Medzhitov R, Janeway CA. Decoding the Patterns of Self and Nonself by the Innate Immune System. *Science*. 2002;296:298–300.
3. Medzhitov R, Janeway CJ. Innate immune recognition: mechanisms and pathways. *Immunol Rev*. 2000;173:89–97.
4. Gust AA, Pruitt R, Nürnberger T. Sensing Danger: Key to Activating Plant Immunity. *Trends Plant Sci*. 2017;22:779–91.
5. Nie L, Cai S-Y, Shao J-Z, Chen J. Toll-Like Receptors, Associated Biological Roles, and Signaling Networks in Non-Mammals. *Front Immunol*. 2018;9:153-160
6. Kalantari P, Kalantari P. The Emerging Role of Pattern Recognition Receptors in the Pathogenesis of Malaria. *Vaccines*. 2018;6:133-141.
7. Jr CAJ, Travers P, Walport M, Shlomchik MJ, Jr CAJ, Travers P, et al. *Immunobiology*. 5th edition. Garland Science; 2001.
8. Bloom BR, Bennett B. Mechanism of a Reaction in Vitro Associated with Delayed-Type Hypersensitivity. *Science*. 1966;153:80–92.
9. David JR. Delayed hypersensitivity in vitro: its mediation by cell-free substances formed by lymphoid cell-antigen interaction. *Proc Natl Acad Sci U S A*. 1966;56:72–77.
10. Pastrana DV, Raghavan N, FitzGerald P, Eisinger SW, Metz C, Bucala R, et al. Filarial Nematode Parasites Secrete a Homologue of the Human Cytokine Macrophage Migration Inhibitory Factor. *Infect Immun*. 1998;66:59–63.
11. Sharma R, Hoti SL, Meena RL, Vasuki V, Sankari T, Kaliraj P. Molecular and functional characterization of macrophage migration inhibitory factor (MIF) homolog of human from lymphatic filarial parasite *Wuchereria bancrofti*. *Parasitol Res*. 2012;111:35–47.
12. Wasala NB, Jaworski DC. *Dermacentor variabilis*: Characterization and modeling of macrophage migration inhibitory factor with phylogenetic comparisons to other ticks, insects and parasitic nematodes. *Exp Parasitol*. 2012;130:32–38.
13. Wasala NB, Bowen CJ, Jaworski DC. Expression and regulation of macrophage migration inhibitory factor (MIF) in feeding American dog ticks, *Dermacentor variabilis*. *Exp Appl Acarol*. 2012;57:179–187.
14. Cho Y, Jones BF, Vermeire JJ, Leng L, DiFedele L, Harrison LM, et al. Structural and Functional Characterization of a Secreted Hookworm Macrophage Migration Inhibitory Factor (MIF) That Interacts with the Human MIF Receptor CD74. *J Biol Chem*. 2007;282:447–456.

15. Augustijn KD, Kleemann R, Thompson J, Kooistra T, Crawford CE, Reece SE, et al. Functional Characterization of the *Plasmodium falciparum* and *P. berghei* Homologues of Macrophage Migration Inhibitory Factor. *Infect Immun*. 2007;75:11-16.
16. Baeza Garcia A, Pierce RJ, Gourbal B, Werkmeister E, Colinet D, Reichhart J-M, et al. Involvement of the Cytokine MIF in the Snail Host Immune Response to the Parasite *Schistosoma mansoni*. *PLoS Pathog*. 2010;6:57-64.
17. Furukawa R, Tamaki K, Kaneko H. Two macrophage migration inhibitory factors regulate starfish larval immune cell chemotaxis. *Immunol Cell Biol*. 94:315–321.
18. Huang S, Cao Y, Lu M, Peng W, Lin J, Tang C, et al. Identification and functional characterization of *Oncomelania hupensis* macrophage migration inhibitory factor involved in the snail host innate immune response to the parasite *Schistosoma japonicum*. *Int J Parasitol*. 2017;47:485–499.
19. Huang W-S, Duan L-P, Huang B, Wang K-J, Zhang C-L, Jia Q, et al. Macrophage migration inhibitory factor (MIF) family in arthropods: Cloning and expression analysis of two MIF and one D-dopachrome tautomerase (DDT) homologues in mud crabs, *Scylla paramamosain*. *Fish Shellfish Immunol*. 2016;50:142–159.
20. Elodie Naessens GD, Philippe Giordanengo OB, Naïma Minet-Kebdani HK, Christine Coustau. A secreted MIF cytokine enables aphid feeding and represses plant immune responses. *Curr-Biol*. 2015;56-63.
21. Panstruga R, Baumgarten K, Bernhagen J. Phylogeny and evolution of plant macrophage migration inhibitory factor/D-dopachrome tautomerase-like proteins. *BMC Evol Biol*. 2015;15:154-161.
22. George M, Vaughan JH. In vitro cell migration as a model for delayed hypersensitivity. *Proc Soc Exp Biol Med Soc Exp Biol Med N Y N*. 1962;111:514–521.
23. Weiser WY, Temple PA, Witek-Giannotti JS, Remold HG, Clark SC, David JR. Molecular cloning of a cDNA encoding a human macrophage migration inhibitory factor. *Proc Natl Acad Sci*. 1989;86:75–86.
24. Calandra T, Roger T. Macrophage migration inhibitory factor: a regulator of innate immunity. *Nat Rev Immunol*. 2003;3:791–800.
25. Bucala R. MIF : A Most Interesting Factor. Barnes & Noble®. World Scientific Publishing Company; 2007.
26. Sugimoto H, Taniguchi M, Nakagawa A, Tanaka I, Suzuki M, Nishihira J. Crystal Structure of Human d-Dopachrome Tautomerase, a Homologue of Macrophage Migration Inhibitory Factor, at 1.54 Å Resolution,. *Biochemistry (Mosc)*. 1999;38:68–79.
27. Odh G, Hindemith A, Rosengren AM, Rosengren E, Rorsman H. Isolation of a New Tautomerase Monitored by the Conversion of D-Dopachrome to 5,6-Dihydroxyindole. *Biochem Biophys Res Commun*. 1993;197:619–624.

28. Bernhagen J, Mitchell RA, Calandra T, Voelter W, Cerami A, Bucala R. Purification, bioactivity, and secondary structure analysis of mouse and human macrophage migration inhibitory factor (MIF). *Biochemistry (Mosc)*. 1994;33:144–155.
29. Kozak CA, Adamson MC, Buckler CE, Segovia L, Paralkar V, Wistow G. Genomic Cloning of Mouse MIF (Macrophage Inhibitory Factor) and Genetic Mapping of the Human and Mouse Expressed Gene and Nine Mouse Pseudogenes. *Genomics*. 1995;27:405–411.
30. Bozza M, Kolakowski LF, Jenkins NA, Gilbert DJ, Copeland NG, David JR, et al. Structural Characterization and Chromosomal Location of the Mouse Macrophage Migration Inhibitory Factor Gene and Pseudogenes. *Genomics*. 1995;27:412–419.
31. Sun HW, Bernhagen J, Bucala R, Lolis E. Crystal structure at 2.6-Å resolution of human macrophage migration inhibitory factor. *Proc Natl Acad Sci U S A*. 1996;93:91–105.
32. Sun H-W, Swope M, Cinquina C, Bedarkar S, Bernhagen J, Bucala R, et al. The subunit structure of human macrophage migration inhibitory factor : evidence for a trimer. *Protein Eng*. 1996;9:631–645.
33. Dobson SE, Augustijn KD, Brannigan JA, Schnick C, Janse CJ, Dodson EJ, et al. The crystal structures of macrophage migration inhibitory factor from *Plasmodium falciparum* and *Plasmodium berghei*. *Protein Sci Publ Protein Soc*. 2009;18:78–91.
34. Buchko GW, Abendroth J, Robinson H, Zhang Y, Hewitt SN, Edwards TE, et al. sequence of GI-MIF and the other Crystal structure of a macrophage migration inhibitory factor from *Giardia lamblia*. *J Struct Funct Genomics*. 2013;14:155–159.
35. Tan TH, Edgerton SA, Kumari R, McAlister MS, Roe SM, Nagl S, et al. Macrophage migration inhibitory factor of the parasitic nematode *Trichinella spiralis*. *Biochem J*. 2001;14:121–129.
36. Zang X, Taylor P, Wang JM, Meyer DJ, Scott AL, Walkinshaw MD, et al. Homologues of Human Macrophage Migration Inhibitory Factor from a Parasitic Nematode gene cloning, protein activity, and crystal structure. *J Biol Chem*. 2002;77:261–277.
37. Wasiel AA, Rozeboom HJ, Hauke D, Baas B-J, Zandvoort E, Quax WJ, et al. Structural and Functional Characterization of a Macrophage Migration Inhibitory Factor Homologue from the Marine Cyanobacterium *Prochlorococcus marinus*,. *Biochemistry (Mosc)*. 2010;49:72–81.
38. Merk M, Mitchell RA, Endres S, Bucala R. D-dopachrome tautomerase (D-DT or MIF-2): Doubling the MIF cytokine family. *Cytokine*. 2012;59:10–27.
39. Flieger O, Engling A, Bucala R, Lue H, Nickel W, Bernhagen J. Regulated secretion of macrophage migration inhibitory factor is mediated by a non-classical pathway involving an ABC transporter. *FEBS Lett*. 551:78–86.
40. Shao D, Zhong X, Zhou Y-F, Han Z, Lin Y, Wang Z, et al. Structural and functional comparison of MIF ortholog from *Plasmodium yoelii* with MIF from its rodent host. *Mol Immunol*. 2010;47:726–737.

41. Subbannayya T, Variar P, Advani J, Nair B, Shankar S, Gowda H, et al. An integrated signal transduction network of macrophage migration inhibitory factor. *J Cell Commun Signal*. 2016;10:165–170.
42. Leng L, Metz CN, Fang Y, Xu J, Donnelly S, Baugh J, et al. MIF Signal Transduction Initiated by Binding to CD74. *J Exp Med*. 2003;197:67–76.
43. Schröder B. The multifaceted roles of the invariant chain CD74 — More than just a chaperone. *Biochim Biophys Acta BBA - Mol Cell Res*. 2016;63:69–81.
44. Xu L, Li Y, Li D, Xu P, Tian S, Sun H, et al. Exploring the binding mechanisms of MIF to CXCR2 using theoretical approaches. *Phys Chem Chem Phys*. 2015;17:70–82.
45. Weber C, Kraemer S, Drechsler M, Lue H, Koenen RR, Kapurniotu A, et al. Structural determinants of MIF functions in CXCR2-mediated inflammatory and atherogenic leukocyte recruitment. *Proc Natl Acad Sci U S A*. 2008;105:78–83.
46. Rajasekaran D, Gröning S, Schmitz C, Zierow S, Drucker N, Bakou M, et al. Macrophage Migration Inhibitory Factor-CXCR4 Receptor Interactions. *J Biol Chem*. 2016;291:81–95.
47. Alampour-Rajabi S, El Bounkari O, Rot A, Müller-Newen G, Bachelier F, Gawaz M, et al. MIF interacts with CXCR7 to promote receptor internalization, ERK1/2 and ZAP-70 signaling, and lymphocyte chemotaxis. *FASEB J*. 2015;29:497–511.
48. Jean-Charles P-Y, Snyder JC, Shenoy SK. Chapter One - Ubiquitination and Deubiquitination of G Protein-Coupled Receptors. In: Shenoy SK, editor. *Progress in Molecular Biology and Translational Science*. Academic Press; 2016.6:1-10.
49. Levoe A, Balabanian K, Baleux F, Bachelier F, Lagane B. CXCR7 heterodimerizes with CXCR4 and regulates CXCL12-mediated G protein signaling. *Blood*. 2009;113:85–93.
50. Mitchell RA, Metz CN, Peng T, Bucala R. Sustained Mitogen-activated Protein Kinase (MAPK) and Cytoplasmic Phospholipase A2 Activation by Macrophage Migration Inhibitory Factor (MIF) Regulatory role in cell proliferation and glucocorticoid action. *J Biol Chem*. 1999;27:156-161.
51. Roger T, David J, Glauser MP, Calandra T. MIF regulates innate immune responses through modulation of Toll-like receptor 4. *Nature*. 2001;414:201–214.
52. Roger T, Froidevaux C, Martin C, Calandra T. Macrophage migration inhibitory factor (MIF) regulates host responses to endotoxin through modulation of Toll-like receptor 4 (TLR4). *J Endotoxin Res*. 2003;9:119–123.
53. Hudson JD, Shoaibi MA, Maestro R, Carnero A, Hannon GJ, Beach DH. A Proinflammatory Cytokine Inhibits P53 Tumor Suppressor Activity. *J Exp Med*. 1999;190:75–82.
54. Mitchell RA, Liao H, Chesney J, Fingerle-Rowson G, Baugh J, David J, et al. Macrophage migration inhibitory factor (MIF) sustains macrophage proinflammatory function by inhibiting p53: Regulatory role in the innate immune response. *Proc Natl Acad Sci U S A*. 2002;99:45–50.

55. Yabunaka N, Nishihira J, Mizue Y, Tsuji M, Kumagai M, Ohtsuka Y, et al. Elevated serum content of macrophage migration inhibitory factor in patients with type 2 diabetes. *Diabetes Care*. 2000;23:56–68.
56. Liu R-M, Sun D-N, Jiao Y-L, Wang P, Zhang J, Wang M, et al. Macrophage migration inhibitory factor promotes tumor aggressiveness of esophageal squamous cell carcinoma via activation of Akt and inactivation of GSK3 β . *Cancer Lett*. 2018;412:89–96.
57. Leech M, Metz C, Hall P, Hutchinson P, Gianis K, Smith M, et al. Macrophage migration inhibitory factor in rheumatoid arthritis: Evidence of proinflammatory function and regulation by glucocorticoids. *Arthritis Rheum*. 1999;42:160–178.
58. Griga T, Wilkens C, Wirkus N, Epplen J, Schmiegel W, Klein W. A polymorphism in the macrophage migration inhibitory factor gene is involved in the genetic predisposition of Crohn's disease and associated with cumulative steroid doses. *Hepatogastroenterology*. 2007;54:84–96.
59. Dambacher J, Staudinger T, Seiderer J, Sisic Z, Schnitzler F, Pfennig S, et al. Macrophage migration inhibitory factor (MIF) -173G/C promoter polymorphism influences upper gastrointestinal tract involvement and disease activity in patients with Crohn's disease. *Inflamm Bowel Dis*. 2007;13:71–82.
60. Hermanowski-Vosatka A, Mundt SS, Ayala JM, Goyal S, Hanlon WA, Czerwinski RM, et al. Enzymatically inactive macrophage migration inhibitory factor inhibits monocyte chemotaxis and random migration. *Biochemistry (Mosc)*. 1999;38:41–59.
61. Lue H, Kapurniotu A, Fingerle-Rowson G, Roger T, Leng L, Thiele M, et al. Rapid and transient activation of the ERK MAPK signalling pathway by macrophage migration inhibitory factor (MIF) and dependence on JAB1/CSN5 and Src kinase activity. *Cell Signal*. 2006;18:88–76.
62. Calandra T, Bucala R. Macrophage migration inhibitory factor: a counter-regulator of glucocorticoid action and critical mediator of septic shock. *J Inflamm*. 1995;47:39–51.
63. Koebernick H, Grode L, David JR, Rohde W, Rolph MS, Mittrücker H-W, et al. Macrophage migration inhibitory factor (MIF) plays a pivotal role in immunity against *Salmonella typhimurium*. *Proc Natl Acad Sci U S A*. 2002;81–96.
64. Das R, LaRose MI, Hergott CB, Leng L, Bucala R, Weiser JN. Macrophage Migration Inhibitory Factor Promotes Clearance of Pneumococcal Colonization. *J Immunol*. 2014;193:64–72.
65. Bozza M, Satoskar AR, Lin G, Lu B, Humbles AA, Gerard C, et al. Targeted Disruption of Migration Inhibitory Factor Gene Reveals Its Critical Role in Sepsis. *J Exp Med*. 1999;189:41–56.
66. Bacher M, Metz CN, Calandra T, Mayer K, Chesney J, Lohoff M, et al. An essential regulatory role for macrophage migration inhibitory factor in T-cell activation. *Proc Natl Acad Sci*. 1996;93:49–54.
67. Abe R, Peng T, Sailors J, Bucala R, Metz CN. Regulation of the CTL Response by Macrophage Migration Inhibitory Factor. *J Immunol*. 2001;166:47–53.

68. Rosengren E, Bucala R, Åman P, Jacobsson L, Odh G, Metz CN, et al. The Immunoregulatory Mediator Macrophage Migration Inhibitory Factor (MIF) Catalyzes a Tautomerization Reaction. *Mol Med*. 1996;2:143-158.
69. Lubetsky JB, Dios A, Han J, Aljabari B, Ruzsicska B, Mitchell R, et al. The Tautomerase Active Site of Macrophage Migration Inhibitory Factor Is a Potential Target for Discovery of Novel Anti-inflammatory Agents. *J Biol Chem*. 2002;277:76–82.
70. Swope M, Sun H-W, Blake PR, Lolis E. Direct link between cytokine activity and a catalytic site for macrophage migration inhibitory factor. *EMBO J*. 1998;17:34–41.
71. Kleemann R, Kapurniotu A, Frank RW, Gessner A, Mischke R, Flieger O, et al. Disulfide analysis reveals a role for macrophage migration inhibitory factor (MIF) as thiol-protein oxidoreductase. *J Mol Biol*. 1998;280:85–102.
72. Thiele M, Bernhagen J. Link Between Macrophage Migration Inhibitory Factor and Cellular Redox Regulation. *Antioxid Redox Signal*. 2005;7:34–48.
73. Wang Y, An R, Umanah GK, Park H, Nambiar K, Eacker SM, et al. A nuclease that mediates cell death induced by DNA damage and poly(ADP-ribose) polymerase-1. *Science*. 2016;354:156-161.
74. Toso C, Emamaullee JA, Merani S, Shapiro AMJ. The role of macrophage migration inhibitory factor on glucose metabolism and diabetes. *Diabetologia*. 2008;51:37–46.
75. Hirokawa J, Sakaue S, Tagami S, Kawakami Y, Sakai M, Nishi S, et al. Identification of Macrophage Migration Inhibitory Factor in Adipose Tissue and Its Induction by Tumor Necrosis Factor- α . *Biochem Biophys Res Commun*. 1997;235:94–108.
76. Hirokawa J, Sakaue S, Furuya Y, Ishii J, Hasegawa A, Tagami S, et al. Tumor Necrosis Factor- α Regulates the Gene Expression of Macrophage Migration Inhibitory Factor through Tyrosine Kinase-Dependent Pathway in 3T3-L1 Adipocytes. *J Biochem* . 1998;123:33–49.
77. Sakaue S, Nishihira J, Hirokawa J, Yoshimura H, Honda T, Aoki K, et al. Regulation of macrophage migration inhibitory factor (MIF) expression by glucose and insulin in adipocytes in vitro. *Mol Med*. 1999;5:61–71.
78. Benigni F, Atsumi T, Calandra T, Metz C, Echtenacher B, Peng T, et al. The proinflammatory mediator macrophage migration inhibitory factor induces glucose catabolism in muscle. *J Clin Invest*. 2000;106:91–100.
79. Skurk T, Herder C, Kräft I, Müller-Scholze S, Hauner H, Kolb H. Production and Release of Macrophage Migration Inhibitory Factor from Human Adipocytes. *Endocrinology*. 2005;146:1–11.
80. Atsumi T, Cho Y-R, Leng L, McDonald C, Yu T, Danton C, et al. The Proinflammatory Cytokine Macrophage Migration Inhibitory Factor Regulates Glucose Metabolism during Systemic Inflammation. *J Immunol*. 2007;179:99–106.
81. Xu X, Wang B, Ye C, Yao C, Lin Y, Huang X, et al. Overexpression of macrophage migration inhibitory factor induces angiogenesis in human breast cancer. *Cancer Lett*. 2008;261:147–157.

82. Richard V, Kindt N, Saussez S. Macrophage migration inhibitory factor involvement in breast cancer (Review). *Int J Oncol.* 2015;47:27–33.
83. Galluzzi L, Maiuri MC, Vitale I, Zischka H, Castedo M, Zitvogel L, et al. Cell death modalities: classification and pathophysiological implications. *Cell Death Differ.* 2007;14:37–43.
84. Mukhtar MS, McCormack ME, Argueso CT, Pajerowska-Mukhtar KM. Pathogen Tactics to Manipulate Plant Cell Death. *Curr Biol.* 2016;183–191.
85. Renehan AG, Booth C, Potten CS. What is apoptosis, and why is it important? *BMJ.* 2001;322:36–48.
86. Kerr JFR, Wyllie AH, Currie AR. Apoptosis: A Basic Biological Phenomenon with Wideranging Implications in Tissue Kinetics. *Br J Cancer.* 1972;26:39–57.
87. Galluzzi L, Vitale I, Abrams JM, Alnemri ES, Baehrecke EH, Blagosklonny MV, et al. Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death Differ.* 2012;19:107–120.
88. Green DR, Llamby F. Cell Death Signaling. *Cold Spring Harb Perspect Biol.* 2015;7:153-161.
89. Kroemer G, Mariño G, Levine B. Autophagy and the Integrated Stress Response. *Mol Cell.* 2010;40:80–93.
90. Takeshige K, Baba M, Tsuboi S, Noda T, Ohsumi Y. Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *J Cell Biol.* 1992;119:1–11.
91. Mizushima N. Autophagy in Protein and Organelle Turnover. *Cold Spring Harb Symp Quant Biol.* 2011;76:97–106.
92. Münz C. Autophagy Beyond Intracellular MHC Class II Antigen Presentation. *Trends Immunol.* 2016;37:55–63.
93. Levine B, Kroemer G. Autophagy in the Pathogenesis of Disease. *Cell.* 2008;132:27–42.
94. Arico S, Petiot A, Bauvy C, Dubbelhuis PF, Meijer AJ, Codogno P, et al. The Tumor Suppressor PTEN Positively Regulates Macroautophagy by Inhibiting the Phosphatidylinositol 3-Kinase/Protein Kinase B Pathway. *J Biol Chem.* 2001;276:43–56.
95. Harris J. Autophagy and cytokines. *Cytokine.* 2011;56:140–154.
96. Chuang Y-C, Su W-H, Lei H-Y, Lin Y-S, Liu H-S, Chang C-P, et al. Macrophage Migration Inhibitory Factor Induces Autophagy via Reactive Oxygen Species Generation. *PLOS ONE.* 2012;7:14-26.
97. Harris J, Deen N, Zamani S, Hasnat MA. Mitophagy and the release of inflammatory cytokines. *Mitochondrion.* 2018;41:2–8.
98. Vanlangenakker N, Berghe TV, Vandenabeele P. Many stimuli pull the necrotic trigger, an overview. *Cell Death Differ.* 2012;19:75–86.

99. Gudipaty SA, Conner CM, Rosenblatt J, Montell DJ. Unconventional Ways to Live and Die: Cell Death and Survival in Development, Homeostasis, and Disease. *Annu Rev Cell Dev Biol.* 2018;34:311–332.
100. Kaczmarek A, Vandenabeele P, Krysko DV. Necroptosis: The Release of Damage-Associated Molecular Patterns and Its Physiological Relevance. *Immunity.* 2013;38:29–33.
101. Fuchslocher Chico J, Saggau C, Adam D. Proteolytic control of regulated necrosis. *Biochim Biophys Acta BBA - Mol Cell Res.* 2017;1864:47–61.
102. Sparkes A, De Baetselier P, Roelants K, De Trez C, Magez S, Van Ginderachter JA, et al. Reprint of: The non-mammalian MIF superfamily. *Immunobiology.* 2017;222:58–67.
103. Shen Y, Thompson DL, Kuah M-K, Wong K-L, Wu KL, Linn SA, et al. The cytokine macrophage migration inhibitory factor (MIF) acts as a neurotrophin in the developing inner ear of the zebrafish, *Danio rerio*. *Dev Biol.* 2012;363:84–94.
104. Oh M, Kasthuri SR, Wan Q, Bathige SDNK, Whang I, Lim B-S, et al. Characterization of MIF family proteins: MIF and DDT from rock bream, *Oplegnathus fasciatus*. *Fish Shellfish Immunol.* 2013;35:58–68.
105. Du J, Yu Y, Tu H, Chen H, Xie X, Mou C, et al. New insights on macrophage migration inhibitory factor: Based on molecular and functional analysis of its homologue of Chinese amphioxus. *Mol Immunol.* 2006;43:83–88.
106. Jin H-J, Xiang L-X, Shao J-Z. Molecular cloning and identification of macrophage migration inhibitory factor (MIF) in teleost fish. *Dev Comp Immunol.* 2007;31:31–44.
107. Kim B-S, Pallua N, Bernhagen J, Bucala R. The macrophage migration inhibitory factor protein superfamily in obesity and wound repair. *Exp Mol Med.* 2015;47:161-171.
108. Kim S, Cox CM, Jenkins MC, Fetterer RH, Miska KB, Dalloul RA. Both host and parasite MIF molecules bind to chicken macrophages via CD74 surface receptor. *Dev Comp Immunol.* 2014;47:19–26.
109. Qiu R, Li J, Xiao Z, Sun L. Macrophage migration inhibitory factor of *Sciaenops ocellatus* regulates immune cell trafficking and is involved in pathogen-induced immune response. *Dev Comp Immunol.* 2013;40:32–39.
110. Suzuki M, Takamura Y, Maéno M, Tochinali S, Iyaguchi D, Tanaka I, et al. *Xenopus laevis* Macrophage Migration Inhibitory Factor Is Essential for Axis Formation and Neural Development. *J Biol Chem.* 2004;279:16–24.
111. Wang B, Zhang Z, Wang Y, Zou Z, Wang G, Wang S, et al. Molecular cloning and characterization of macrophage migration inhibitory factor from small abalone *Haliotis diversicolor supertexta*. *Fish Shellfish Immunol.* 2009;27:57–64.
112. Kamir D, Zierow S, Leng L, Cho Y, Diaz Y, Griffith J, et al. A *Leishmania* Ortholog of Macrophage Migration Inhibitory Factor Modulates Host Macrophage Responses. *J Immunol.* 2008;180:50–61.

113. Miska KB, Fetterer RH, Lillehoj HS, Jenkins MC, Allen PC, Harper SB. Characterisation of macrophage migration inhibitory factor from *Eimeria* species infectious to chickens. *Mol Biochem Parasitol.* 2007;151:173–183.
114. Sommerville C, Richardson JM, Williams RAM, Mottram JC, Roberts CW, Alexander J, et al. Biochemical and Immunological Characterization of *Toxoplasma gondii* Macrophage Migration Inhibitory Factor. *J Biol Chem.* 2013;288:33–41.
115. Ajonina-Ekoti I, Kurosinski MA, Younis AE, Ndjonka D, Tanyi MK, Achukwi M, et al. Comparative analysis of macrophage migration inhibitory factors (MIFs) from the parasitic nematode *Onchocerca volvulus* and the free-living nematode *Caenorhabditis elegans*. *Parasitol Res.* 2013;112:35–46.
116. Chauhan N, Sharma R, Hoti SL. Identification and biochemical characterization of macrophage migration inhibitory factor-2 (MIF-2) homologue of human lymphatic filarial parasite, *Wuchereria bancrofti*. *Acta Trop.* 2015;142:71–88.
117. Liu J, Shao D, Lin Y, Luo M, Wang Z, Yao M, et al. PyMIF enhances the inflammatory response in a rodent model by stimulating CD11b+Ly6C+ cells accumulation in spleen. *Parasite Immunol.* 2016;38:77–83.
118. Miller JL, Harupa A, Kappe SHI, Mikolajczak SA. *Plasmodium yoelii* Macrophage Migration Inhibitory Factor Is Necessary for Efficient Liver-Stage Development. *Infect Immun.* 2012;80:39–47.
119. Park SK, Cho MK, Park H-K, Lee KH, Lee SJ, Choi SH, et al. Macrophage Migration Inhibitory Factor Homologs of *Anisakis simplex* Suppress Th2 Response in Allergic Airway Inflammation Model via CD4+CD25+Foxp3+ T Cell Recruitment. *J Immunol.* 2009;182:7–14.
120. Prieto-Lafuente L, Gregory WF, Allen JE, Maizels RM. MIF homologues from a filarial nematode parasite synergize with IL-4 to induce alternative activation of host macrophages. *J Leukoc Biol.* 2009;85:44–54.
121. Blackman R L., Eastop VF. *Aphids on the World's Crops: An Identification and Information Guide*, 2nd Edition. WILEY; 2000.
122. Giordanengo P, Brunissen L, Rusterucci C, Vincent C, van Bel A, Dinant S, et al. Compatible plant-aphid interactions: How aphids manipulate plant responses. *C R Biol.* 2010;333:16–23.
123. Jaouannet M, Rodriguez PA, Thorpe P, Lenoir CJG, MacLeod R, Escudero-Martinez C, et al. Plant immunity in plant–aphid interactions. *Front Plant Sci.* 2014;5:23-32.
124. Elzinga DA, De Vos M, Jander G. Suppression of plant defenses by a *Myzus persicae* (green peach aphid) salivary effector protein. *Mol Plant-Microbe Interact.* 2014;27:47–56.
125. Bos JIB, Prince D, Pitino M, Maffei ME, Win J, Hogenhout SA. A Functional Genomics Approach Identifies Candidate Effectors from the Aphid Species *Myzus persicae* (Green Peach Aphid). *PLOS Genet.* 2010;6:145-156

126. Rodriguez PA, Stam R, Warbroek T, Bos JIB. Mp10 and Mp42 from the Aphid Species *Myzus persicae* Trigger Plant Defenses in *Nicotiana benthamiana* Through Different Activities. *Mol Plant Microbe Interact.* 2013;27:30–39.
127. Genome Sequence of the Pea Aphid *Acyrtosiphon pisum*. *PLoS Biol.* 2010;8:45-53.
128. Dubreuil G, Deleury E, Crochard D, Simon J-C, Coustau C. Diversification of MIF immune regulators in aphids: link with agonistic and antagonistic interactions. *BMC Genomics.* 2014;15:13-23.
129. Peyraud R, Dubiella U, Barbacci A, Genin S, Raffaele S, Roby D. Advances on plant–pathogen interactions from molecular toward systems biology perspectives. *Plant J.* 2017;90:20–37.
130. Sparkes A, De Baetselier P, Roelants K, De Trez C, Magez S, Van Ginderachter JA, et al. The non-mammalian MIF superfamily. *Immunobiology.* 2017;222:73–82.
131. Vermeire JJ, Cho Y, Lolis E, Bucala R, Cappello M. Orthologs of macrophage migration inhibitory factor from parasitic nematodes. *Trends Parasitol.* 2008;24:55–63.
132. Holowka T, Castilho TM, Garcia AB, Sun T, McMahon-Pratt D, Bucala R. Leishmania-encoded orthologs of macrophage migration inhibitory factor regulate host immunity to promote parasite persistence. *FASEB J.* 2016;30:49–65.
133. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics.* 2014;30:36–40.
134. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* 2014;42:43–48.
135. Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, et al. Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Res.* 2016;44:74–80.
136. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2012;40:13–25.
137. Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* 2014;42:26–31.
138. Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, Harb OS, et al. FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Res.* 2012;40:75–81.
139. Attrill H, Falls K, Goodman JL, Millburn GH, Antonazzo G, Rey AJ, et al. FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*. *Nucleic Acids Res.* 2016;44:86–92.
140. Gauthier J-P, Legeai F, Zasadzinski A, Rispe C, Tagu D. AphidBase: a database for aphid genomic resources. *Bioinformatics.* 2007;23:83–94.

141. Logan-Klumpler FJ, De Silva N, Boehme U, Rogers MB, Velarde G, McQuillan JA, et al. GeneDB—an annotation database for pathogens. *Nucleic Acids Res.* 2012;40 Database issue:98–108.
142. Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, et al. Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine. *Nucleic Acids Res.* 2016;44:93–100.
143. Howe KL, Bolt BJ, Cain S, Chan J, Chen WJ, Davis P, et al. WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res.* 2016;44:74–80.
144. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite – a comprehensive resource for helminth genomics. *Mol Biochem Parasitol.* 2017;215:2-10.
145. Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, et al. Cyanophora paradoxa Genome Elucidates Origin of Photosynthesis in Algae and Plants. *Science.* 2012;335:43–47.
146. Heiges M, Wang H, Robinson E, Aurrecochea C, Gao X, Kaluskar N, et al. CryptoDB: a Cryptosporidium bioinformatics resource update. *Nucleic Acids Res.* 2006;34:19-22
147. Blanc-Mathieu R, Perfus-Barbeoch L, Aury J-M, Rocha MD, Gouzy J, Sallet E, et al. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLOS Genet.* 2017;13:77-80.
148. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:92–107.
149. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 2002;30:59-68.
150. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 2009;25:89–91.
151. Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics.* 2012;28:23–39.
152. Pond SLK, Frost SDW. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics.* 2005;21:25-31.
153. Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res.* 2007;35:11-21.
154. Pond K, L S, Frost SDW. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Mol Biol Evol.* 2005;22:12–22.
155. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Pond SLK. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLOS Genet.* 2012;8:57-64.
156. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10:45–58.

157. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014;45:52.
158. Bienert S, Waterhouse A, de Beer TAP, Tauriello G, Studer G, Bordoli L, et al. The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Res.* 2017;45:13–29.
159. Rodriguez PA, Bos JIB. Toward Understanding the Role of Aphid Effectors in Plant Infestation. *Mol Plant Microbe Interact.* 2012;26:25–30.
160. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep.* 2017;7:80-104.
161. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective - Wiley Online Library. 2018;12:123-129.
162. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 2008;9:40-54.
163. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc.* 2010;5:25–38.
164. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods.* 2015;12:7–9.
165. Mills JE, Dean PM. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J Comput Aided Mol Des.* 1996;10:17–22.
166. Hartley JL, Temple GF, Brasch MA. DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Res.* 2000;10:88–95.
167. Burki F. The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. *Cold Spring Harb Perspect Biol.* 2014;12:156-161.
168. Holterman M, Karegar A, Mooijman P, van Megen H, van den Elsen S, Vervoort MTW, et al. Disparate gain and loss of parasitic abilities among nematode lineages. *PLoS ONE.* 2017;12:14-21
169. Blaxter M, Koutsovoulos G. The evolution of parasitism in Nematoda. *Parasitology.* 2015;142:26–39.
170. Dieterich C, Sommer RJ. How to become a parasite – lessons from the genomes of nematodes. *Trends Genet.* 2009;25:13–19.
171. Giorgi CD, Veronico P, Luca FD, Natilla A, Lanave C, Pesole G. Structural and evolutionary analysis of the ribosomal genes of the parasitic nematode *Meloidogyne artiellia* suggests its ancient origin. *Mol Biochem Parasitol.* 2002;124:91–104.
172. Caspermeyer J. Next Generation TimeTree: An Expanded History of Life on Earth at Your Fingertips. *Mol Biol Evol.* 2017;34:22–33.

173. Mitreva M, Blaxter ML, Bird DM, McCarter JP. Comparative genomics of nematodes. *Trends Genet TIG*. 2005;21:73–81.
174. Trautwein MD, Wiegmann BM, Beutel R, Kjer KM, Yeates DK. Advances in Insect Phylogeny at the Dawn of the Postgenomic Era. *Annu Rev Entomol*. 2011;57:49–68.
175. Chauhan N, Hoti SL. Role of cysteine-58 and cysteine-95 residues in the thiol di-sulfide oxidoreductase activity of Macrophage Migration Inhibitory Factor-2 of *Wuchereria bancrofti*. *Acta Trop*. 2016;153:14–20.
176. Alam A, Goyal M, Iqbal MS, Bindu S, Dey S, Pal C, et al. Cysteine-3 and cysteine-4 are essential for the thioredoxin-like oxidoreductase and antioxidant activities of *Plasmodium falciparum* macrophage migration inhibitory factor. *Free Radic Biol Med*. 2011;50:59–68.
177. Chanteau S, Glaziou P, Plichart C, Luquiaud P, Moulia-Pelat JP, N’Guyen L, et al. *Wuchereria bancrofti* filariasis in French Polynesia: Age-specific patterns of microfilaremia, circulating antigen, and specific IgG and IgG4 responses according to transmission level. *Int J Parasitol*. 1995;25:81–95.
178. Zhou B, Zeng L. Immunity-Associated Programmed Cell Death as a Tool for the Identification of Genes Essential for Plant Innate Immunity. Springer New York; 2018. 51–63.
179. Ronald PC, Salmeron JM, Carland FM, Staskawicz BJ. The cloned avirulence gene *avrPto* induces disease resistance in tomato cultivars containing the *Pto* resistance gene. *J Bacteriol*. 1992;11:164:174.
180. Gimenez-Ibanez S, Solano R. Nuclear jasmonate and salicylate signaling and crosstalk in defense against pathogens. *Front Plant Sci*. 2013;4:123-127
181. Dios A, Mitchell RA, Aljabari B, Lubetsky J, O’Connor K, Liao H, et al. Inhibition of MIF Bioactivity by Rational Design of Pharmacological Inhibitors of MIF Tautomerase Activity. *J Med Chem*. 2002;45:2-16.
182. Dziedzic P, Cisneros JA, Robertson MJ, Hare AA, Danford NE, Baxter RHG, et al. Design, Synthesis, and Protein Crystallography of Biaryltriazoles as Potent Tautomerase Inhibitors of Macrophage Migration Inhibitory Factor. *J Am Chem Soc*. 2015;137:29-33.
183. Wells L, Whelan SA, Hart GW. O-GlcNAc: a regulatory post-translational modification. *Biochem Biophys Res Commun*. 2003;302:35–41.
184. Coll NS, Epple P, Dangl JL. Programmed cell death in the plant immune system. *Cell Death Differ*. 2011;18:47–56.
185. Elmore S. Apoptosis: A Review of Programmed Cell Death. *Toxicol Pathol*. 2007;35:495–516.
186. Meng X, Zhang S. MAPK Cascades in Plant Disease Resistance Signaling. *Annu Rev Phytopathol*. 2013;51:45–66.
187. Bernhagen J, Calandra T, Mitchell RA, Martin SB, Tracey KJ, Voelter W, et al. MIF is a pituitary-derived cytokine that potentiates lethal endotoxaemia. *Nature*. 1993;365:56–59.

188. Morand EF, Leech M, Bernhagen J. MIF: a new cytokine link between rheumatoid arthritis and atherosclerosis. *Nat Rev Drug Discov.* 2006;5:399–411.
189. Pawig L, Klasen C, Weber C, Bernhagen J, Noels H. Diversity and Inter-Connections in the CXCR4 Chemokine Receptor/Ligand Family: Molecular Perspectives. *Front Immunol.* 2015;6:19-31.
190. Tilstam PV, Qi D, Leng L, Young L, Bucala R. MIF family cytokines in cardiovascular diseases and prospects for precision-based therapeutics. *Expert Opin Ther Targets.* 2017;21:71–83.
191. Crichlow GV, Fan C, Keeler C, Hodsdon M, Lolis EJ. Structural Interactions Dictate the Kinetics of Macrophage Migration Inhibitory Factor Inhibition by Different Cancer-Preventive Isothiocyanates. *Biochemistry (Mosc).* 2012;51:1–14.
192. Thomma BPHJ, Eggermont K, Penninckx IAMA, Mauch-Mani B, Vogelsang R, Cammue BPA, et al. Separate jasmonate-dependent and salicylate-dependent defense-response pathways in *Arabidopsis* are essential for resistance to distinct microbial pathogens. *Proc Natl Acad Sci U S A.* 1998;95:107-111.
193. Roux BL, Rouanet H. Geometric Data Analysis: From Correspondence Analysis to Structured Data Analysis. Springer Science & Business Media; 2004;78:45-53.
194. Dos Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PCJ, Yang Z. Uncertainty in the Timing of Origin of Animals and the Limits of Precision in Molecular Timescales. *Curr Biol.* 2015;25:39–50.
195. Matari NH, Blair JE. A multilocus timescale for oomycete evolution estimated under three distinct molecular clock models. *BMC Evol Biol.* 2014;14:101-115.
196. Kim H, Lee S, Jang Y. Macroevolutionary Patterns in the Aphidini Aphids (Hemiptera: Aphididae): Diversification, Host Association, and Biogeographic Origins. *Plos one.* 2011;6:36-52
197. Lacy M, Kontos C, Brandhofer M, Hille K, Gröning S, Sinitski D, et al. Identification of an Arg-Leu-Arg tripeptide that contributes to the binding interface between the cytokine MIF and the chemokine receptor CXCR4. *Sci Rep.* 2018;8:51-71.
198. Pantouris G, Ho J, Shah D, Syed MA, Leng L, Bhandari V, et al. Nanosecond Dynamics Regulate the MIF-Induced Activity of CD74. *Angew Chem Int Ed.* 2018;57:16–29.
199. Merk M, Zierow S, Leng L, Das R, Du X, Schulte W, et al. The D-dopachrome tautomerase (DDT) gene product is a cytokine and functional homolog of macrophage migration inhibitory factor (MIF). *Proc Natl Acad Sci U S A.* 2011;108:77–85.
200. Parisi M-G, Toubiana M, Mangano V, Parrinello N, Cammarata M, Roch P. MIF from mussel: Coding sequence, phylogeny, polymorphism, 3D model and regulation of expression. *Dev Comp Immunol.* 2012;36:88–96.
201. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–410.

202. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, et al. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* 2017;45:190–209.
203. Rahlf T. *Data Visualisation with R: 100 Examples.* Springer International Publishing; 2017.
204. Madsen H, Thyregod P. *Introduction to General and Generalized Linear Models.* CRC Press. Taylor and Francis; 2010;4:125-131.
205. Burnham KP, Anderson DR, Huyvaert KP. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav Ecol Sociobiol.* 2011;65:23–35.
206. Neerchal NK, More JG, Huang X, Moluh A. A Stepwise Algorithm for Generalized Linear Mixed Models. :18;22–34.
207. Dray S, Dufour AB. The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw.* 2007;22:111–120.
208. Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. *Bioinformatics.* 2004;20:26–37.
209. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25:72–83.
210. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 2003;52:96–104.
211. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinforma Oxf Engl.* 2001;17:754–765.
212. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics.* 2011;27:64–75.
213. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol.* 2010;59:7–21.
214. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol Evol.* 2010;27:21–34.
215. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics.* 2003;19:72–84.
216. Posada D. jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol.* 2008;25:53–66.
217. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 2006;34:15-25.

Analyse évolutive et fonctionnelle des « Macrophage Migration Inhibitory Factor » chez les eucaryotes

Les cytokines MIFs (Macrophage Migration Inhibitory Factor) sont des protéines multifonctionnelles qui, chez les mammifères, interviennent dans plusieurs processus majeurs tels que le contrôle du cycle et de la mobilité cellulaire, l'activation de la réponse immunitaire et l'inhibition de l'apoptose. Des travaux récents montrent que les protéines MIFs peuvent également jouer un rôle majeur dans l'immunité des invertébrés, et être utilisées par des organismes parasites d'animaux ou de végétaux pour inhiber les défenses de leurs hôtes respectifs, ce qui soulève la question de leur diversité, de leur histoire évolutive et des potentielles différences fonctionnelles.

L'objectif général de ce travail de thèse était d'explorer la diversité et l'histoire évolutive des protéines MIFs à une échelle trans-règne, puis de rechercher leurs éventuelles différences fonctionnelles, en se focalisant sur les systèmes plantes-pathogènes.

Nous avons tout d'abord identifié les MIFs chez 803 espèces de plantes, champignons, protistes, et métazoaires, et analysé leur présence/absence et histoire évolutive en fonction des taxa, de l'écologie et du mode de vie (libre ou parasitaire) des espèces. Nous avons montré que l'histoire évolutive des MIFs, chez les eucaryotes, est complexe et implique des duplications ancestrales ainsi que des pertes multiples ou des re-duplications récentes. Les plantes (espèces libres autotrophes) et les parasites de plantes (autres que champignons) possèdent un nombre médian de trois MIFs, alors que les espèces hétérotrophes et les parasites d'animaux ont un nombre de MIF plus faible et/ou plus variable. De plus, les protéines MIFs semblent essentielles et fortement conservées, avec de nombreux résidus sous sélection purifiante, chez certains groupes comme les plantes, alors que dans d'autres groupes, elles semblent facultatives (e.g. champignons) ou présentes en plusieurs copies divergentes (e.g. nématodes, insectes), ce qui suggère de potentielles néofonctionnalisations.

Nous avons ensuite analysé l'effet des protéines MIFs de plusieurs espèces sur la mort cellulaire en système végétal. Tous les organismes testés (plantes oomycètes, protozoaires, insectes et nématodes), y compris ceux n'ayant pas d'interaction avec les plantes, possèdent au moins une protéine MIF capable d'inhiber cette mort cellulaire. Cela suggère que l'inhibition de la mort cellulaire en plante ne correspond pas à une néofonctionnalisation des MIFs de parasites de plantes, mais serait liée à des propriétés structurales et conservées des MIFs. Toutefois, aucun des paramètres étudiés (localisation subcellulaire) ou prédits *in silico* (présence de motifs, structures 3D, oligomérisation, modifications post-traductionnelles) ne semble lié à cette activité d'inhibition de la mort cellulaire. De futures études fonctionnelles poussées sont nécessaires à l'élucidation des relations structure/fonction de ces protéines complexes.

Mots clé: Macrophage Migration Inhibitory factor; reconstructions phylogénétiques, eucaryotes, Inhibition de mort cellulaire

Evolutionary and functional analysis of MIF in eukaryotes.

Macrophage migration inhibitory factors (MIF) are multifunctional proteins regulating major processes in mammals, including control of the cell cycle and migration, activation of innate immune responses, and prevention of p53-mediated apoptosis. MIF proteins also play a role in innate immunity of invertebrate organisms or serve as virulence factors in parasitic organisms, raising the question of their evolutionary history and of a putative differential evolution of structure/function relationships.

The general aim of this PhD was to explore the diversity and evolutionary history of MIF proteins across kingdoms, and to investigate their potential functional differences, with a special emphasis on host-parasite systems.

We first performed a broad survey of MIF presence or absence and evolutionary relationships across 803 species of plants, fungi, protists, and animals, and explored a potential relation with the taxonomic status, the ecology, and the lifestyle of individual species. We show that MIF evolutionary history in eukaryotes is complex, involving ancestral duplications, multiple gene losses and recent clade-specific re-duplications. Of note, plants and plant parasites (other than fungi) harbour a median number of three MIFs, while heterotrophic and animal parasite species harbour a lower or/and variable MIF number. Intriguingly, MIFs seem to be essential and highly conserved with many sites under purifying selection in some kingdoms (e.g. plants), while in other kingdoms they appear more dispensable (e.g. in fungi) or present in several diverged variants (e.g. insects, nematodes), suggesting potential neofunctionalizations within the protein superfamily.

We then analysed the effect of MIF proteins from selected species on plant cell death. All organisms tested (plant, oomycetes, protozoa, insects, and nematodes) including species that are not in interaction with plants, possess at least one MIF protein showing a significant cell death inhibitory effect. This suggests that plant cell death inhibition does not result from a neofunctionalization of MIF from plant-parasites, and is related to conserved structural features of MIF proteins. However, none of the parameters predicted *in silico* (sequence motifs, 3D structures, oligomerization, post-translational modifications) appeared to be related to the cell death inhibitory activity. Future extensive functional studies are necessary to unravel the structure-function relationship of these evolutionarily and functionally complex proteins.

Key words: Macrophage Migration Inhibitory factor; phylogenetic reconstructions, eukaryotes, cell death inhibition